A dynamic analysis of stock markets using a latent Markov model

Luca De Angelis*

Department of Statistical Sciences, University of Bologna, Italy

Leonard J. Paas

Department of Marketing, Faculty of Economics and Business, VU University, Amsterdam, The Netherlands

October 29, 2010

* Please send all correspondence to: Luca De Angelis, Department of Statistical Sciences, Alma Mater Studiorum University of Bologna, Via delle Belle Arti, 41, 40126, Bologna. Italy. E-mail: l.deangelis@unibo.it; Phone: +39-051-2094628.

A dynamic analysis of stock markets using a latent Markov model

This paper proposes an innovative framework to detect financial crises, pinpoint the end of a crisis and predict future developments in stock markets. This proposal is based on a latent Markov model and allows for a specific focus on conditional mean returns. By analyzing weekly changes in the U.S. stock market indexes over a period of 20 years, this study obtains an accurate detection of stable and turmoil periods and a probabilistic measure of switching between different stock market conditions. The results contribute to the discussion of the capabilities of latent Markov models and give financial operators some appealing investment strategies.

Keywords: Stock market pattern analysis; Regime-switching; Forecasting; Latent Markov model; Financial crises; Market stability periods.

1. Introduction

Since the seminal work by Hamilton [14], Markov-switching approaches have been applied frequently to analyze stock price index data. The initial autoregressive framework, with unobserved changes in regimes modelled by a first-order Markov chain, has been extended [15] with an autoregressive conditional hetereoskedasticity (ARCH) specification with regime switching, governed by an unobserved Markov chain (SWARCH), which can evaluate changes in stock market volatilities. This method prevents an excess of persistence, which ARCH models impute to stock index volatility, and improves forecasting performance. Further extensions of the Markov-switching model specification [15] investigate changes in stock market volatility [4, 9, 12, 13, 16, 19, 23]. This article offers another type of extension that can model stock index dynamics.

Previous research mainly has investigated conditional variances, categorizing time periods according to variances in stock market indexes, such that periods with similar volatilities have relatively high probabilities of being allocated to the same category. However, we classify time periods on the basis of mean returns, which provide crucial proxies of expected returns. Accordingly, we investigate conditional means that characterize different stock market regimes, for which purpose we undertake an exploratory analysis of the latent stochastic process that underlies the observed time-series of the stock market return distribution, using a latent Markov model (LMM) [3]. The latent process comprises a discrete number of states that can be interpreted as different market regimes. Therefore, as our first research contribution, we provide a probabilistic classification of each time observation into market regimes, according to the value of the observed return at that time and the correlation structure of the series.

Furthermore, prior financial literature has addressed the topic of conditional means only marginally, claiming that "it is well known that conditional means are hard to estimate" [2] and thus rarely investigated them in depth. Some authors who consider differences in variances find similar or non-significantly different means across regimes that could be constrained to be equal [2, 9]. But an accurate evaluation of conditional means might improve time-series classification. Stable periods, crises, and financial bubbles should be characterized by significantly different mean returns, and our analysis reveals that in the U.S. stock market, the conditional means differ statistically significantly across time periods.

We therefore undertake an endogenous detection of different market phases, which contributes to extant research involving Markov-switching models that usually identify the number of market regimes a priori, often predetermining the number of latent states that characterize the unobservable Markov chain [4, 9, 23]. For example, a study might select two predetermined latent states that represent a low-volatility regime (i.e., the bull market phase) and a high-volatility regime (i.e., the bear market phase) [17, 18, 22]. However, model complexity and the many parameters that must be estimated in SWARCH and MS-GARCH specifications make it impractical to include more than four latent states [23]. As an alternative, we determine the number of latent states by turning to statistical procedures and the more parsimonious LMM model, which introduces a new methodological dimension to this step and allows for more than two or four latent states. Therefore, LMM enables a researcher to focus on dynamics and regime switches across different stock market phases, which offers valuable insights for financial variables analyses.

In Section 2, we discuss a formal introduction to the applied LMM. Section 3 includes the theoretical framework for our application of LMM, followed by a discussion in Section 4 of the analyzed data, analysis, and results. We conclude in Section 5.

2. Latent Markov Model

In this section, we introduce the latent Markov model specification, parameter estimation and model selection. Furthermore, we describe the classification procedure used for allocating the time observations to the estimated latent states.

2.1. Model specification

Although the LMM model originally was introduced to analyze categorical indicators [26, 28], recent work has exploited its potential to analyze continuous variables [5, 18, 25, 29]. Our work extends such recent work by focusing on the stock market return distribution.

Specifically, let z_t denote the return observation of a stock market index at time t (t = 1, …, T). The LMM analyzes $f(z)$, or the probability density function of the return distribution of the market index over time, using a latent transition structure defined by a first-order Markov process. For each time point t, the model defines a single discrete latent variable denoted by y_t, which consists of S latent classes (usually referred to as latent states). Thus the LMM includes T latent variables and can be specified as

$$f(z) = \sum_{y_1=1}^{S}\sum_{y_2=1}^{S}\cdots\sum_{y_T=1}^{S} f(y_1,...,y_T)f(z;y_1,...,y_T), \tag{1}$$

where

$$f(y_1,...,y_T) = f(y_1)\prod_{t=2}^{T} f(y_t \mid y_{t-1}), \tag{2}$$

and

$$f(z;y_1,...,y_T) = \prod_{t=1}^{T} f(z_t \mid y_t). \tag{3}$$

Equation (1) reveals that the model is a mixture, with S^T latent classes (mixture components). In the other mixture models, $f(z)$ can be obtained by marginalizing with respect to the latent variables. Because the y are discrete variables, Equation (1) is a weighted average of probability densities $f(z; y_1,..., y_T)$, where the latent class membership probabilities (or prior probabilities) $f(y_1,..., y_T)$ are weights [20]. Furthermore, Equations (2) and (3) depict the conditional independence assumption implied by the LMM, which can simplify the density functions $f(y_1,..., y_T)$ and $f(z; y_1,..., y_T)$. Equation (2) implies an additional model assumption, namely, that $f(y_1,..., y_T)$ follows a first-order Markov process. Thus, latent state y_t is associated with y_{t-1} and y_{t+1} only. Furthermore, $f(y_1)$ denotes the (latent) initial-state probability function. According to Equation (3), the return observation at time t is independent of observations at other time points, conditional on the latent state occupied at time t.

