CREATES Research Paper 2010-51

A Duration Hidden Markov Model for the Identification of Regimes in Stock Market Returns

Christos Ntantamis

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

A Duration Hidden Markov Model for the Identification of Regimes in Stock Market Returns

Christos Ntantamis *

CREATES, University of Aarhus

August 25, 2010

Abstract

This paper introduces a Duration Hidden Markov Model to model bull and bear market regime switches in the stock market; the duration of each state of the Markov Chain is a random variable that depends on a set of exogenous variables. The model not only allows the endogenous determination of the different regimes and but also estimates the effect of the explanatory variables on the regimes' durations. The model is estimated here on NYSE returns using the short-term interest rate and the interest rate spread as exogenous variables. The bull market regime is assigned to the identified state with the higher mean and lower variance; bull market duration is found to be negatively dependent on short-term interest rates and positively on the interest rate spread, while bear market duration depends positively the short-term interest rate and negatively on the interest rate spread.

JEL Classification Numbers: C13, C22, G1

Keywords: Hidden Markov Model, Variable-dependent regime duration, Regime Switching, Interest rate effect.

*Address: School of Economics & Management, University of Aarhus, Bartholins Alle 10,Building 1322, 8000 Aarhus C, Denmark, telephone: +45 89422140, email: *cntantamis@creates.au.dk*. The author is grateful to John W. Galbraith for his helpful comments. I also benefited from comments at seminars at McGill University, Cornerstone Research, CREATES, and participants in Econometric Society European Meeting, 2009. Center for Research in Econometric Analysis of Time Series, CREATES, is funded by The Danish National Research Foundation.

1 Introduction

Non-linearity of returns in the stock market is well established (Scheinkman and LeBaron 1989, Tong 1990, and McMillan 2001). The existence of different regimes in stock market returns has been proposed as one potential explanation for nonlinear conditional returns. Although in principle such processes might be represented as having any number of regimes, the existing theoretical and empirical literature has tended to concentrate on two regimes, the traditional 'bull' and 'bear' markets. Blanchard and Watson (1982) presented a model of stochastic bubbles that either survive or collapse in each period; in such a world, returns would be drawn from one of two distributions: surviving bubbles or collapsing bubbles. Cecchetti et al. (1990) considered a Lucas asset pricing model in which the endowment switches between high economic growth and low economic growth, and showed that such switching in fundamentals accounts for several features of stock market returns. Cecchetti, Lam, and Mark (2000) introduced belief distortions that vary over expansions and contractions and lead to systematic predictability in returns. Gordon and St-Amour (2000) examined a model in which risk-aversion changes according to an exogenous regime-switching process lead to the generation of bull and bear markets. Finally, Schaller and van Norden (1997) presented strong empirical evidence of regime switching in US stock market returns, evidence which was robust to different specifications of the nature of switching.

Where this division into regimes is of interest, there are issues regarding their identification, and the distributional characteristic of the returns, after accounting for regime switches. Moreover, information about the duration of each regime is important. Investment decisions can of course be influenced, but (less evidently) policy questions may also arise. The share of stock market wealth in total household wealth has increased over recent years, and plays an important role in the household's consumption decisions [1]. The effect of these decision, when aggregated, is a determinant of economic growth and hence is of concern for governments and central banks.[2]

There are two streams of literature related to these classes of question. One stream of the econometric literature treats parameter estimation and regime identification (and therefore, duration analysis) as two independent exercises. Under this view, the estimation of the parameters of interest is performed, and the calculations of the regime of the stock market and its durations are made with the appropriately parameterized model. The identification of the regimes is made exogenously with the help of data-based rules. A key consequence of such a modeling strategy is that the duration of cycles is simply a by-product of the parameter estimates and not an intrinsic feature of the model. Examples of this strain of literature include Neftçi (1984), who tested for asymmetries over the business cycles using a two-state Markov model; the identification of the different states was made using data-driven criteria. Diebold and Rudebusch (1990) modeled the hazard rates for the state duration using nonparametric methods, whereas Cohran and Defina (1995) considered a parametric approach to the state hazard rates using an exponential distribution. Lunde and Timmermann (2004) modeled the hazard rates for the durations of bull and bear markets, using covariates that included exogenous variables such as interest rates, and the age of the state.

