A New HMM Learning Algorithm for Event Studies: Empirical Evidence from the French Stock Market

Mohamed Saidane* and Christian Lavergne

I3M, Department of Mathematics, University of Montpellier II
Place Eugène Bataillon CC 051 – 34095 – FRANCE

Abstract

In this article we propose a new approach in event studies based on a hidden Markov chain combined with a classical event study model. The number of states inform us about the number of significant events affecting the related market and the identification of the hidden states determines exactly the delimiters of the event period. Studying each state parameters allows us to examine the events effect on the related market and to compare results to traditional event analysis. Extensive Monte Carlo simulations and preliminary examination of real data in the French stock market show promising results.

Keywords: Constant Mean Return Model, EM Algorithm, Event study, HMM, Market Model, Model selection.

1 Introduction

The event study has been applied to economy-wide events in finance to test either if financial markets are efficient, or if (presuming markets are efficient) there is news provided to market participants given by certain events.

The classical event study proceeds from the known event to measure its effects on the stock price. It starts with the selection of exactly what is the event of interest. However, the market reaction usually is better captured in price movements over time and not on a single event day. Thus, the duration of event-period should be determined first. The determination of the starting point of event period, or the endpoint of estimation period, have to be carefully chosen so that the estimation period will not contaminate any potential effects due to the event. In fact, the initial step of the analysis involves determining what constitutes the event of interest, when it is considered to happen, how long it takes to occur, and finally the length of the "event window", the period the asset will be observed after the event occurs. Actually, in empirical studies, less effort is devoted to this step and it is generally made intuitively.

Our approach proceeds differently: it seeks on one hand, to determine whether significant events have occurred anywhere in a given set of time series. On the other hand, it seeks to fix accurately the starting and the endpoint of the event window. The core idea is that we model the event occurrence by a hidden Markov chain combined with a classical event study

*Correspondence to: Mohamed Saidane, Institut de Mathématiques et de Modélisation de Montpellier, Université Montpellier II, Place Eugène Bataillon CC 051 – 34095, France. Tel.: +33-6-8601-0239, Fax: +33-4-6714-3558, **E-mail**: saidane@math.univ-montp2.fr

model. The number of states inform us about the number of significant events affecting the related market and the identification of the hidden states determines exactly the delimiters of the event period. Finally, studying each state parameters allows us to examine the events effect on the related market and to compare results to traditional event analysis.

The paper is organized as follows. In the second section, the classical methodology is presented. We give an overview on its steps and we criticize some aspects to which we will try to provide a solution. In the third section of the paper, we present the statistical framework of statistical models and algorithms we will apply next to resolve our proposed model. The fourth section consists in describing the methodology we propose and its implementation steps tested on simulated data in section 5. In section 6, we apply our proposed model to an examination of real data in the French stock market.

2 Some of the main results

According to the state of the art provided by Mackinlay (1997), the event study beginning may be assigned to Dolley (1933) who studied the price effects of stock splits. Mayers and Bakay (1948), then Baker (1956, 1957, 1958) and Ashley (1962) worked on the same subject and added the elimination of the general price movements impact on the market and separating the confused events.

In the late 1960's, two studies of serendipity made event studies an important reference tool in finance, especially corporate finance. The first is the Ball and Brown (1968) paper where the authors considered the information content of earnings. The second is the Fama, Fisher, Jensen and Roll (1969) paper that examines the effects of stock splits on firm returns after removing the effects of simultaneous dividend increases. These two studies introduced the Cumulated Abnormal Return methodology (CAR) based on the Market Model.

After these pioneer works, applications have been so numerous that it would be impractical to try to list them exhaustively. In fact, the original (Fama, Fisher, Jensen and Roll) methodology has been improved in order to tackle numerous problems such as event clustering (Salinger, 1992), event-induced variance (Bohemer, Musumeci and Poulsen, 1991), and event-date uncertainty (Ball and Torous (1988) and Aktas, Bodt and Cousin (2003)).

2.1 Classical procedure for event study

Before presenting the classical procedure used in the event study, we should define some terminologies to set the stage for the event study in finance that is the principal focus of this paper.

2.1.1 Event definition

According to Grar (1997), an event is defined as an information which is made public on the market and which may affect the value of one or several firms at the same time. An event may be general or specific, periodical or occasional, exogenic or decided by the firm managers. As event examples, we note: stock splits, mergers and acquisitions, earning announcements, issues of new debt or equity, regulation change, and announcement of macroeconomic variables such as the trade deficit.

Event study methods exploit the fact that, given rationality in the marketplace, the effects of an event will be reflected immediately in security prices. Thus the impact can be measured by examining security prices surrounding the event. In general, a significant event affects the equilibrium to which is subjected the couple return-risk of a security.

2.1.2 Event window

The event window is the time span in which all relevant information needed to assess outcomes linked to the event being examined becomes available. Logically, the event window would begin on the event date, the date when the event is formally announced by the firm or other sources of information. However, news about events often becomes public knowledge even before the event date as details about the impending event are leaked by insiders, spread as rumors or are inferred through other means by investors in capital markets. The event period can be symmetrical or asymmetrical around the announcement day.

2.1.3 Estimation window

The estimation window, also called pre-event period, is a time span before the event period. During this period, the considered market is supposed to be in a normal situation which means that no events have occurred. The estimation period is the period which is used as the basis for estimating what the values of the observed time series during the event period would have been if the event had not occurred. It is typical for the estimation window and the event window not to overlap because including the event window in the estimation of the normal model parameters could lead to the event returns having a large influence on the normal return measure. However, the methodology is built around the assumption that the event impact is captured by the abnormal returns. This would be problematic especially with the uncertainty in fixing windows delimiters.

2.1.4 Post-event window

The post event window is a period chosen after the event period. It is a period where all statistically significant effects of the event on the stock price are presumed to be finished. It is a period where the asset will be observed after the event occurs. On occasion, the estimation period combines the pre-event and post-event windows in order to increase the robustness of of the normal market return measure to gradual changes in its parameters. Note that in the classical approach, the three time periods of analysis are common to all firms.

2.1.5 Classical methodology

1. **Time line definition**: Traditional event analysis starts with the selection of exactly what is the event of interest. This initial step of the study involves determining what constitutes the event of interest, when it is considered to happen, how long it takes to occur, and finally the length of the event, the pre-event and the post-event windows.

2. **Data sample determination**: The second step is the selection of the sample set of firms to include in the analysis. This selection includes the choice of an appropriate market reacting to the considered event, the firms sample is chosen in this market.

3. **Normal return measurement**: In the absence of the event and especially before the event, the return is considered as a normal return. The following step in the event study is of course the prediction of the normal return according to a definite model.

The model parameters are classically computed over the estimation window. To specify a model, a number of approaches are available. The approaches are grouped into two categories: statistical and economic (MacKinlay, 1997). Statistical models of returns are derived purely from statistical assumptions about the behavior of returns whereas but, economic models apply restrictions to a statistical model that result from assumptions about investor behavior motivated by theory (e.g., capital asset pricing model, arbitrage pricing theory).

In this framework, we present only the two most popular statistical models in practice: Constant Mean Return Model and Market Model. For these two models, we should firstly assume that returns on stocks are jointly multivariate normal and are *iid* distributed through time.

- **Constant Mean Return Model**

The constant mean return model (CMRM) relates the return of any given security to its previous average return.

Let μ_a be the mean return of asset a. The constant mean return model assumes that asset returns are given by:

$$R_{a,t} = \mu_a + \varepsilon_{a,t} \quad \text{with} \quad \varepsilon_{a,t} \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_a}\right) \tag{1}$$

where $R_{a,t}$ is the period-t return on security a and $\varepsilon_{a,t}$ is the time period t disturbance term for security a with an expectation of zero and variance $\sigma^2_{\varepsilon_a}$.

Brown and Warner (1980, 1985) find that, in spite of its simplicity, the constant mean returns model often yields results similar to those of more sophisticated models because the variance of abnormal returns is not reduced much by choosing a more sophisticated model.

- **Market Model**

The market model relates the return of any given security to the return of the market portfolio. The Market Model (MM) is the most frequently chosen model. Relying the event study on the market model was suggested by Fama, Fisher, Jensen and Roll (1969), perpetuated by Brown and Warner (1980), Brown and Warner (1985), Dodd and Warner (1983), Mackinlay (1997), Cable and Holland (1999) and used by Beitel, Schiereck and Wahrenburg (2003). It is written as follows

$$R_{a,t} = \mu_a + b_a R_{m,t} + \varepsilon_{a,t} \quad \text{with} \quad \varepsilon_{a,t} \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon_a}\right) \tag{2}$$

where $R_{a,t}$ and $R_{m,t}$ are the period-t returns on security a and the market portfolio, respectively, and $\varepsilon_{a,t}$ is the zero mean disturbance term. μ_a, b_a and $\varepsilon_{a,t}$ are the parameters of the MM.

Since the ideal market index does not exist and even the broadest defined market index does not perfectly represent all traded assets, only approximations of the market index are used. Usually a broad based stock index represents for the market portfolio. According to Mackinlay (1997) the market model represents a potential improvement over the constant mean return model, since by removing the portion of the return that is related to variation in the market's return, the variance of the abnormal return is reduced. McWilliams and Siegel (1997) appraise the market model as the best currently available model, although it has its shortcomings. Brown and Warner (1985) conclude that the market model is both well-specified and relatively powerful under a wide variety of conditions.