Furthermore, $f(y_t \mid y_{t-1})$ denotes the latent transition probability function, which provides the probability of being in a particular latent state at time t, conditional on the state occupied at the previous time point, t – 1. Assuming a homogenous transition process with respect to time, we achieve a latent transition matrix in which the generic element $p_{jk} = \Pr\text{ob}(y_t = k \mid y_{t-1} = j)$ denotes the probability of switching from latent state j at time t – 1 to latent state k at time t, for j, k = 1, …, S.

This LMM can be considered a restricted version of the regime-switching autoregressive model proposed by Hamilton [14], which restricts the autoregressive component to 0. Using the specification provided by Equations (1)–(3), we intend to model the latent stochastic process y_t to gain insights into stock market dynamics and a specific focus on the different conditional means $\mu(z_t \mid y_t = k)$ for k = 1, …, S.

2.2. Parameter estimation

For the parameter estimation, we maximize the log-likelihood function (*LL*) according to the expectation-maximization (EM) algorithm [7]. However, the iterative procedure of the EM algorithm is often impractical for estimating a LMM. For the expectation step, it must compute and store S^T entries of the joint posterior latent distribution, $f(y_1,...,y_T \mid z)$, so computational time increases exponentially with *T*, and even a moderate time-series length may prevent the convergence of the algorithm. We therefore use a variant of the EM algorithm, called the forward-backward or Baum-Welch algorithm [3], as extended by Paas *et al.* [21] for an application to data sets with multiple observed indicators and implemented in the Latent GOLD 4.5 computer program [27].

The forward-backward algorithm exploits the conditional independence assumption of the LMM to compute the joint posterior latent distribution by estimating the missing data, which in the LMM are unobserved state memberships. This estimation is realized by computing the expected value of the log-likelihood function, given the current parameter values and the observed data. The maximization step uses standard maximum likelihood estimation methods for complete data to update the model parameters. The algorithm cycles between these steps until it reaches a previously defined convergence criterion.

2.3. Model selection and class membership

Model selection involves the choice of the number of latent states *S*, which in our framework represents the number of market regimes. This extension of existing approaches addresses their inability to estimate Markov switching models with $S > 4$ because of their complexity [23]. The choice of the appropriate number of latent classes is based on the Akaike information criterion (AIC) [1]:[1]

$$AIC = -2LL + 2NPar\,,$$

where *NPar* is the number of model parameters.

For our analysis, stock index returns are the indicators z_t, for $t = 1, \ldots, T$. Each z_t is classified into one latent state according to the estimated posterior probabilities. That is, z_t is

[1] Markov chain order identification in LMM remains an unresolved issue (see [6], Chap. 15 for a recent discussion), and there are several concerns about the robustness and reliability of information criteria. We also

allocated to latent state j if $\hat{f}(y_t = j \mid z_t) > \hat{f}(y_t = k \mid z_t)$ for every $k = 1, \ldots, S$. In this modal classification, time points with a similar development are more likely to be allocated to the same latent state than are those with highly divergent developments.

3. Theoretical framework

The use of the LMM for these purposes is promising, because financial markets are characterized by frequent changes in regimes. If stock market index returns are subject to discrete changes in regimes, including periods in which the dynamic pattern of the series differs markedly, a nonlinear model should exploit the time path of the observed series to draw inferences about a set of discrete latent states [14]. Different market regimes thus should be characterized by different means and standard deviation values or—using the terminology of portfolio theory framework—by different risk–return profiles. During a financial crisis, the stock market experiences a strong negative mean return, and the standard deviation, used as a proxy of risk, is large. During more stable phases, stock returns fluctuate around a constant mean, and the standard deviation of the index value is lower. Different regimes in different time periods imply the ability to cluster time observations, according to the similarity in the dynamics of the index value and the volatility of that index [11]. Time periods with more (less) similar index dynamics have a higher (lower) probability of being allocated to the same cluster. Moreover, empirical analyses clearly show stock returns are characterized by asymmetry and larger kurtosis than the Gaussian distribution [10], which invalidates inferences. By modelling regime changes using a mixture of normal distributions, LMM provides an effective solution to these issues [8].

The LMM applied in this study classifies different observations into a limited set of regimes, on the basis of stock market price index dynamics. For example, a week

---

agree with the concerns about a uncritical use of these indicators. However, we believe that they can contribute characterized by a strong decline in the stock market price index may be allocated to the large value decrease market regime, whereas weeks defined by small changes likely appear in the stable market regime. Switches between regimes are modelled as a Markov process. Using equations previously introduced by Paas *et al.* [19], we also can employ the LMM to predict future stock market dynamic patterns. These one-step ahead forecasts, based on latent transition probabilities, reveal which regimes the stock market is likely to experience in the next week.

4. Empirical analysis using the latent Markov model

In this section, the proposed model is applied to the U.S. stock market index S&P 500. First, data description is provided. Next, empirical results and applications of the estimated model are discussed. Finally, the estimated model forecasting performance is evaluated.

4.1. Data description

Our analysis is based on weekly returns for the U.S. stock market price index S&P 500, calculated as the percentage achieved in the relative variation in index prices, p_t: $z_t = \frac{p_t - p_{t-1}}{p_{t-1}} \times 100$. The data set covers the period from January 5, 1990, to January 1, 2010, which includes T = 1044 time points. As Figure 1 shows, our data set includes at least three periods with high volatility, which reflect stronger fluctuations and rapid changes from positive to negative peaks: prior to 1991, from 1997 to 2003, and after 2008. Our data also contain several stable periods, such as those from 1992 to 1997 and from mid-2003 to the end of 2007. According to NBER-defined business cycles,[2] the total study period contains three crisis periods: the "savings and loan" crisis (July 1990–March 1991), the Internet bubble burst and September 11 attacks (March 2001–November 2001), and the credit crisis (starting in

---

to current procedures, for which the choice of latent states is somewhat arbitrary.