None of these papers presented models that treat data generating processes and regimes jointly. For an

[1] This accounts for 24.4 % of the total household wealth in 2006, based on the Flow of Accounts of United States reported by the Federal Reserve System.

[2] Consumption accounts for 64.5% of the US Gross Domestic Product during the period 1947-2007 and for about 70 % for the period after 2001. The impact of the stock market wealth in consumption is found to be 2-5 cents increase in consumptions for every dollar increase in stock market wealth (Poterba 2000, Dynan and Maki 2001, Case, Quiegly and Shiller 2005).

analyst who wishes to model the regime switches and the expected durations, the advantages of Markov Regime Switching models are clear. Since the regimes are jointly estimated with the parameters of the data generating process, the analyst can identify them without resorting to data-based rules. Examples, following Hamilton's (1989) seminal work, include Engel (1994) who modeled exchange rate regimes, Hamilton and Susmel (1994), Gray (1996), and Dueker (1997) who discussed regime switching in GARCH models, and Schaller and van Norden (1997) for US stock market returns.

The above models provide a very simple link between the transition probabilities and the expected durations, thus do not allowing much flexibility. By assuming time-homogeneity of the transition probability matrix, the expected durations of the different regimes are different but remain constant across time. The expected duration of the different regimes should be allowed to vary across time, according to some explanatory variables that describe the underlying economic and market conditions because this will be more consistent with economic intuition. Since time-homogeneous transition probabilities models cannot capture this behavior, extensions have subsequently been introduced that allow the transition probabilities to vary across time. Diebold et al. (1994), Filardo (1994), and Schaller and van Norden (1997) allowed the regime-transition probabilities to be functions of a set of explanatory variables. Durland and McCurdy (1994), and Maheu and McCurdy (2000) considered the transition probabilities to be logistic functions of the number of periods that the process has been in that state, during the past introducing the duration-dependent Markov Switching model. Even though these approaches do not explicitly model state durations, state duration expectations are allowed to vary over time as they are now a function of the time-varying transition probabilities.

A more direct approach is proposed in this paper: that is to model explicitly the duration of each regime. Duration is now considered to be itself a random variable, with a distribution that can be modeled and its parameters estimated. Speech recognition applications have followed this approach, since it was found that standard Markov models are not adequate in order to model the duration and temporal structure of words. Ferguson (1980) considered a nonparametric approach, Russell and Moore (1985) assumed a Poisson distribution, Levinson (1986) estimated the model using a Gamma distribution, Mitchell and Jamieson(1993) suggested the use of the exponential family of distributions. The transition probability matrix is considered to be time-homogeneous with its diagonal elements to be equal to zero.

While attractive, these methods are restrictive in economic applications. Although the duration itself is now explicitly modeled, it is determined solely by the properties of the series under scrutiny without considering additional information in the analysis. A set of explanatory variables should be introduced in the duration specification in order to provide more information about its expected values, in a similar fashion to Lunde and Timmermann (2004) analysis of regimes' hazard rates.

In this paper, a Duration Hidden Markov Model structure is proposed. The Hidden Markov Model structure, which is assumed to have time-homogeneous transition probabilities, allows for the endogenous identification of regimes, i.e. the identification of the regimes and the estimation of the parameters governing their distributional characteristics are jointly obtained within the model. The novelty of this model is that it assumes the duration of each regime to be a random variable that depends on a set of explanatory variables, whose values are available at the time that the transition from one regime to another takes place, within the Hidden Markov Model framework. The difference relative to models of survival analysis is that regime durations are now not provided as part of the data set or as outcomes of data-based rules, but instead are obtained jointly with all the other parameters of the model. The functional form for the duration specification is the same across regimes, but the parameters are allowed to differ. Thus, the effect of explanatory variables in expected durations in now direct, in contrast to models of time-varying transition probabilities.