4. **Abnormal return measurement**: After calculating the normal return parameters, the following step is to appraise the abnormal return within the event window which reflects investor response to the event. The abnormal return is deduced from the normal return. So, it is defined as the difference between the actual and the predicted returns according to the normal return parameters. For asset a and event date t, the abnormal return is

$$AR_{a,t} = R_{a,t} - \mathbb{E}\left(R_{a,t}|I_{\tau}\right) \tag{3}$$

where $AR_{a,t}$, $R_{a,t}$ and $\mathbb{E}\left(R_{a,t}|I_{\tau}\right)$ are the abnormal, the actual, and normal returns respectively for time period t. I_{τ} is the conditioning information for the normal return model computed in the estimation period.

5. **Significance tests**: The last step in the event study is testing the significance of the abnormal return (checking whether the abnormal return is statistically different from zero) and checking whether the event impact is positive or negative. Generally, the cumulative abnormal return (CAR) is used instead of the abnormal return. The tests consist mainly on:

   - Testing the hypothesis that abnormal return equals zero for a single security,
   - Testing the hypothesis that the average abnormal return equals zero for many securities.

2.2 Criticized issues

We are interested in time windows uncertain choice and parameters evolution from a window to another for which we found many different hypothesis in the literature.

2.2.1 Dates uncertainty

1. **Event date uncertainty**: The event date is the date when the event information is made public (Grar, 1997). This date is assumed to be the date of the first announcement of the event. However, in some studies it may be difficult to identify the exact date. A common example is when collecting event dates from financial publications such as the Wall Street Journal because we cannot be certain if the market was informed prior to the close of the market the prior trading day. Thus, we aren't sure whether the current or the previous day is the event date.

   Foster (1973) reported that announcement of an "earnings estimate" by a company official effectively usurped the information content of the subsequent earnings announcement. Empirical studies suggest that event-date uncertainty affects the power of the tests which are designed to detect the presence of abnormal performance associated with an event, and that events, both treatment events and confounding events, can be hard to find and even harder to date. See, for example, Brown and Warner (1980, 1985).

2. **Event period identification**: Under theoretical assumptions of market efficiency and perfect information, the event period should be the event announcement day: the market reaction is supposed to be instantaneous and immediate.

   However, on one hand, the information may become known to a wide segment of the market prior to the first public announcement through a news leak, a rumors spread or it may be inferred through other means by investors in capital markets released in a form which effectively communicates the information but which is not considered to be a public announcement of the event itself. On the other hand, the market reaction usually is better captured in price movements over time and not on a single day. Thus, the duration of event-period may be greater than one day and may be different from a firm to another.

   Dyckman, Philbrick, and Stephan (1984) investigated event periods of 1 through 5 days in length. They reported that a longer event period should be used when the bounds of the uncertain period are known ex-ante. Brown and Warner (1985) reported that the power of statistical tests decreased with longer event periods, but that event study test statistics continued to be well specified when the event period was longer than one day. Brown, Lockwood, and Lummer (1985) suggested that the event period should be selected on a case-by-case basis, which suggests the use of some analytical method on which to base the selection. Event periods of various lengths are found in the literature but always made arbitrarily or intuitively.

3. **Estimation period identification**: Standard approach supposes that no events affect the market during the estimation period. Therefore, the determination of the starting point, or the endpoint of estimation period, must be carefully chosen so that the estimation period will not contaminate any potential effects due to any event. This hypothesis is so strong and rarely possible since defining a normal situation of the market is a difficult task.

   Moreover, the pre-event window is most often defined as a sufficiently long period preceding the event (between 250 and 30 days in case of daily data). Aktas, Bodt, and Cousin (2003) showed that this approach may generate a significant risk of bias for the analysis of specific kinds of corporate events.

The usual method used to cope with dates uncertainty is to expand the windows studied. Actually, studying the inferences with dates uncertainty has attached less interest in empirical studies. The authors are only aware of two papers dealing with this issue. The first paper is the Ball and Torous (1988) paper where they develop a maximum likelihood estimation procedure estimating, simultaneously, for each day of the event window, the abnormal returns, their variance and the probability of an event. The results indicate that such approach detects more frequently simultaneous abnormal returns than the classical approach. The second work is the Aktas, Bodt, and Cousin (2003) paper, in which a Markov Switching Regression test is proposed. This test divided the estimation period into two regimes (low variance for event period and high variance for non event periods) in order to distinguish an event period from a non event period.

2.2.2 Hypothesis on the model parameters

The second criticized issue concerns some hypothesis generally made for the model parameters in an event study. These hypothesis are criticized in many empirical studies but no definitive solution has been suggested so far.

To examine these hypothesis, let's firstly remember that according to the equation (3), the expectation and the variance of returns conditionally to the information available during the estimation period are given in the table below (table 1).

Table 1: Expectation and variance of normal and abnormal returns

	During the estimation period		During the event period	
	Normal return	Abormal return	Normal return	Abormal return
Expectation	$\mu_{a,t}$	0	$\mu_{a,t} + K_{a,t}$	$K_{a,t}$
Variance	$\sigma^2_{a,t}$	$\sigma^2_{a,t}$	$\eta^2_{a,t}$	$\eta^2_{a,t}$

In this presentation, $\mu_{a,t}$ refers to the expected return of asset a at the date t belonging to a normal period and $K_{a,t}$ is the jump reflected to the expected return during the event period. The event study is aimed at estimating the $K_{a,t}$ for each security a at each date t of the event period. The mean of these jumps provides an estimation of the mean market reaction to the related event.

1. **Event incidence on the variance**: Most of standard methodologies in event study apply tests based on a variance computed during the estimation period, so it is implicitly assumed that the variance doesn't change in the event period, which means according the previous presentation (table 1) that the following constraint is added to the model: $\eta^2_{a,t} = \sigma^2_{a,t} = \sigma^2_a$.

   However, Several papers have confirmed a temporary or persistent increase of the event returns variance (see for example Ball and Torous 1988). Therefore, the constraint $\eta^2_{a,t} = \sigma^2_{a,t}$ should be rejected. In fact, if the variance is understated, as has been found to be the normal case, the null hypothesis of zero abnormal returns will be rejected more frequently than it should, even though the average abnormal return is significantly close to zero. Hence, many studies tried to find abnormal performance caused by a failure to consider event induced variance and to find which methods provide the most accurate results under these conditions. The main results show that event-related variance increases cause standard parametric tests to report a price reaction where none actually exists more often than expected. That's why nonparametric tests were introduced since they do not use the return variance. Thus, they may perform better under variance increases than parametric tests that assume stable variances (see for example Arnold 1992). Whereas, few studies tried to consider the induced-variance effect and the solutions they proposed consists mainly in:

   • Introducing a model describing the dynamic variance during time over both of the periods, such as GARCH models to describe the error.

   • Considering a constant variance but submitted to a jump during the event period. For example, Ball and Torous (1988) presented the following model parameters:

	Estimation period	Event period
Expectation	μ_a	$\mu_a + \alpha_a \cdot \sigma_a$
Variance	σ_a^2	$\omega^2 \cdot \sigma_a^2$

2. **Event incidence on the market portfolio coefficient**: This issue concerns only the market model. In fact, according to the market model, a security return may be divided into a constant component (the mean), a risk premium proportional to the beta (the market portfolio coefficient denoted b in this work) of the security, and a random component of mean null. According to the previous presentation in table 1, only the constant component may be affected.

   However, some papers show that some events (such as stock splits) may cause an increase of the beta. This would be problematic because, even if some studies consider the increase in the variance, no study consider the increase in the beta value. Therefore, in the case of increasing beta, the mean return will be affected twice, directly by the constant increase and indirectly because of the risk premium. Thus, the abnormal return will be overestimated.

Two types of constraints was presented in this paragraph (constant variance and constant beta) but, in the literature we find more constraints concerning the model parameter evolution from a period to another, such as mean is constant, mean is null, parameters are the same during the pre-event and the post-event windows. These constraints are made in a case by case manner and aren't checked.

3 Proposed approach

In the previous section, we outlined some issues generally neglected in the empirical event studies. In fact, in this work, we will propose a new approach to deal with event study without neglecting these issues. As we presented in the previous section, the classical event study proceeds from the known event to measure its effects on the market. Our approach proceeds differently. First, it seeks to determine whether significant events have occurred anywhere in a given time series, to determine with certainty the delimiters of each period (pre-event, event and post-event periods). The periods identification will be made separately for each firm of the considered market. Then, our next goal is to choose the appropriate constraint model for the parameters estimation and finally to estimate the parameters.

The main idea of this approach is to consider that the market's reaction to the event marks a transition from a state to another state. In fact, our approach consists in the use of a mixed model combining:

1. a classical model used in event study (Constant mean return model or market model) in order to model the returns evolution during the event study time-line,

2. a hidden Markov chain to model the succession of the three time phases: the pre-event phase, the event phase and the post-event phase.

Our initial intuition is to consider the three periods defined in the event study time line as three states of a Markov chain. Identifying the hidden states allows us to find the event window $[T_1, T_2]$ giving the highest probability when assuming correctness of the classical model used. But this strategy will be applied in the case of three states Markov chain, so when we already know that a unique event has occurred in the time bar we consider.