December 2007). A pressing question during such periods is when the economic situation might improve. Therefore, we apply the LMM to discriminate endogenously the stable from the crisis periods and recognize the end of a crisis, according to the mean returns of the stock market price index. The LMM also enables us to predict what will happen during the subsequent week in a period of crisis and which market-regime is most likely.

INSERT FIGURE 1 ABOUT HERE

Table 1 contains the different values of the mean returns and standard deviations for the entire time sample and five subperiods, which can be associated with low or high volatility market phases: According to the standard deviation values, the five subperiods are characterized by different levels of variability. In particular, the levels differ greatly for periods II and IV and periods I, III, and V, as well as across the three high-volatility phases. The latter finding implies that each financial crisis creates its own peculiarities.

The Jarque-Bera normality test results are significant for the entire data set, implying a significant difference between the observed and a normal distribution. We also can reject the normality assumption for subperiods III and V, according to the Jarque-Bera test. Therefore, the LMM may be a desirable alternative to traditional financial econometric models, because it accounts for both asymmetry and more kurtosis than a normal distribution.

INSERT TABLE 1 ABOUT HERE

4.2. Model estimation and class profiling

We estimate the LMM for 1 to 8 latent states (S = 1, …, 8) and provide, in Table 2, the maximum log-likelihood function, number of estimated parameters, and AIC values. According to the AIC criterion, the LMM with seven latent states provides the best fit to the

[2] Source: http://www.nber.org/cycles.html.

data. In our framework, these latent states represent seven different stock market regimes. According to the return means in each state, the S&P 500 index reveals three negative and four positive regimes; in Table 3, we label the profiles of the seven market regimes using the return means. For example, latent state 1 has an average return of -13.45% and constitutes 0.31% of the $T = 1044$ analyzed weeks.

INSERT TABLE 2 AND 3 ABOUT HERE

As Table 3 shows, the LMM can define different regimes of the stock market. The return means differ significantly across latent states, according to both the Wald test ($W = 817.02$, $df = 6$, $p\text{-}value < 0.001$) and ANOVA ($F = 358.24$, $df = 6; 1037$, $p\text{-}value < 0.001$), so we reject the null hypothesis of equality between conditional means. Furthermore, the dispersion within each latent state is relatively low, according to the similar standard deviation values in Table 3, with the exception of latent state 1, which represents the biggest stock market drops.

Figure 2 displays the weekly return time-series of the S&P 500 index and that obtained through the seven-state LMM. The estimated series is plotted using the latent state return means. The LMM approximates the observed time series of the S&P 500 index quite accurately. Moreover, it detects two stable periods, corresponding to latent state 4, as represented by the straight lines from March 1, 1991–March 14, 1997, and from May 23, 2003–July 20, 2007 (Figure 2). The LMM results also show that the three periods, characterized by high volatility, include frequent switches between regimes with positive and negative conditional means. These three periods correspond to the three crises and recessions we noted previously, though the 2001 crisis was preceded by a period of turmoil that started in 1997, which may indicate a spillover of the Asian crisis to the U.S. stock market [24].

INSERT FIGURE 2 ABOUT HERE

4.3. Latent transition analysis

In the transition probability matrix estimated by the LMM in Table 4, the transition probabilities define the stock market regime-switching. The values on the diagonal represent state persistence, that is, the probabilities of remaining in a particular market regime. The modal latent state 4 has high persistence (p_{44} = 0.995) and represents the stable market regime. As Figure 2 reveals, this result indicates that the U.S. stock market tended to remain in that regime (T = $(1 - p_{44})^{-1} \approx 217$ weeks). The off-diagonal p_{jk} values indicate the probabilities of market regime-switching. It is quite likely that the S&P 500 index switches from a very negative phase to a period of fast growth (p_{16} = 0.9606), whereas the opposite switch is unlikely (p_{61} = 0.0005).

The probabilities in Table 4 thus underline some important features of market regime-switching. First, for latent states 2 to 5, the transition probabilities p_{jj} are relatively high, whereas for latent states 1 and 7, persistance is unlikely (p_{66} and p_{77} < 0.01), and state 6 has a persistance probability of 0.10. Second, when the S&P 500 declines (states 1 or 2) at time t, at time t + 1, the market may continue in a negative phase (p_{21} = 0.069 and p_{22} = 0.427) or switch to a positive regime (p_{16} = 0.961, p_{25} = 0.168, p_{26} = 0.126, and p_{27} = 0.203). The other states rarely occur after state 1 or 2.

Overall, 31 of the 49 transition probabilities are less than 0.05 in the transition matrix, which indicates most regime switches are very unlikely for the S&P 500 index. Accordingly, our results offers interesting insights for future market phase predictions.

INSERT TABLE 4 ABOUT HERE

4.4. Recognition of the stable market phase

The model also can predict a stable period, after a previous stable period or after a period in which the market was not categorized in the stable latent phase 4. The latent state characterized by a moderate positive mean return is most common and has a persistence probability of close to 1. These features denote a stable market regime, as mentioned in Section 4.2. In these periods, which correspond to subperiods II and IV in Table 1, the stock price index value does not experience large and frequent changes. The ranges between the minimum and the maximum returns of the S&P 500 index in periods II and IV are 8.44 and 7.95, respectively, and the standard deviations are 1.375 and 1.435. In contrast, periods I, III, and V are characterized by ranges of 10.37, 19.38, and 30.23 and standard deviations of 2.333, 2.789, and 3.893, respectively (see Table 1). Therefore, the time points classified into latent state 4 can be interpreted as belonging to a low-volatility period of market stability, whereas the other six latent states refer to the high-volatility periods I, III, and V, characterized by frequent switches between high- and low-return regimes.