Otherwise, if we consider that instead of having three states there are J states, determining the states number allows us to detect the number of events affecting each firm. In addition, determining the duration of each state allows us to correctly divide the time bar into different periods. Furthermore, determining the securities parameters (return and risk) in each state will allow us to examine the event impact on the securities behavior (definitive or temporary, see section 2.1). But, before the parameters estimation, the appropriate constraint model should be selected. In fact, the constraints models are a set of models to which we add constraints such as constant mean or variance. We consider also a no constraint model where we presume that all the model parameters may change from a state to another. The selection of the appropriate model will depends on the related data.

For modeling hypothesis, we firstly fix a sample of firms to study. So, we suppose that we have N assets belonging to the same market and being subjected to the same event announcements. We choose a classical model for normal returns that we combine to a hidden Markov chain. Our two classical models application requires to suppose that: returns on stocks are jointly multivariate normal and are *iid* distributed through time. For the market model, we suppose also that the securities returns are uncorrelated with the market portfolio.

Furthermore, we assume that the N assets have the same behavior during the three states, it means that we suppose that the N assets have the same mean and variance during each state but they do not need to pass by all the states, which means that an event is able to affect all the securities or just some of them and the event effect may be longer in a security than another one.

Doing our analysis requires firstly to apply model selection criteria in order to choose the appropriate model and the appropriate states number. Then, the estimation of the model parameters which will be done by the maximization of the likelihood function. The parameters will be computed with the Expectation-Maximization (EM) algorithm. Finally, restoring the different states will be done with the Viterbi (1967) algorithm.

4 Statistical models and parameters estimation

In this section we consider the same hypothesis of section 3. So, we have a market where there are N assets subjected to the same event announcements. We suppose that there are J events affecting this market. In the context of this work, the observed data is the observed returns of the N assets during a period of time that we will denote $[1, T]$. The hidden data is the states of the Markov chain defined as presented in section 3. So, we assume that the market reaction to each event marks a changement of the return behavior: a transition from a state to another.

The hidden Markov chain is defined by its transition matrix and its initial probabilities vector. The hypothesis of changing the state with market reaction to events implies that the properties of the considered hidden chain change over time: as time increases, the state index increases (or stay the same), no transition are allowed to states whose indices are lower than the current state. For states $J = 3$, the state transition probabilities have the properties:

$$\begin{cases} p_{ij} \geq 0 & j > i+1 \\ p_{ij} = 0 & j \leq i+1 \\ \sum_{j=1}^{J} p_{ij} = 1 \end{cases}$$

So, no jump for more than one state are allowed. The form of the transition matrix, in the case of three states, is thus:

$$A = \begin{bmatrix} p_{11} & p_{12} & 0 \\ 0 & p_{22} & p_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

furthermore, the initial transition probabilities have the property:

$$\pi_j = \begin{cases} 1, & j = 1 \\ 0, & j \neq 1 \end{cases}$$

since the state sequence must begin in state 1 and end in state J.

The market behavior is determined by the classical model parameters: mean and variance (respectively mean, variance and the market return coefficient) in the case of the CMRM (respectively MM). These parameters depend on the state.

Parameters will be estimated with the EM algorithm and the hidden states will be restored thanks to the Viterbi algorithm.

5 Implementation issues and simulation results

In this section, we deal with practical implementation issues including initial parameters estimates, model selection and choice of resolution algorithms that will be illustrated by results made on simulated data.

5.1 EM limitations

In spite of its elegance and simplicity, EM algorithm has two major limits:

1. Before running EM, we should know the exact number of states of the hidden Markov chain. For that, we can compare a set of models with different states number and then choose the best one describing our data sample. This problem is obviously similar to the problem of selecting the best constraint model which should also be resolved before estimating the model parameters.

2. EM solution depend highly on its starting position. Indeed, two cases are possible:
   - Convergence to a stationary point of the likelihood,
   - Crippling slow convergence because the starting points are far away true values.

   Therefore, it is straightforward that a good initialization process can help us in finding good results. A key question is therefore how do we choose initial estimates of our model so that the local maximum is the global maximum of the likelihood function.

5.2 Model selection

The model selection problem will be raised twice: firstly, for choosing the number of states in the hidden Markov chain and secondly, for selecting the best constraint model. We set eight constraints models for the constant mean return model and eleven for the market model. Therefore, comparing the adequacy of the possibly models may be done by computing a criterion for each model and comparing the criteria values. We have chosen to use two of the most popular criteria for comparing multiple models, taking both descriptive accuracy and parsimony into account: AIC (Akaike, 1974) and BIC (Schwarz, 1978). These criteria are based on a penalized likelihood computation.

5.3 Choice of the starting position

Many simple methods are proposed in the literature to choose sensible starting values for the EM algorithm to get maximum likelihood parameter estimation. However, none of the experimental strategies can be regarded as the best one and it is difficult to characterize situations where a particular strategy can be expected to outperform the other ones. In this work, we'll be interested in the initialization strategies proposed by Biernacki, Celeux, and Govaert (2002) where the authors experimented an efficient three step Search/Run/Select (S/R/S) strategy for maximizing the likelihood.

The Search step consists in generating n initial positions based on random starts or the output from an algorithm like a Classification EM (CEM) algorithm (see McLachlan and Peel 2000, Section 2.21), a Stochastic EM (SEM) algorithm (see Celeux and Diebolt, 1985) or short runs of the standard EM algorithm. The Run step consists in running the EM algorithm a set number of times at each initial position with a fixed number of iterations. The Selection step consists in selecting the solution providing the best likelihood among the n trials, say θ^*. This three-step strategy can be compounded by repeating the three steps x times and using the $\theta_1^*, ..., \theta_x^*$ as the starting positions in the first step.

In this paper we have used short runs of EM from random positions followed by a long run of EM from the solution maximizing the observed log-likelihood. Short EM runs provide a rapid initiating procedure visiting a large parameter space thanks to random positions it take. So, without consuming a large amount of CPU time, several short runs are performed before passing to EM with the best solution among them. By a short run of EM, it is meant that instead of waiting to convergence, the algorithm is stopped as soon as

$$\frac{l^k - l^{k-1}}{l^k - l^0} \leq s$$

l^k denotes the observed log-likelihood at *k-th* iteration. s represents a threshold value chosen in a pragmatic ground. In the constant mean return model, we choose it equal to 10^{-3} and in the market model, we assign it to 10^{-2}. Furthermore, x repetitions of the previous strategy lead to the so called "*xem*-EM" strategy. In *xem*-EM, the basic three steps are random search, short runs of EM, and selection, compounded by x outer loops with EM runs.

5.4 Model building procedure and simulation examples

As we stated earlier, given a sample set of N assets belonging to the same market and subjected to the same event announcements during a period time of length T, our goal is to detect the events affecting this sample and to examine their impact during time. To fulfill this objective, we choose a mixed model combining a standard return estimation model (the constant mean return model or the market model) and a hidden Markov chain. We raised the issue of estimating this model parameters and we presented the appropriate algorithms for it. Our purpose now consists in describing the algorithmic scheme we propose and studying its performance on simulated data.

The procedure we propose to follow is divided into four steps. In fact, our input is the data sample of N assets returns (and the market portfolio return in the case of the market model) during a period of time $[t_1, t_T]$. Firstly, we should define the number of states affecting the sample during this time period. Obviously, the number of states denotes the market regimes before, during and after the events. Then, knowing the states number, we select the best constraint model describing the related data sample. The next step is to estimate the parameters in each sate according to the selected constraint model. The mixed model parameters are Markov chain parameters: the transition probabilities matrix and the constant mean return model (or the market model) parameters: mean and variance (and the market portfolio coefficient). Finally, we restore the hidden states which allow us to find the transition dates from a state to another.

In the following subsections, we will give the detail of each step and illustrate it by some results found under simulated data. Indeed, we consider the example of three states model with a transition matrix

$$\begin{pmatrix} 0.9 & 0.1 & 0 \\ 0 & 0.85 & 0.15 \\ 0 & 0 & 1 \end{pmatrix}$$

We consider the same parameters for the constant mean return model and the market model: $\mu = [0.7; 0; 1]$ and $\sigma^2 = [0.3; 0.7; 0.5]$. For the market model, we add the market portfolio coefficient: $b = [2; 1.2; 2.4]$.

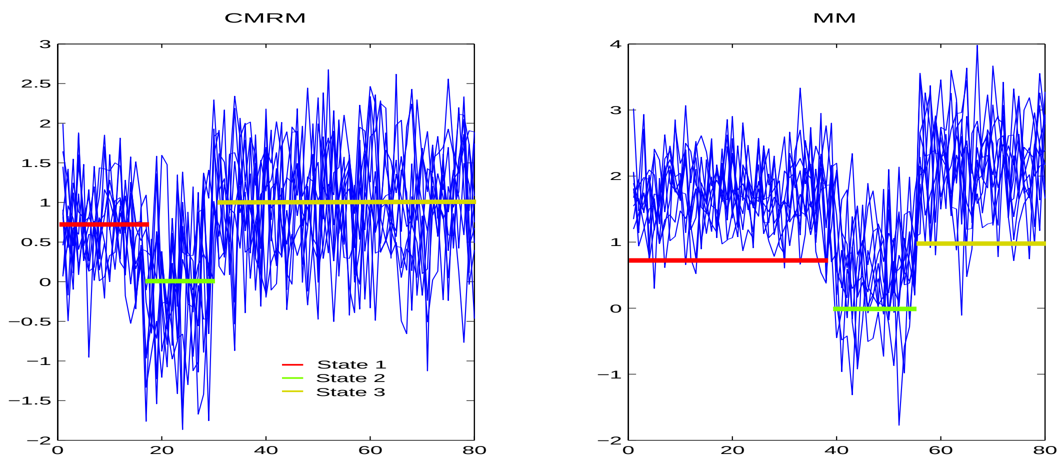

Figure 1: Simulated data for the constant mean return model and the market model.

5.4.1 Choice of the number of states

Description

To find the number of states during a given time period $[t_1, t_T]$, we use a selection criterion in order to choose between M models. We denote by *m-th* model, $m \in [1, M]$ the model of m states. We run the EM algorithm without constraints in order to maximize the log-likelihood of the complete data of each model $m \in [1, M]$. Minimizing the selection criteria computed after each EM running allows us to find the best model among the M models. M is firstly chosen randomly (or according to a prediction on the eventual events number). Then, if the best model is the *M-th* model, we run the no constraint EM algorithm again but for models with m states $m \geq M$ until we find that the best model is $m < M_{max}$. This strategy is the same for the constant mean return or the market model.

Numerical experiments

Simulations show BIC to perform better than AIC. The AIC criterion generally fails in finding the exact number of states. Whereas, the BIC criterion detect the exact number (see results in table 2).

According to the literature, such result is not surprising. On one hand, despite the widespread use of the AIC, some believe that it is too liberal and tends to select overly complex models (Kass and Raftery, 1995). It has been pointed out that the AIC neglects the sampling variability of the estimated parameters. When the likelihood values for these parameters are not highly concentrated around their maximum value, this can lead to overly optimistic assessments. Furthermore, the AIC is not consistent. That is, as the number of observations grows very large, the probability that the AIC recovers a true low dimensional model does not approach unity. On the other hand, a comparison of BIC to AIC shows that the BIC penalty term is larger than the AIC penalty term. The BIC assumes that the true generation model is in the set of candidate models, and it measures the degree of belief that a certain model is the true data-generating model. As we assume that the true model is in the candidate set and that it is relatively low dimensional, we favor BIC over AIC. Hence, only the BIC criterion will be used in finding the states number.

Table 2: Selection criteria values in searching the states number

Number of states	CMRM		MM	
	AIC	BIC	AIC	BIC
2	1.9371	1.9511	1.9730	2.0064
3	1.8899	**1.9458**	1.8027	**1.8552**
4	1.8833	1.9602	1.7870	1.8587
5	1.8784	1.9762	1.7882	1.8791
6	1.8793	1.9981	1.7986	1.9085
7	1.8771	2.0169	1.8034	1.9324
8	1.8786	2.0393	1.7864	1.9346
9	**1.8710**	2.0526	**1.7851**	1.9524
10	1.8788	2.0814	1.7860	1.9724

5.4.2 Choice of the best constraint model

The choice of the best constraint model is also a selection problem in which we will use the selection criteria AIC and BIC. The procedure consists in choosing the constraint model that leads to the lowest selection criterion value after several EM runnings.

Unlike the previous case, AIC and BIC criteria find easily the best constraint model and are equivalent. Table 3 gives an illustration of algorithms run on the simulated data (see figure 1).

Table 3: Selection criteria values in searching the best constraint model

	CMRM		MM	
	AIC	BIC	AIC	BIC
no constraints	**1.8859**	**1.9128**	**2.0298**	**2.0728**
mean is null	2.3217	2.3358	3.2550	3.2980
reduced variance	1.9997	2.0137	2.0960	2.1534
constant mean	2.0190	2.0378	2.2433	2.3271
constant variance	1.9052	1.9240	2.0361	2.1367
same mean for 2 states	1.8993	1.9227	2.0940	2.2282
same variance for 2 states	1.8893	1.9140	2.1362	2.2895
same parameters for 2 states	1.9031	1.9219	2.2470	2.3764
$b = 0$			2.3612	2.5050
$b =$ constant			2.1546	2.3391
same b for 2 states			2.1566	2.3867

5.4.3 Parameters estimation

The parameters estimation is the hardest part of the work since it is very sensible to the initialization parameters. The initialization issue was raised in section 5.3, where we anticipated the use of EM short runs. In this case, we fix a threshold of 10^{-3} for constant mean return model and 10^{-2} for the market model to get short runs in the order of 5 EM iterations per run. This procedure will be iterated about 30 times from different random positions.

After the initialization, the EM algorithm is run with a maximal iterations number of 1000 according to the convergence criterion given by $\mathcal{L}^{(k)} - \mathcal{L}^{(k-1)} \leq 10^{-16}$, where $\mathcal{L}^{(k)}$ is the log-likelihood in the k-th iteration. This procedure is iterated about 10 times where the parameters related to the greatest likelihood are kept. We should remember that the parameters estimation is made according to the number of states and the best constraint model found in the previous steps.

The numerical experiments show a good performance in estimating the parameters. By applying the methodology described in the previous paragraph.

5.4.4 Hidden states restoration

Restoring the hidden Markov chain is done by applying the Viterbi algorithm described in Rabiner (1989). Running the Viterbi algorithm requires the estimated parameters we found thanks to the previous implementation steps.

The numerical experiments show that, in general we restitute the exact hidden state we simulate. But, in some cases, we find an error in one or two dates.

6 Empirical study

In this section, we apply the methodology outlined previously to an investigation of the market reaction to some events. Indeed, our empirical study is based on the assets constituting the French CAC40 Index. We have chosen to analyze the impact of three types of events on the market and on the related stock. The events we consider are stock splits, earnings announcements and dividends distribution. We consider also a special event which is the eleven of September event. For each event, we use the CMRM and the MM.

This section will be divided into two subsections. In the first one, we display our findings for each event we consider. The second consists in giving some remarks and comparisons of the results.

6.1 Results

As we stated, the data sample we consider is the stocks composing the CAC40 Index. The events we consider are the following:

- Accor Stock splits on the 21-*st* December 1999
- LVMH dividend distribution on the 4-*th* December 2001
- Eleven of September event

The methodology is the methodology stated in the previous section. Furthermore, the stocks returns are computed according to the following formula

$$R_{a,t} = \frac{C_{a,t}}{C_{a,t-1}} - 1$$

with $C_{a,t}$ is stock price of asset a at the time period-t.

6.1.1 Accor Stock splits on the 21-*st* of December 1999

We know that Accor has announced a stock splits on the 21-*st* of December 1999. So, we firstly took our time series a month before and 20 days after this date. The a priori time line is going from the 24-*th* of September to the 10-*th* of January. The first findings are reported in the next paragraph for the CMRM and the MM.

Constant Mean Return Model

• Findings

Our first finding show two states in the considered market with a constant mean and a variance that doubles in the second state (see table 4). Nonetheless, some securities stay in the first state during all the time bar such as: Carrefour, Total, L'oreal, Suez, Sanofi Synthelabo, Axa, Danone, Sodexho Alliance, Pinault Printemps, Peugeot, Schneider Electric, Saint Gobain, Vinci, Casino Guichard, Agf, Aventis, Société Generale, BNP Paribas, Renault and Eads. Moreover, as it is shown in table 5, the transition date from the first to the second state varies from the 26/11/99 for TF1,Bouygues , France telecom, Thomson and ST microelectronics to the 07/01/00 for Sodexho Alliance.

• Interpretations

Before interpreting this result, we should remember that the considered time bar includes also earnings announcement dates. So, even if the event is detected, we can't be sure that

Table 4: CMRM: Properties of the model parameters associated to the Accor stock splits of the 21/12/99

Number of states: 2	
Constraint model: mean is constant	
$\mu_1 = 0.0030$	$\mu_2 = 0.0030$
$\sigma_1 = 0.0224$	$\sigma_2 = 0.0430$

the detected event is the stock splits. If we consider that variance increase corresponds to an event occurrence (as showed in Aktas, Bodt and Cousin, 2003)), we can say that the second state is an event window. But this event impact is seen prematurely on all the securities except Carrefour and Sodexho Alliance. Furthermore, there is no transition to the second state for the majority of the securities. We should mainly notice that Accor stays in the first state. So Two cases are possible:

1. The event is not detected by the security on which it occurred,
2. The event detection may be before the beginning of the considered time line or after its end. To check this eventuality, we should consider a larger time line.

Table 5: Accor stock splits of the 21/12/99: Transition dates from the first to the second state (CMRM)

26/11/99	TF1,Bouygues, France telecom, Thomson and ST microelectronics
08/12/99	Lagardere S.C.A
14/12/99	Pinault Printemps
15/12/99	Air liquide
17/12/99	Lvmh, Sanofi Synthelabo, Renault
04/01/00	Carrefour
07/01/00	Sodexho Alliance

Market Model

• Findings

The Market model allow us to recuperate also two states with an increase in the variance and a constant mean but which is null (see table 6). Nonetheless, there less securities that stay in the first state according to MM. They are: TF1, Bouygues, Accor, Lagardere S.C.A.,Thomson, Stmicroelectronics and Eads.

Moreover, as it is shown in table 7, the transition date from the first to the second state varies from the 24/11/99 for 19 assets to the 15/12/99 for Alcatel and France Telecom.