In Figure 3, the estimated posterior probabilities for latent state 4, $\hat{f}(y_t = 4 \mid z_t)$, underlie the high level of confidence with which the LMM determines the two stable periods characterized by low level of volatility: Of the 536 observations classified into latent state 4, only 38 have a posterior probability less than 0.90. In other words, the probability of remaining in latent state 4 across time points is quite high, as correctly predicted by the model.

INSERT FIGURE 3 ABOUT HERE

To evaluate the model's capability to detect a stable period after a period of crisis, we estimate the LMM with 7 latent states for shorter time series. The beginning of the second stable regime (period IV in Table 1) provided by the LMM, when applied to the entire time series, starts on May 23, 2003. We use the crisis before May 2003 to evaluate the model's capacity to detect a stable period, because period II (Table 1) is preceded by a very short unstable period in our data set, and the crisis that started in 2001 had not ended. We assess how many weeks of stability are required to detect the end of the financial crisis, which the LMM estimates as May 23, 2003. Therefore, we first estimate the model using data from January 5, 1990–May 23, 2003, and then from January 5, 1990–May 30, 2003, and so on. A stable period emerges when multiple weeks, latest in time, are allocated to the stable latent state 4.

Our analysis reveals that LMM can detect the stable market phase within 13 weeks of May 23, 2003. That is, a period containing only stable regimes after May 23, 2003, appears when we use the data set with stock index returns from January 5, 1990–August 22, 2003. In the analysis in which we included fewer than 13 weeks, the last few observations are not allocated to the stable latent state 4. This feature of LMM is potentially useful for detecting when the financial crisis that started in 2007 will end. That is, by the time we concluded our analysis (January 1, 2010), there were not 13 consecutive weeks allocated to the stable latent state 4; the crisis had not ended by January 1, 2010. However, Figure 3 shows that the posterior probability for latent state 4 increased in the most recent observations in our sample, reaching $\hat{f}(y_{1/1/2010} = 4 \mid z_{1/1/2010}) = 0.34$, though still not representing the modal state.

Of course, great care should be taken in interpreting the results of this application of the LMM. Each crisis has idiosyncratic characteristics, which implies that different periods of stability may be required to detect different crises. This topic remains for further study.

Other model characteristics also emerge from the analysis for predicting the end of the crisis that occured prior to May 2003. Figures 4 and 5 compare the original time-series with respect to the LMM estimate derived from the whole data set and the estimate of a LMM with 7 latent states applied to the data from January 5, 1990–August 22, 2003. The return means of the LMM estimate, based on the shorter time series, differ slightly from the means of the overall LMM estimated time series. Nevertheless, latent state memberships derived from the shorter time series are almost the same as the LMM estimates achieved with the entire data set. In particular, the observations from March 1, 1991–March 14, 1997, can be allocated to the stable regime (latent state 4, Table 3) in both data sets, in support of the robustness of the LMM classification procedure and its power to detect low-volatility stable periods without referring directly to the analysis of any volatility measure. However, after only five weeks, the LMM can identify the beginning of a "potential" stable period; it classifies the previous 12 weekly return observations into latent state 5. Despite a low transition probability (p_{54} = 0.0116 in Table 4), this latent state is the regime that the stock market experiences just before switching to the stable latent state 4.[3]

This feature underlines an interesting behaviour of the S&P 500, which tends to stabilize and consolidate after a positive regime. Our analysis instead shows that once the stable market phase ends, instability occurs for a quite long period in the three crises in our data set. For instance, the high-volatility period III in Table 1 has approximately the same length of the stable period II. This feature can be generalized to other crises.

INSERT FIGURES 4 AND 5 ABOUT HERE

4.5. Predictive power of LMM

In Section 4.3, we reported on the latent transition matrix for the S&P 500 (Table 4). In this section, we exploit the information provided by the transition probabilities to evaluate the forecasting accuracy of the LMM. In particular, we investigate the power of the model to predict the next discrete latent state, using a one-step ahead dynamic forecast. More formally, we evaluate $\hat{y}_{T+h|T+h-1}$, the out-of-sample forecast of observation T + h, given the LMM

[3] However, p_{54} is the highest transition probability with respect to the other transition probabilities p_{j4} for $j \neq 4$

estimate prior to time $T + h - 1$. The *ex-ante* estimate refers to the transition probabilities in matrix P. If the observation at time T has been classified into latent state k by the model, the observation at time $T + 1$ will be classified into latent state j with a specific probability $\hat{p}_{jk}$. After collecting the observation for time $T + 1$ and reestimating the model, as a proxy of forecast error, we compare *ex-post* the prediction based on the highest p_{jk}s with the actual classification of the observation $y_{T+1} = j$. We first must prevent transition probabilities from changing over time [21] and require that the latent Markov chain be homogenous [29]. We thus show that the LMM can forecast the next week's market regime accurately.

As Table 4 shows, some regime switching can be predicted quite accurately, because of the high transition probabilities. The persistence of the stable regime is highly predictable, as is the switching from latent state 1 to state 6. However, for some latent states, at least three transition probabilities are greater than 0.10, which complicates our prediction. For example, latent state 2 has four transition probabilities higher than 0.10.

The LMM, developed on the weekly price index of the S&P 500 from January 5, 1990–January 1, 2010, applies to predict weekly index regimes during the period from January 1, 2010–April 2, 2010, with the forecasting results summarized in Table 5. We report the one-step ahead forecasts for the out-of-sample observations, starting from the last observation of the time series (January 1, 2010), which we denote as T, to observation $T + 13$, which corresponds to April 2, 2010. The second column of Table 5 reports the actual return observations z_{T+h} of the S&P 500 index from January 1, 2010 to April 2, 2010, for $h = 0, 1, \dots, 13$. The third colomn of Table 5 shows the latent state $\hat{y}_{T+h} = j$ obtained by estimating the LMM up to observation $T + h$, for $j = 1, \dots, 7$; in the fourth column, we provide the relative conditional mean $\hat{\mu}(z_{T+h} \mid y_{T+h} = j)$. Finally, the p_{jk} column shows the transition

(see Table 4), and latent state 5 is the last visited regime before the switch to latent state 4 in both cases in the analyzed data set.

probability of that particular switch, $\hat{f}(y_{T+h} = j \mid y_{T+h-1} = k)$. Therefore, all the switches are predicted according to the most probable p_{jk} in latent transition matrix P, except for observation $T + 5$, for which p_{35} is the second highest transition probability for latent state 3.