• Interpretations

To explain the difference in the number of states found by each model, we suggest two possibilities, either the Market Model detect the event with a delay if we compare it with the CMRM or the CMRM detect the event with an advance. In addition, we can affirm that the stock split event had no impact on the Accor return since as it is shown by the MM, the

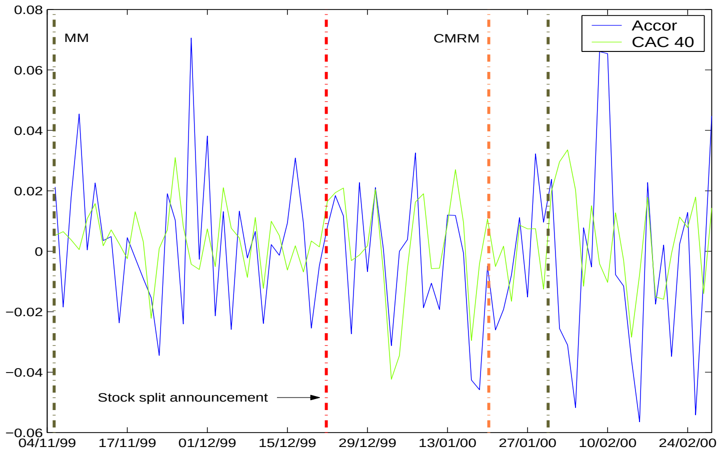

Figure 2: Windows delimiters for the Accor security between 04/11/99 and 29/02/00.

Table 6: MM: Properties of the model parameters associated to the Accor stock splits of the 21/12/99

Number of states: 2	
Constraint model: mean is null and b is constant	
$\mu_1 = 0$	$\mu_2 = 0$
$b_1 = 0.7775$	$b_2 = 0.7677$
$\sigma_1 = 0.0219$	$\sigma_2 = 0.0394$

Table 7: Accor stock splits of the 21/12/99: Transition dates from the first to the second state (MM)

24/11/99	Air Liquide, Total, L'oreal, Suez, Lafarge, Sanofi Synthelabo, Axa, Peugeot, Schneider Electric, Saint Gobain, Agf, Vivendi Universal, Pernod-Ricard, Lvmh, Sodexho Alliance, Michelin, Pinault Printemps, Société Générale and BNP Paribas
25/11/99	Danone
29/11/99	Vinci, Renault
30/11/99	Casino Guichard
02/12/99	Carrefour
03/12/99	Aventis
15/12/99	Alcatel, France Telecom

second state, which includes this event announcement date lasts three months (see figure 2), which is a long period for an event window. So, returns during this window are normal.

Conclusion

The first conclusion is that the Accor stock splits of the 21/12/99 hasn't affected the performance of the stock price Accor.

The second conclusion is that the CMRM and the MM provide different results (a delay

Table 8: Accor stock split of the 21/12/99, programs runned for Accor time series from 04/11/99 to 29/02/00

	CMRM		MM		
Number of states	2		3		
Constraint model	mean is null		mean is null		
Transition date	18/01/00		04/11/99	02/02/00	
μ	0	0	0	0	0
σ	0.0219	0.0314	0.0085	0.0201	0.0337
b			2.3770	0.5775	0.0773

is approximated to a month in the time windows definition). This result was explained in the literature (see Brown and Warner, 1980). When the abnormal performance is present, the differences between the methodologies based on CMRM and MM are quite small. Whereas, in spite of the good performance of the CMRM in many cases, it is highly sensitive to the events generation. For example, when there is an event month clustering, methodologies which incorporate information about the market's realized return perform substantially better than constant mean returns.

6.1.2 LVMH dividends distribution on the 4-*th* of December 2001

We know that Lvmh has announced a dividends distribution on the *4-th* of December 2001. Our study will be limited to apply the CMRM and the MM to the Lvmh return time series from the 01/11/01 to 28/02/02. The findings are reported in table 9.

• Findings

Both of the methodologies based on the CMRM and the MM find that there is only two states, the mean is constant (null in the MM) and the variance is higher in the second state (the CMRM variance is greater than the MM variance).

• Interpretations

Both of the methodologies based on the CMRM and the MM show that there is an abnormal return performance starting in the beginning of September. As the first state is characterized by a low variance in contrast with the second state, it is possible to confirm an event detection but we are not able to affirm that it is the dividend distribution since it may be an earnings announcement or an other event that affected the market. To confirm our assertion, we should consult financial journals to check event date announcements in the related time bar.

Conclusion

To conclude this dividend distribution analysis, we confirm an event detection in the considered time line by both of the used methodologies which provided similar results. The delay in detecting the event date between our two models is short (4 days), so in this case we confirm the fact that the CMRM, in spite of its simplicity, has a favorable performance in detecting the abnormal return. The variance decrease in the MM is also reported in the literature since its one of this model objectives.

Table 9: Windows delimiters for the Accor security between 01/11/01 and 28/02/02

	CMRM		MM	
Number of states	2		2	
Constraint model	mean is constant		mean is null	
Transition date	05/11/01		01/11/01	
μ	0.0048	0.0048	0	0
σ	$7.7324e-004$	0.0287	$3.1449e-007$	0.0187
b			0.9508	1.5612

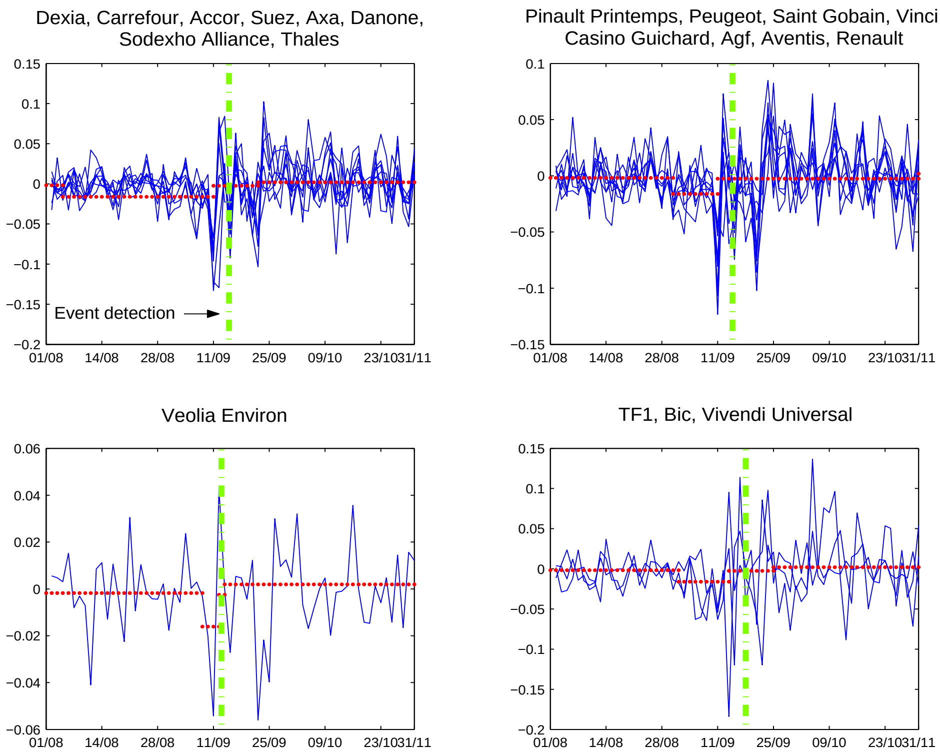

Figure 3: CMRM: Dates of detection of the 11 of September event.

6.1.3 Eleven of September event

The eleven of September event isn't a financial event like stock splits or earnings announcement but it may be considered as an information which is made public on the market and which may affect the value of one or several firms, which correspond to the event definition we stated in section 2. So, this event impact may be studied by the event study methodology (see for example Andrew and Thomas (2004), where the authors applied the Fama, Fisher, Jensen, and Roll methodology (1969) to identify abnormal returns using a Mean adjusted return approach and concluded that this terrorist attack negatively affected U.S. capital markets).

In this last application, we will study the eleven of September event impact on the French

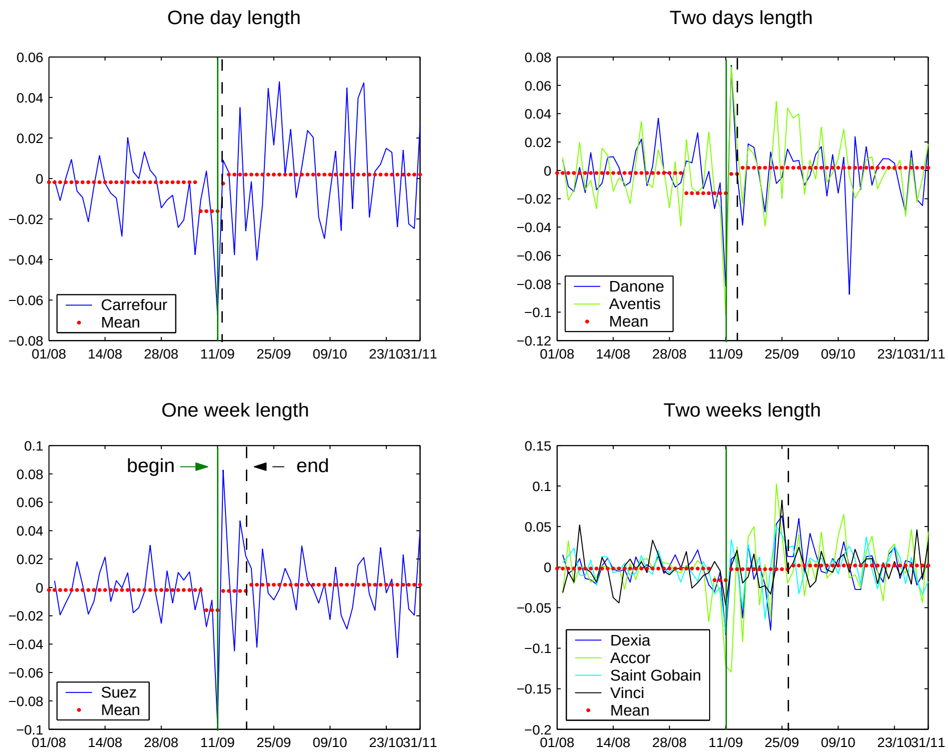

Figure 4: CMRM: Length of the 11 of September event.

market. So we consider the forty assets of the CAC40 at this date. Then we apply our methodology based on the CMRM and on the MM.

We consider a time line of three months going from the first of August to the *31-th* of October 2001.

Constant Mean Return Model

- Findings

Table 10: CMRM: Model parameters for the 11/09/01 event detection

Number of states: 4			
Constraint model: No constraint			
$\mu_1 = -0.0018$	$\mu_2 = -0.0161$	$\mu_3 = -0.0025$	$\mu_4 = 0.0019$
$\sigma_1 = 0.0147$	$\sigma_2 = 0.0291$	$\sigma_3 = 0.0515$	$\sigma_4 = 0.0223$

Results presented in table 10 show that we have four states with a mean and variance changing from a state to another. The highest variance is the third state variance. In table 11 are presented the event occurrence dates. This table illustrates that the eleven of September event has affected the market. The third state is obviously this event window.

• Interpretations

The first remark is that all the securities returns of our sample were affected by the eleven of September event. Sixteen firms detected this event exactly on its date occurrence. These firms are: Dexia, Carrefour, Accor, Suez, Axa, Danone, Sodexho Alliance, Thales, Pinault Printemps, Peugeot, Saint Gobain, Vinci, Casino Guichard, Agf, Aventis and Renault. Moreover, the other firms detect the event with a delay but the majority were affected at greatest two days after its occurrence. In figure 3, we superpose some returns time series of securities that detected the event at the same time.

The second remark concerns the event window length that depends on the firms. It varies from one day (Carrefour) to two weeks (Dexia, Accor, Saint Gobain, ...) as it is shown in figure 4.

Table 11: Constant mean return model: Data around the eleven of September

Dexia	06/09/01	11/09/01	26/09/01
Carrefour	05/09/01	11/09/01	12/09/01
Accor	06/09/01	11/09/01	26/09/01
Suez	06/09/01	11/09/01	18/09/01
Axa	30/08/01	11/09/01	01/11/01
Danone	07/09/01	11/09/01	13/09/01
Sodexho Alliance	30/08/01	11/09/01	22/09/01
Thales	04/08/01	11/09/01	21/09/01
Pinault Printemps	11/09/01	14/09/01	25/09/01
Peugeot	08/09/01	11/09/01	25/09/01
Saint Gobain	07/09/01	11/09/01	26/09/01
Vinci	11/09/01	25/09/01	26/09/01
Casino Guichard	11/09/01	12/09/01	13/09/01
Agf	24/08/01	11/09/01	24/09/01
Aventis	30/08/01	11/09/01	13/09/01
Renault	30/08/01	11/09/01	31/10/01
Veolia Environ.	08/09/01	12/09/01	13/09/01
TF1	06/09/01	13/09/01	26/09/01
Bic	07/09/01	13/09/01	22/09/01
Vivendi Universal	01/09/01	13/09/01	25/09/01

Market Model

• Findings

Table 12 shows that we have four states with a mean and variance changing from a state to another. The highest variance is the third state variance. But, the methodology based on the MM finds that some firms didn't be affected by any events during the considered time bar. These firms are: Carrefour, BNP Parisbas, Société générale and Lafarge (see figure 5).

Table 12: MM: Model parameters for the 11/09/01 event detection

Number of states: 4 Constraint model: No constraint			
$\mu_1 = 0.0012$	$\mu_2 = -0.0021$	$\mu_3 = 0.0069$	$\mu_4 = 0.0001$
$b_1 = 0.9625$	$b_2 = 1.1963$	$b_3 = -1.9057$	$b_4 = 0.7452$
$\sigma_1 = 0.0143$	$\sigma_2 = 0.0363$	$\sigma_3 = 0.0663$	$\sigma_4 = 0.0189$

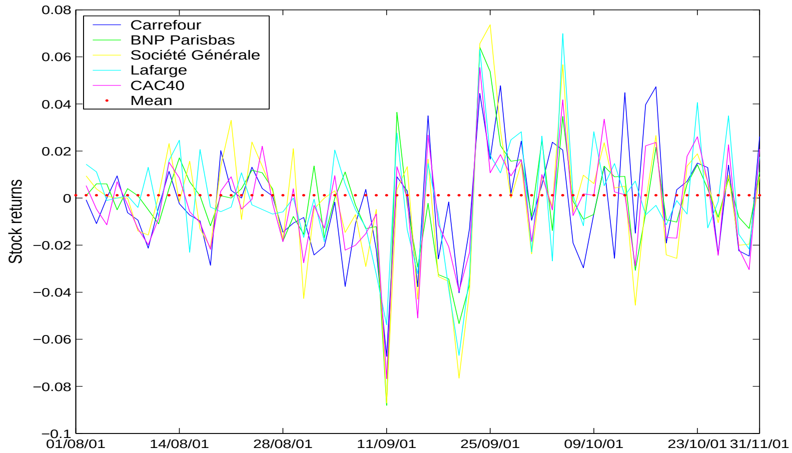

Figure 5: MM: Securities where no events are detected.

- Interpretations

The methodology based on the MM found the same number of states as the methodology based on the CMRM. The highest variance corresponds also to the third state and is the 11 of September event window as it is shown by figure 6. However, the length of this window is too short comparing with the result found by the CMRM (one or two days). In addition, the variance of the third state is higher in the case of the MM, which is in contrast with the fact that the MM reduces the variance. This result may be explained by the impact of incorporating information about the market's realized return (the CAC40 return) in the MM which summarizes the event impact in a short period with a high variance. Moreover, for some securities like Carrefour, for which the CMRM detected a one day event window, the MM didn't find an event impact.

Conclusion

The methodologies based on CMRM and MM detected an impact on the market induced by the 11 of September event. However, this impact is shorter in time with the MM and sometimes absent.

6.2 General remarks and conclusions

Our empirical study wasn't limited to the results reported below. In fact, we studied also the earnings announcement event which occurs generally at the end of the year. So, we run

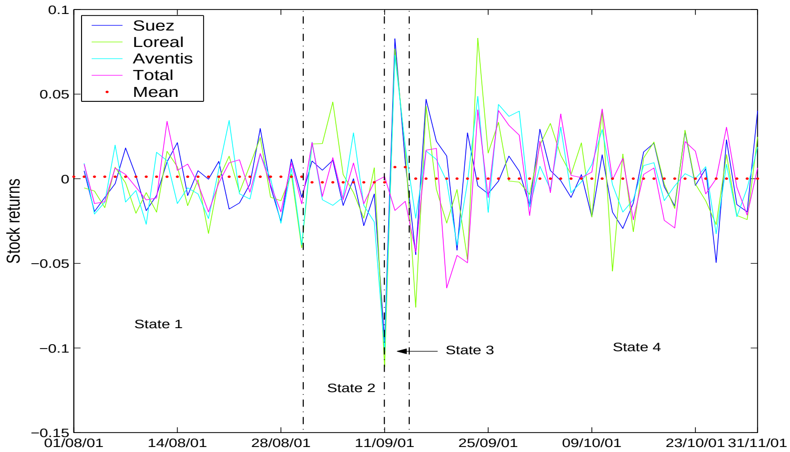

Figure 6: MM : Eleven of September event detection.

our programs on the CAC40 stock returns time series. We selected a period of four months: November, December, January and February. Then, we repeated this study for five years.

In this section, we don't report our findings in detail but we give general remarks about these results.

6.2.1 Model parameters

- The number of states:

Let's remember that the considered time bar is a four months period. With the CMRM, we always obtain two states (except in the 2001/2002 data and with the 11/09/2001 event). However, the results based on MM varies from two states to four states.

- The model constraints:

We obtain that the variance differs from a state to another, but the mean is generally constant (null in the case of the MM). This finding confirms the hypothesis usually used in event studies of supposing the mean null. The beta factor (Market Model) is generally different from a state to another, but sometimes it is constant.

6.2.2 Comparing the methodologies based on the CMRM and the MM

The first comparison point is that with the MM, the variance is always lower and the mean is null. These are two common hypothesis in event studies.

The second comparison point consists in choosing the best performing model. In fact, as we stated in the previous section, the CMRM performs well in some cases but sometimes, it provides non good results. Let's remember that if $b = 0$ in the MM, it becomes the CMRM and that we stated a candidate constraint model with the constraint $b = 0$. This model was never selected. So according to our methodology, empirical results show that the MM performs better than the CMRM. So, incorporating information about the market's realized return performs substantially better.