INSERT TABLE 5 ABOUT HERE

We can assess the forecasting accuracy of the LMM by referring to the in-sample one-step ahead predictions. In Table 6, we report the number of times the LMM predicted the next week market regime correctly for the sample period, according to the four highest latent transition probabilities for each state. Column 1 indicates the number of times the LMM predicts the next market regime by referring to the most probable $\hat{p}_{jk}$ in the latent transition matrix, column 2 contains the count of times the LMM forecasts correctly, according to the second modal transition probability, and so on. For example, the December 25, 2009, observation was classified into latent state 5, whereas the January 1, 2010, observation was in state 3. The transition probability of switching from state 5 to state 3 is p_{53} = 0.377, the second highest probability for latent state 3, following p_{55}, as in column 2 of Table 6. The last column of Table 6 shows the number of times the model was unable to predict the next week's regime by referring to the four most probable latent transition probabilities. The percentage of column "-," or the proportion of times that LMM failed for any reason to predict the week market regime, is consistently low: 0.29%. The percentages in column 1 are higher, and the model prediction accuracy (columns 1 and 2) exceeds 95%.

INSERT TABLE 6 ABOUT HERE

5. Discussion and conclusions

We have investigated the dynamic patterns of stock markets by exploiting the potential of the LMM for defining different market regimes and providing transition probabilities for regime-switching. On the basis of the AIC, we find empirical evidence for a LMM with seven latent states for the U.S. S&P 500 index. The regimes, represented by the seven latent states, are clearly defined and characterized by different return means. Therefore, we show that stock markets can be analyzed by referring to a simple and flexible model specification with a specific focus on conditional means that differ significantly and substantially across latent states. Our approach represents an efficient alternative to the more sophisticated but much less flexible Markov-switching models that attempt to evaluate the conditional variance without estimating more than four latent states [23] and without consideration of conditional means [2].

The LMM endogenously detects crises, including the 1990–91 U.S. recession, the turmoil of 1997–99 and 2000–01, and the crisis that started in late 2007. It also detects two long, stable periods between these crises. A stable market regime is defined by a particular latent state, characterized by a moderate positive return mean and a high state persistence probability, comparable to the low-volatility regime achieved in volatility-based Markov-switching models. Furthermore, the model distinguishes relatively moderate fluctuations in stable periods from stronger fluctuations during periods of crisis. With respect to volatility, our approach describes the fluctuations during high-volatility periods with six latent states and therefore enhances understanding of crises, in terms of switching between regimes with low and high (conditional) mean returns. That is, the LMM provides straightforward insights into high-volatility regimes, which cannot be achieved by Markov-switching volatility approaches that are useful for defining periods characterized by a high conditional variance value but cannot investigate fluctuations within these periods.

With regime characterization and latent transition probabilities, we can achieve two additional important goals. First, with LMM, we recognize the beginning of stable periods within 13 weeks. This feature may provide a highly pertinent opportunity to detect the end of the current financial crisis that started in 2007. Despite some preceding positive weeks, the crisis had not ended by January 2010 or April 2010 (according to the out-of-sample forecast in Section 4.5). It also enables us to recognize the beginning a crisis promptly, based on a switch from the stable market phase represented in our analysis by latent state 4 to one of the other six latent states. Unstable periods last for many weeks before “bouncing back” to a new stable phase. Second, with LMM, we can predict which regime the stock market is going to experience the following week. Additional studies should apply this methodology to other periods and countries as well to determine if the latent states we have found, as well as our other findings, hold in different circumstances.

REFERENCES

[1] Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transaction on Automatic Control*, **19(6)**, pp. 716-723.

[2] Ang, A. and Bekaert, G. (2002) International asset allocation with regime shifts. *Review of Financial Studies*, **15(4)**, pp. 1137-1187.

[3] Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970) A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.*, **41**, pp. 164-171.

[4] Bauwens, L., Preminger, A. and Rombouts, J. (2007) Theory and inference for a Markov-switching GARCH model. *CIRPEE Working Paper*, **07-33**.

[5] Bulla, J. and Bulla, I. (2006) Stylized facts of financial time series and hidden semi-Markov models. *Comput Statistics and Data Analysis*, **51**, pp. 2192-2209.

[6] Cappé, O., Moulines, E., and Rydén, T. (2005) *Inference in Hidden Markov Models*. New York: Springer.

[7] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977) Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *J. R. Statist. Soc.* ***B***, **39**, pp. 1-38.

[8] Dias, J.G., Vermunt, J.K., and Ramos, S. (2010) Mixture hidden Markov models in finance research. *Advances in Data Analysis, Data Handling and Business Intelligence, Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin: Springer-Verlag, pp. 451-459.

[9] Duekel, M.J. (1997) Markov switching in GARCH processes and mean-reverting stock-market volatility. *J. of Bus. & Ec. Statist.*, **15**(**1**), pp. 26-34.

[10] Fama, E.F. (1977) *Foundation of finance: Portfolio decisions and securities prices*. Blackwell.

[11] Frühwirth-Schnatter, S., and Kaufmann, S. (2008) Model-based clustering of multiple time series. *J. of Bus. & Econ. Statist.*, **26**, pp. 78–89.

[12] Gelman, S., and Wilfling, B. (2009) Markov-switching in target stocks during takeover bids. *J. of Empirical Finance,* **16**, pp. 745-758.

[13] Haas, M., Mittnik, S., and Paolella, M. (2004) A new approach to Markov-switching GARCH models. *J. of Fin. Econometrics*, **2**, pp. 27-62.

[14] Hamilton, J.D. (1989) A new approach to the economic-analysis of non-stationary time-series and the business-cycle. *Econometrica*, **57**, pp. 357-384.