6.2.3 Event detection

The most important conclusion to report is that the empirical study shows that our approach performs well since it detects the events and provides reasonable windows delimiters for the event, pre event and post event periods.

Moreover, even if this approach doesn't provide a way to differentiate a period of abnormal return performance (event window) and a period of normal return performance (estimation window), recognizing each one of them isn't difficult if we have an idea about the market agents and information.

Appendix: EM algorithm for the Market Model

One observed sequence

We combine the market Model (2) and a hidden Markov chain with J states $\{S_t\}_{t=1,...,T}$ in the first order. In addition, the hidden sequence state $S_1 = s_1, ..., S_T = s_T$ will be written $S_1^T = s_1^T$. According to this combination, we have the following model:

$$R_t = \mu_{s_t} + b_{s_t} R_{m,t} + \varepsilon_t, \quad t = 1, ..., T \tag{4}$$

where R_t is the period-t observed return, μ_{s_t} is the mean return in the state s_t, $R_{m,t}$ is the period-t returns on the market portfolio and it is observed, b_{s_t} is the weight of the market portfolio returns in the state s_t, ε_t is the time period-t disturbance term for the security during the state s_t and $\varepsilon_t | S_t = s_t \sim \mathcal{N}(0, \sigma_{s_t}^2)$. Therefore, we notice that $R_t | S_t = s_t \sim \mathcal{N}(\mu_{s_t} + b_{s_t} R_{m,t} \, , \, \sigma_{s_t}^2)$.

Complete density for one sequence

We denote by $R_{m,1}^T$ the observed return of the market portfolio at the dates $t = 1, ..., T$; $\theta_1 = (\mu_j, b_j, \sigma_j^2)$, $j = 1, ..., J$ the vector of parameters to estimate relatively to the J MM models and $\theta_2 = (\pi_j, p_{ij})$, $j = 1, ..., J$ the vector of parameters to estimate relatively to the Markov chain $\{S_t\}$. In order to specificate the complete data problem, we assume that the sequences R_1^T and s_1^T are observed. According to the conditioning properties, the complete data density for a T-length sequence is written:

$$f(R_1^T, s_1^T; \theta) = f(R_1^T | s_1^T; \theta_1) f(s_1^T; \theta_2)$$

As $\{S_t\}$ is a first order Markov chain, the joint probability of the state sequences is given by

$$\begin{aligned} f(s_1^T; \theta_2) &= p(S_1^T = s_1^T; \theta_2) \\ &= p(S_1 = s_1) \prod_{t=2}^{T} p(S_t = s_t | S_{t-1} = s_{t-1}) \\ &= \pi_{s_1} \prod_{t=2}^{T} p_{s_{t-1} s_t} \end{aligned}$$

The observed returns conditionally to the state sequence are independent, so:

$$f(R_1^T|s_1^T;\theta_1) = \prod_{t=1}^{T} f(R_t|s_t;\theta_1) \tag{5}$$

$R_t|s_t \sim \mathcal{N}(\mu_{s_t} + b_{s_t} R_{m,t}\ ,\ \sigma^2_{s_t})$ means that

$$f(R_t|s_t;\theta_1) = \frac{1}{\sigma_{s_t}\sqrt{2\pi}} \exp\left[-\frac{(R_t - \mu_{s_t} - b_{s_t} R_{m,t})^2}{2\sigma^2_{s_t}}\right]$$

Therefore the complete data density for the sequence relative to the asset a for a period length T is written:

$$f_a(R^T_{a,1}, s^T_{a,1};\theta) = \prod_{t=1}^{T} \frac{1}{\sigma_{s_{a,t}}\sqrt{2\pi}} \exp\left[-\frac{\left(R_{a,t} - \mu_{s_{a,t}} - b_{s_{a,t}} R_{m,t}\right)^2}{2\sigma^2_{s_{a,t}}}\right] \pi_{s_{a,1}} \prod_{t=2}^{T} p_{s_{a,t-1}s_{a,t}}$$

where $R^{T_a}_{a,1}$ is the observed return of the asset a at the dates $t = 1, ..., T_a$, $s^{T_a}_{a,1}$ is the hidden state sequence of the asset a at dates $t = 1, ..., T_a$, $\theta = (\pi_j, p_{ij}, b_j, \mu_j, \sigma^2_j)$, $i, j = 1, ..., J$ is the parameters vector to estimate when the process $\{S_t\}$ is in the state j at time period t and in the state i at $t - 1$.

N observed sequences

Now, we consider N assets. The indice a will design the asset a and the indice t remains the temporal indice. The modeling (2) for the asset a will be written as follows:

$$R_{a,t} = \mu_{s_{a,t}} + b_{s_{a,t}} R_{m,t} + \varepsilon_{a,t}, \quad t = 1, ..., T_a \tag{6}$$

where $R_{a,t}$ is the period-t observed return of the asset a; $\mu_{s_{a,t}}$ is the mean return in the state $s_{a,t}$; $R_{m,t}$ is the period-t market portfolio return; and $\varepsilon_{a,t} \sim \mathcal{N}(0, \sigma^2_{s_{a,t}})$ is the time period-t disturbance term of the security a in the state $s_{a,t}$. The $\varepsilon_{a,t}$ are independent identically distributed (*iid*).

Complete density for the N sequences

We assume that the N assets are independent. Therefore the complete data density is written as the product over all the assets of the complete data density of each one of them.

$$\begin{aligned} f(R, S, \theta) &= \prod_{a=1}^{N} f_a(R^{T_a}_{a,1}, s^{T_a}_{a,1};\theta) \\ &= \prod_{a=1}^{N}\prod_{t=1}^{T_a} \frac{1}{\sigma_{s_{a,t}}\sqrt{2\pi}} \exp\left[-\frac{\left(R_{a,t} - \mu_{s_{a,t}} - b_{s_{a,t}} R_{m,t}\right)^2}{2\sigma^2_{s_{a,t}}}\right] \pi_{s_{a,1}} \prod_{t=2}^{T_a} p_{s_{a,t-1}s_{a,t}} \end{aligned}$$

where $R = (R_1^{T_a}, ..., R_N^{T_a})$, $S = (S_{1,1}^{T_a}, ..., S_{N,1}^{T_a})$. Thus, the log-likelihood of the complete data is given by

$$
\begin{aligned}
l(R, S, \theta) &= \log f(R, S, \theta) \\
&= \sum_{a=1}^{N} \Bigg[\log \pi_{s_{a,1}} + \sum_{t=2}^{T_a} \log p_{s_{a,t-1} s_{a,t}} - \frac{T_a}{2} \log 2\pi \\
&+ \sum_{t=1}^{T_a} \left(- \log \sigma_{s_{a,t}} - \frac{\left(R_{a,t} - \mu_{s_{a,t}} - b_{s_{a,t}} R_{m,t}\right)^2}{2\sigma_{s_{a,t}}^2} \right) \Bigg]
\end{aligned}
$$

E-Step

On working with (18) as the complete data log-likelihood function for θ, the E-step on the *(q+1)-th* iteration is effected by finding the conditional expectation of complete data log-likelihood, given the observation data, using the current fit $\theta^{(q)}$ for θ.

$$
\begin{aligned}
&\mathcal{Q}(\theta^{(q+1)}, \theta^{(q)}) = \mathbb{E}\left[l(R, S, \theta) | R; \theta^{(q)} \right] \\
&= -\frac{T}{2} \log 2\pi + \sum_{j=1}^{J} \sum_{a=1}^{N} p(S_{a,1} = j | R; \theta^{(q)}) \log \pi_j \\
&+ \sum_{i,j=1}^{J} \sum_{a=1}^{N} \sum_{t=2}^{T_a} p(S_{a,t-1} = i, S_{a,t} = j | R; \theta^{(q)}) \log p_{ij} \\
&+ \sum_{j=1}^{J} \sum_{a=1}^{N} \sum_{t=1}^{T_a} p(S_{a,t} = j | R; \theta^{(q)}) \left[- \frac{\mathbb{E}\left[\left(R_{a,t} - \mu_j - b_j R_{m,t}\right)^2 | R, s_{a,t} = j; \theta^{(q)} \right]}{2\sigma_j^2} - \log \sigma_j \right]
\end{aligned}
$$

We denote $\gamma_{a,j}(t) = p(S_{a,t} = j | R; \theta^{(q)})$. We know that the vectors R_a are independent. Hence, conditioning $S_{a,t} = j$ on R is equivalent to conditioning $S_{a,t} = j$ on the observation data relative to the asset a. Therefore, we have that

$$
\gamma_{a,j}(t) = p(S_{a,t} = j | R_{a,1}^{T_a}; \theta^{(q)})
$$

and

$$
\mathbb{E}\left[\left(R_{a,t} - \mu_j - b_j R_{m,t}\right)^2 | R, s_{a,t} = j; \theta^{(q)} \right] = \mathbb{E}\left[\left(R_{a,t} - \mu_j - b_j R_{m,t}\right)^2 | R_{a,1}^{T_a}, s_{a,t} = j; \theta^{(q)} \right]
$$