[15] Hamilton, J.D., and Susmel, R. (1994) Autoregressive conditional heteroskedasticity and changes in regime. *J. of Econometrics*, **64**, pp. 307-333.

[16] Klaassen, F. (2002) Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, **27**, pp. 363–394.

[17] Li, M.-Y.L., and Chen, C.-N. (2010) Examining the interrelation dynamics between option and stock markets using the Markov-switching vector error correction model, *J. of App. Statist.*, **37(7)**, pp. 1173-1191.

[18] Maheu, J.M., and McCurdy, T.H. (2000) Identifying bull and bear markets in stock returns. *J. of Business & Economic Statistics*, **18(1)**, pp. 100-112.

[19] Marcucci, J. (2005) Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, **9**, pp. 1-53.

[20] McLachan, G.J., and Peel, D. (2000) *Finite Mixture Models*. New York: John Wiley & Sons.

[21] Paas, L.J., Vermunt, J.K., and Bijmolt, T.H.A. (2007) Discrete time, discrete state latent Markov modelling for assessing and predicting household acquisitions of financial products. *J. R. Statist. Soc. A*, **170**, pp. 955-974.

[22] Ramos, S.B., Vermunt, J.K., and Dias, J.G. (2010). When markets fall down: Are emerging markets all equal? *Intern. J. Finance and Economics*, forthcoming.

[23] Rossi, A., and Gallo, G.M. (2006) Volatility estimation via hidden Markov models. *J. of Empirical Finance,* **13**, 203-230.

[24] Roubini, N. (1998) Chronology of the Asian crisis and its global contagion. http://www.stern.nyu.edu/globalmacro.

[25] Rydén, T., Teräsvirta, T., and Åsbrink, S. (1998) Stylized facts of daily return series and the hidden Markov model. *J. of App. Econometrics*, **13**, pp. 217-244.

[26] Van de Pol, F., and Langeheine, R. (1990) Mixed Markov latent class models. In *Sociological Methodology 1990* (ed. C.C. Clogg), Oxford: Basil Blackwell, pp. 213-247.

[27] Vermunt, J.K., and Magidson, J. (2007) *Technical guide for Latent GOLD 4.5: basic and advanced*. Belmont, MA: Statistical Innovations Inc.

[28] Vermunt, J.K., Langeheine, R., Böckenholt, U. (1999) Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates. *J. of Educat. and Behavioral Statist.*, **24**, 178-205.

[29] Zucchini, W., and MacDonald, I.L. (2009) *Hidden Markov models for time series: An introduction using R*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Table 1: Mean, standard deviation, minimum, maximum, skewness, kurtosis, and Jarque-Bera test of S&P 500 index in different periods

Period	Mean Return	Standard Deviation	Min	Max	Skewness	Kurtosis	Jarque-Bera Test
Entire data set (1990/01/05–2010/01/01, T = 1044)	0,1381	2,3571	-18,20	12,03	-0,4873	6,0027	1608,7**
Period I (high volatility) (1990/01/01–1991/02/15, T = 59)	0,1000	2,3330	-4,98	5,39	0,1090	-0,6003	1,003
Period II (low volatility) (1991/02/22–1997/03/21, T = 318)	0,2465	1,3754	-3,42	5,02	0,0753	0,1637	0,655
Period III (high volatility) (1997/03/28–2003/05/16, T = 321)	0,0970	2,7890	-11,60	7,78	-0,2653	1,2159	23,54**
Period IV (low volatility) (2003/05/23–2007/07/20, T = 218)	0,2328	1,4353	-4,41	3,54	-0,2761	0,2753	3,458
Period V (high volatility) (2007/07/27–2010/01/01, T = 128)	-0,1724	3,8928	-18,20	12,03	-0,3969	3,6494	74,39**

**Significant at 1%.

Table 2: Log-likelihood function, number of parameters, and AIC criterion of the LMM from 1 to 8 latent states for S&P 500

Number of Latent States	*LL*	*NPar*	*AIC*
1	-2375.96	2	4755.91
2	-2329.37	6	4570.99
3	-2279.50	12	4527.28
4	-2253.64	20	4487.63
5	-2228.81	30	4456.48
6	-2207.24	42	4441.23
7	-2192.62	56	4435.94
8	-2181.97	72	4570.99

Table 3: Sizes, return means, standard deviations, and Jarque-Bera tests of 7 latent states for S&P 500 index

Latent State	Size	Return Mean (standard error)	Standard Deviation
1	.0031	-13.454 (0.844)	4.151
2	.0147	-6.349 (0.493)	1.599
3	.1983	-1.913 (0.217)	1.352
4	.5001	0.246 (0.064)	1.398
5	.2328	1.169 (0.240)	1.196
6	.0477	4.604 (0.381)	1.169
7	.0032	11.072 (0.838)	0.833
Entire data set	1.000	0.138	2.357

*Significant at 5%.

** Significant at 1%.

Table 4: Latent transition matrix for S&P 500 index

$j \setminus k$	1	2	3	4	5	6	7
1	.0067	.0065	.0065	.0065	.0065	.9606	.0066
2	.0687	.4270	.0051	.0016	.1684	.1261	.2033
3	.0001	.0149	.4457	.0002	.3581	.1809	.0001
4	.0000	.0000	.0044	.9951	.0003	.0001	.0000
5	.0084	.0018	.3774	.0116	.5921	.0086	.0001
6	.0005	.0850	.3656	.0016	.4446	.1023	.0004
7	.0065	.2882	.6633	.0065	.0145	.0145	.0065

Table 5. One-step ahead forecast for S&P 500 index, January 1, 2010–April 2, 2010 (out-of-sample)