Moreover, we cannot know the probability of an observation $R_{a,t}$ $t = 1, ..., T_a$ unless we know the state $s_{a,t}$ in which it is at the time period t. Then, the probability of $(R_{a,t} - \mu_j - b_j R_{m,t})^2 | R_{a,1}^{T_a}, s_{a,t} = j$ is resumed to the probability of $(R_{a,t} - \mu_j - b_j R_{m,t})^2 | R_{a,t}, s_{a,t} = j$. It follows that

$$\mathcal{Q}(\theta^{(q+1)},\theta^{(q)}) = -\frac{T}{2}\log 2\pi + \sum_{j=1}^{J}\sum_{a=1}^{N}\gamma_{a,j}(1)\log \pi_j$$
$$+\sum_{i,j=1}^{J}\sum_{a=1}^{N}\sum_{t=2}^{T_a} p(S_{a,t-1}=i, S_{a,t}=j|R_{a,1}^{T_a};\theta^{(q)})\log p_{ij}$$
$$+\sum_{j=1}^{J}\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)\left[-\frac{\mathbb{E}\left[(R_{a,t}-\mu_j-b_jR_{m,t})^2|R_{a,t},s_{a,t}=j;\theta^{(q)}\right]}{2\sigma_j^2} - \log\sigma_j\right] \tag{7}$$

We now consider the calculation of the expressions $\gamma_{a,j}(t)$ and $p(S_{a,t-1}=i, S_{a,t}=j|R_{a,t}, s_{a,t}=j;\theta^{(q)})$ which will be effected by the Forward-Backward algorithm described above for the CMRM. For the implementation of this algorithm in the case of market model, we take

$$f(R_{a,t}|s_{a,t}=j) = \frac{1}{\sigma_j\sqrt{2\pi}}\exp\left[-\frac{(R_{a,t}-\mu_j-b_jR_{m,t})^2}{2\sigma_j^2}\right] \tag{8}$$

M-Step

The hidden Markov chain parameters of this model can be easily estimated using the formula as in the case of CMRM. However, the computation of σ_j^2, $j=1,...,J$ in the *q-th* iteration requires the maximization of $\mathcal{Q}(\theta^{(q+1)},\theta^{(q)})$ with respect to σ_j^2. In this case we obtain

$$\sigma_j^{2(q+1)} = \frac{\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)\mathbb{E}\left[(R_{a,t}-\mu_j-b_jR_{m,t})^2|R_{a,t},s_{a,t}=j;\theta^{(q)}\right]}{\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)}$$

The term $(R_{a,t}-\mu_j-b_jR_{m,t})^2|R_{a,t}=\varepsilon_{a,t}$ when $s_{a,t}=j$; $\varepsilon_{a,t}|s_{a,t}=j\sim\mathcal{N}(0,\sigma_j^2)$ and $R_{a,t}|s_{a,t}=j\sim\mathcal{N}(\mu_j+b_jR_{m,t}\,,\,\sigma_j^2)$. In addition, $Cov(\varepsilon_{a,t},R_{a,t}|s_{a,t})=\sigma_j^2$. Thus we can deduce that:

$$\mathbb{E}\left[(R_{a,t}-\mu_j-b_jR_{m,t})^2|R_{a,t},s_{a,t}\right] = (R_{a,t}-\mu_j-b_jR_{m,t})^2$$

therefore $\forall\, j=1,...,J$

$$\sigma_j^{2(q+1)} = \frac{\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)(R_{a,t}-\mu_j^{(q)}-b_j^{(q)}R_{m,t})^2}{\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)}$$

The computation of μ_j, $j=1,...,J$ in the *q-th* iteration requires the maximization of $\mathcal{Q}(\theta^{(q+1)},\theta^{(q)})$ with respect to μ_j. We obtain that

$$\mu_j^{(q+1)} = \frac{\sum\limits_{a=1}^{N}\sum\limits_{t=1}^{T_a}\gamma_{a,j}(t)\left(R_{a,t} - b_j^{(q)}R_{m,t}\right)}{\sum\limits_{a=1}^{N}\sum\limits_{t=1}^{T_a}\gamma_{a,j}(t)}$$

Finally, the computation of the market portfolio coefficient in the *q-th* iteration requires the maximization of $\mathcal{Q}(\theta^{(q+1)}, \theta^{(q)})$ with respect to b_j which means that:

$$\frac{\partial}{\partial b_j}\sum_{a=1}^{N}\sum_{t=1}^{T_a}\gamma_{a,j}(t)\mathbb{E}\left[(R_{a,t} - \mu_j - b_j R_{m,t})^2\,|R_{a,t}, s_{a,t} = j\right] = 0$$

We obtain that, for $j = 1, ..., J$

$$b_j = \frac{\sum\limits_{a=1}^{N}\sum\limits_{t=1}^{T_a}\gamma_{a,j}(t)R_{m,t}\left(R_{a,t} - \mu_j^{(q)}\right)}{\sum\limits_{a=1}^{N}\sum\limits_{t=1}^{T_a}\gamma_{a,j}(t)R_{m,t}^2}$$

References

[1] Akaike, H. (1974). A new look at the statististical identification model. *IEEE Trans. Auto. Control* **19**, 716–723.

[2] Aktas, N., Bodt, E. and Cousin, J-G. (2003). Event study methodology under noisy estimation period. *EFMA Helsinki Meetings Journal* **1**, 1–42.

[3] Andrew, H.C. and Thomas, F.S. (2004). The effects of terrorism on global capital markets. *European Journal of Political Economy* **20**, 349–366.

[4] Cowan, Arnold R. (1992). Nonparametric Event Study Tests. *Review of Quantitative Finance and Accounting* **2**, 343–358.

[5] Ashley, J. (1962). Stock prices and changes in earnings and dividends: Some empirical results. *Journal of Political Economy* **70**, 82–85.

[6] Baker, C. (1956). Effective stock splits. *Havard business review* **34**, 101–106.

[7] Baker, C. (1957). Stock splits in a bull market. *Havard business review* **35**, 72–79.

[8] Baker, C. (1958). Evaluation of stock dividends. *Havard business review* **36**, 99–114.

[9] Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research* **6**, 159–178.

[10] Ball, C. and Torous, W. (1988). Investigating security-price performance in the presence of eventdate uncertainty. *Journal of Financial Economics* **22**, 123–154.

[11] Beitel, P., Schiereck, D. and Wahrenburg, M. (2003). Explaining the M&A-success in European bank mergers and acquisitions. *Working Papers, e-finance lab.* Technische Universitat Darmstadt.

[12] Biernacki, C., Celeux, G. and Govaert, G. (2002). Choosing starting values for the em algorithm for getting the highest likelihood in multivariate gaussian mixture models. *Computational Statistics and Data Analysis* **41**, 561–575.

[13] Bohemer, E., Musumeci, J. and Poulsen, A. (1991). Event study methdology under conditions of event-induced variance. *Journal of Financial Economics* **30**, 253–272.

[14] Brown, K.C., Lockwood, L.J. and Lummer, S.L. (1985). An examination of event dependency and structural change in security pricing models. *Journal of Financial and Quantitative Analysis* **20**, 315–334.

[15] Brown, S.J. and Warner, J.B. (1980). Measuring security price performance. *Journal of Financial Economics* **8**, 205–258.

[16] Brown, S.J. and Warner, J.B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics* **14**, 3–31.

[17] Cable, J. and Holland, K. (1999). Modelling normal returns in event studies: A model selection approach and pilot study. *The European Journal of Finance* **5**, 331–341.

[18] Celeux, G. and Diebolt, J. (1985). The sem algorithm: a probabilisitic teacher algorithm derived from the em algorithm for the mixture problem. *Computational Statistics and Data Analysis* **2**, 73–82.

[19] Dodd, P. and Warner, J.B. (1983). On corporate governance: A study of proxy contests. *Journal of Financial Economics* **11**, 401–438.

[20] Dolley, J. (1933). Characteristics and procedure of common stock split-ups. *Havard business review* **11**, 316–326.

[21] Dyckman, T.R., Philbrick D. and Stephan, J. (1984). A comparison of event study methodologies using daily stock returns: A simulation approach. *Journal of Accounting Research* **22**, 1–33.

[22] Fama, E., Fisher, L., Jensen, M. and Roll, R. (1969). The adjustement of stock prices to new information. *International Economic Review* **10**, 1–21.

[23] Foster, G. (1973). Stock market reaction to estimates of earnings per share by company officials. *Journal of Accounting Research* **11**, 25–37.

[24] Grar, A. (1997). Etude d'événement. *In Encyclopédie des Marchés Financiers*, Economica, Yves Simon Ed., Vol 1, pp. 462–484.

[25] Kass, R.E. and Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773–795.

[26] MacKinlay, A.C. (1997). Event studies in econmics and finance. *Journal of Economic literature* **35**, 13–39.

[27] McLachlan, G. and Peel, D. (2000). Finite Mixture Models. Wiley Series in Probability and Statistics. John Wiley & Sons.

[28] McWilliams, A. and Siegel, D. (1997). Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal* **40**, 626–657.

[29] Myers, J. and Bakay, A. (1948). Influence of stock split-ups on market price. *Havard Business Review* **26**, 251–265.

[30] Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**, 257–286.

[31] Salinger, M. (1992). Standard errors in event study. *Journal of Financial and Quantitative Analysis* **27**, 39–53.

[32] Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics* **6**, 461–464.

[33] Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* **13**, 260–269.