Obs.	Date	Observed Index Return	Estimated Latent State	Conditional Mean	$\hat{p}_{jk}$
T	01/01/10	-1.01	3	-1.913	-
$T+1$	01/08/10	-0.78	3	-1.913	.4457
$T+2$	01/15/10	-3.90	3	-1.913	.4457
$T+3$	01/22/10	-1.64	3	-1.913	.4457
$T+4$	01/29/10	-0.72	3	-1.913	.4457
$T+5$	02/05/10	0.87	5	1.169	.3581
$T+6$	02/12/10	3.13	5	1.169	.5921
$T+7$	02/19/10	-0.42	5	1.169	.5921
$T+8$	02/26/10	3.10	5	1.169	.5921
$T+9$	03/05/10	0.99	5	1.169	.5921
$T+10$	03/12/10	0.86	5	1.169	.5921
$T+11$	03/19/10	0.58	5	1.169	.5921
$T+12$	03/26/10	0.99	5	1.169	.5921
$T+13$	04/02/10	1.38	5	1.169	.5921

Table 6: Number and percentages of correct LMM predictions of next latent state according to the four highest transition probabilities (in-sample forecast)

	1	2	3	4	-	Total
#	806	186	42	6	3	1043
%	77.28	17.83	4.03	0.58	0.29	100

Figure 1: S&P 500 weekly return distributions from January 5, 1990–January 1, 2010

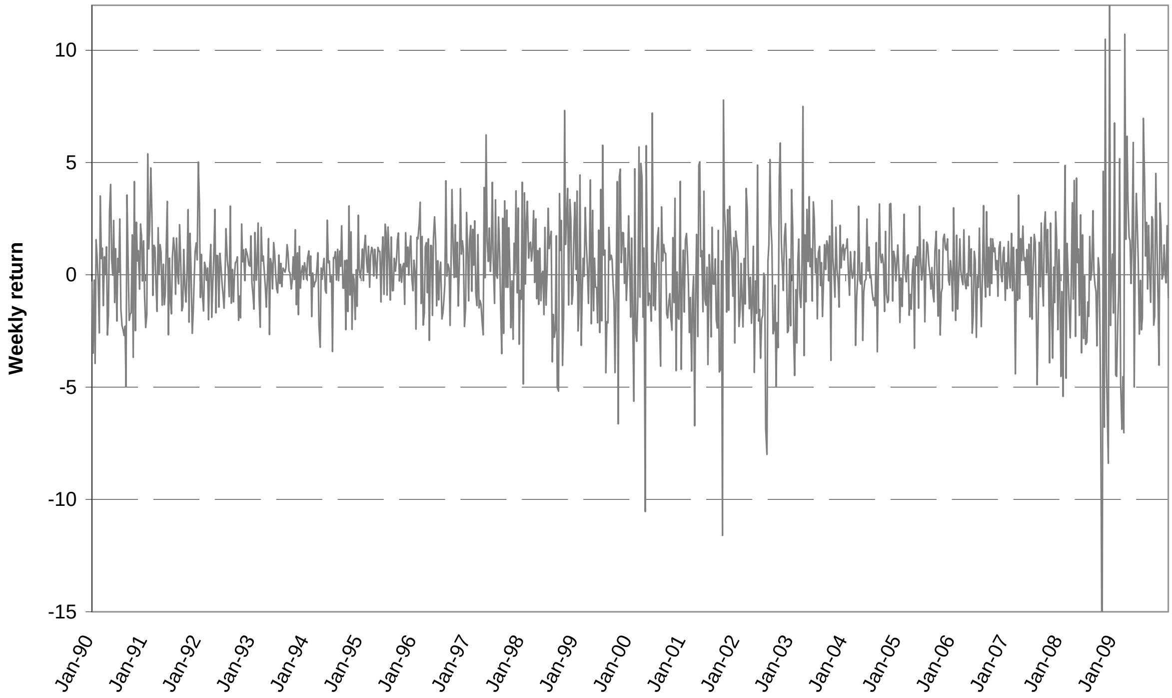

Figure 2: S&P 500 and LMM estimated time series

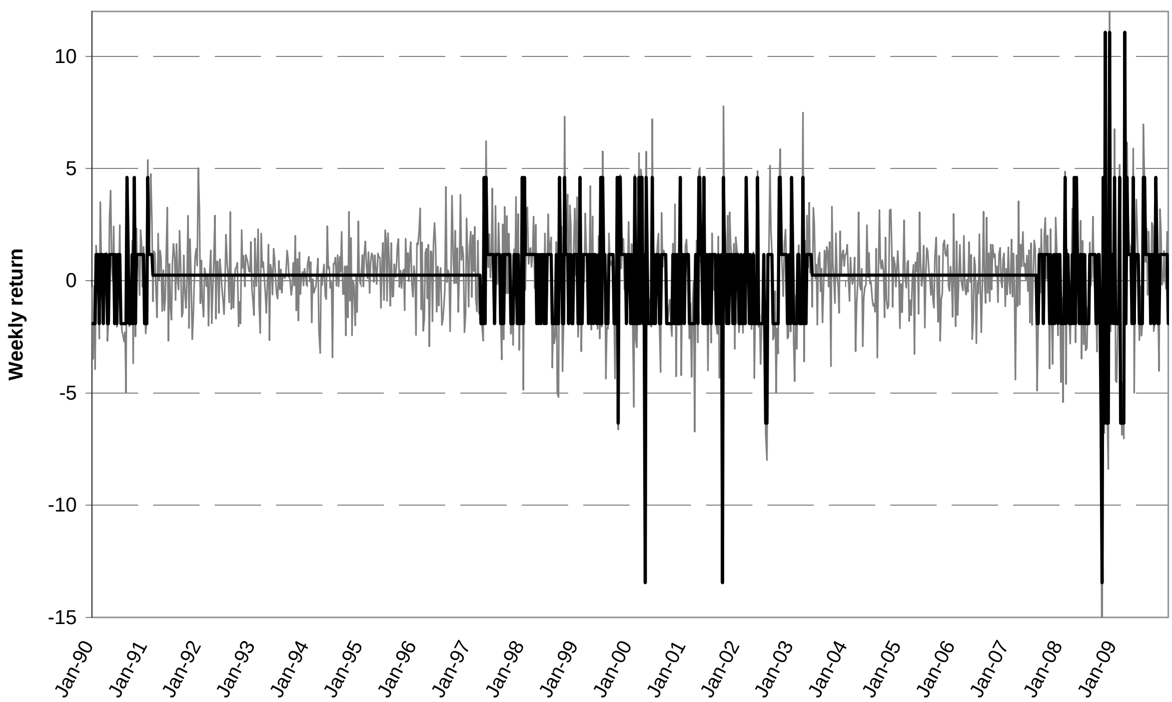

Figure 3: Estimated posterior probabilities for latent state 4

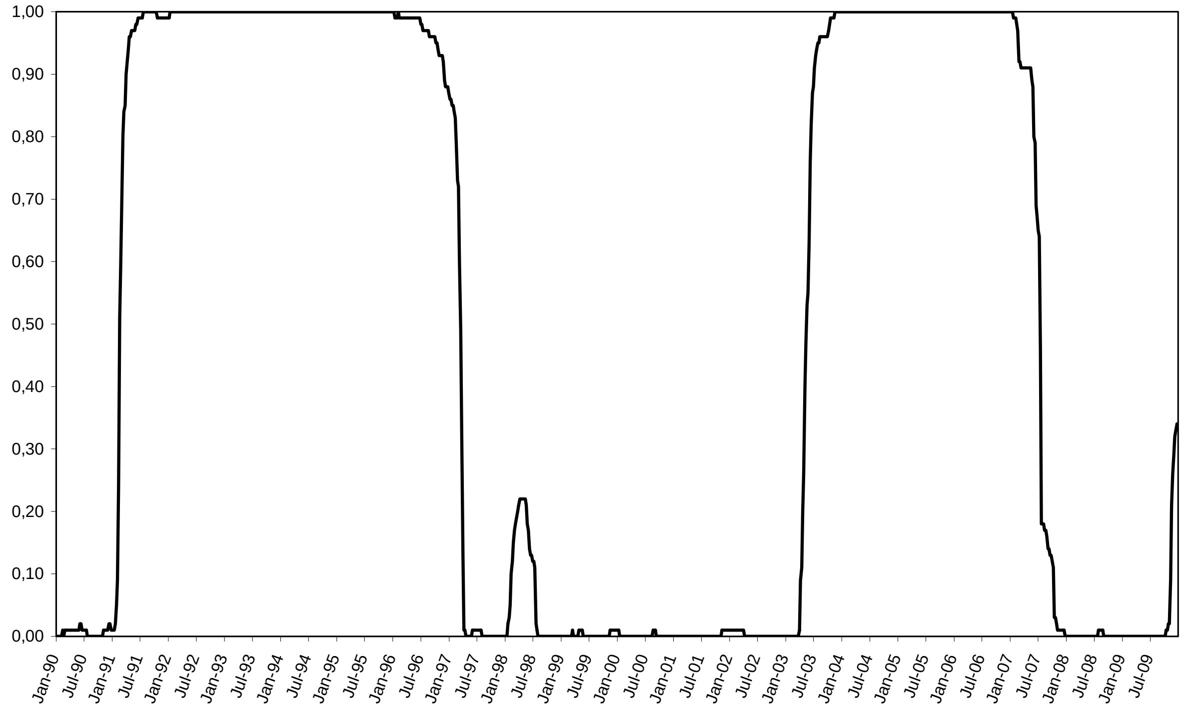

Figure 4: S&P 500 index return distribution, overall LMM estimates, and LMM estimates for the stable regime (LMM stable)

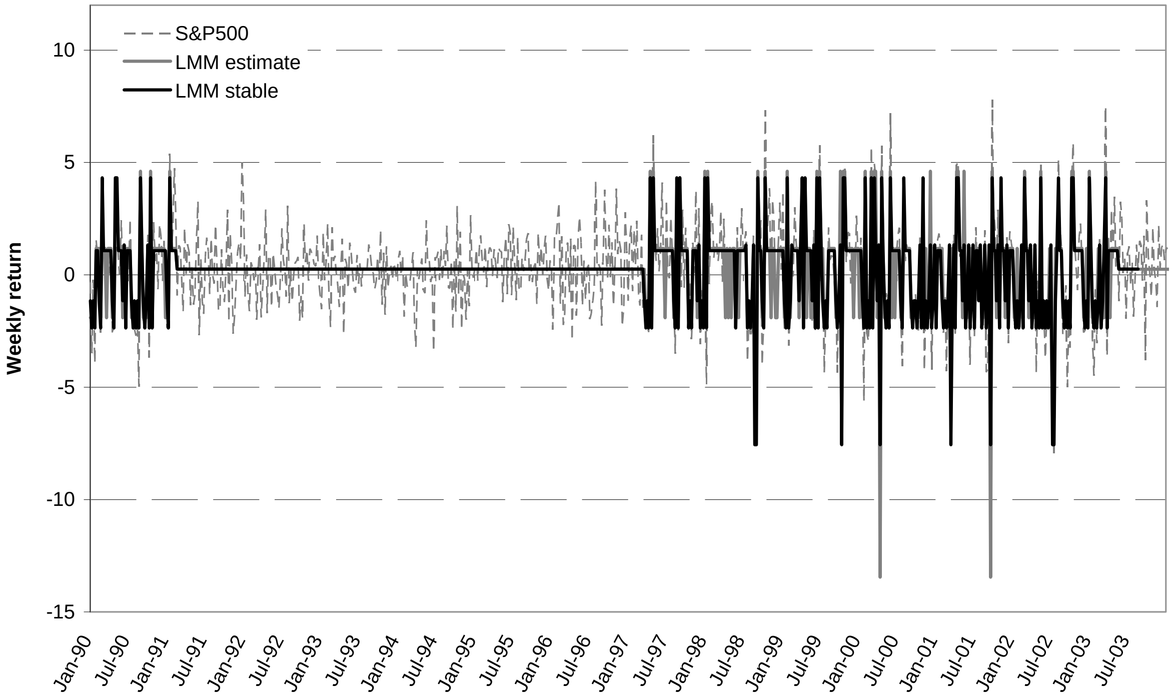

Figure 5: Close-up of S&P 500 index return distribution, overall LMM estimates, and LMM estimates for the stable regime (LMM stable)