Moreover, the observations's distributions are assumed to be a finite mixture of Normal distributions, compared with the standard Markov Regime Switching literature where a single Normal distribution is used. Finite mixtures are known to provide a very flexible distributional form (MacLachlan and Peel 2000), which can accommodate many features of stock market returns (Rydén et al. 1998).

The model is applied to New York Stock Exchange (NYSE) monthly returns. The short-term interest rate and the interest rate spread, defined as the difference between long-term and short-term interest rates, are used as the explanatory variables in the duration function. The choice of these variables was made in order to capture potential information arising from the predictive value of these quantities for economic growth, and consequently future profits of listed companies.

The bull market regime was identified as the state with the higher mean and the lower variance for the returns' distribution. The bear market regime, on the other hand, had a lower mean and a higher variance. In terms of the duration specification results, the bull market's duration is found below to depend negatively on short-term interest rates and positively on the interest rate spread. The duration of bear markets is positively dependent on short-term interest rates and negatively on the interest rate spread. These results suggest that policy makers may have the ability to influence the durations of bull and bear stock markets by changing the levels of short-term interest rates.

The structure of the rest of the paper is as follows. Section 2 introduces the model and discusses the estimation procedure, the derivation of the standard errors, and how the model can be used for forecasting purposes. Section 3 discusses the data that were used for the estimation. Results from the estimation are presented in Section 4. Section 5 provides a brief conclusion and discuss some issues for further research.

2 Duration Hidden Markov Model

2.1 Model description

Consider a Hidden Markov Model (X_t, Y_t), where X_t is the hidden Markov Chain and Y_t are the observed data. HMM parameters are often symbolized collectively as $\lambda \triangleq (\pi, \mathbf{A}, \mathbf{B})$, where $\mathbf{B}$ represents the parameters corresponding to all the observation distributions $(b_j(y_t) = P(Y_t = y_t \mid X_t = j))$, π is such that its i^{th} element is $\pi_i = P(X_1 = i)$, and $\mathbf{A}$, with $A_{ij} = P(X_t = j \mid X_{t-1} = i)$ is the probability transition matrix (see Hamilton 1989, Bilmes 2006, or Rabiner 1989). The HMM parameters, $\lambda \triangleq (\pi, \mathbf{A}, \mathbf{B})$, are usually estimated with an algorithm advocated in a series of papers by Baum and co-authors (e.g. Baum et al 1970), which is known as the Baum-Welsh algorithm. This algorithm belongs to the Expectation Maximization (EM) algorithm family which was formally introduced by Dempster et al (1977), in order to feasibly estimate model parameters using maximum likelihood.

The traditional HMM implies that the probability of remaining in regime i, for a duration τ, $P(\tau \mid X = i)$ is proportional to $A_{ii}^{\tau-1}$ and so regime durations follow an exponential distribution. This approach is rather restrictive. Explicitly modeling the regime duration as a random variable following a given distribution function, as in Ferguson (1980), Russell and Moore (1985) and Levinson (1986) was introduced in order to alleviate this problem (all of these papers are related to speech recognition applications).

Another issue is that the values of the transition probabilities (A_{ii}), are derived based only on the information set defined by the observations sequence. Extending the information set with a second observation sequence will improve the model fit on the data and may have favorable implications for forecasting. There are two ways to model this effect. The first way would be to discard the time-homogeneity of the transition probability matrix and to allow its element to depend on exogenous variables as in Diebold et al. (1994), whereas Durland and McCurdy (1994) considered the transition probability matrix to be duration dependent.

The second way, which is followed in this paper, is to allow the regime durations to directly depend on a set of variables, other than the one provided by the observation sequence. This increases the information set available, and also provides intuition on how durations may change over time.

To do so, a Hidden Markov Model with time-homogeneous transition probability is used. The regime durations are assumed to be a probabilistic function of a set of exogenous variables, whose values are available to the researcher when the transition from one regime to another takes place. The effect of these variables in the duration is allowed to differ across regimes. Overall, the proposed model consists of the following elements:

(I) Hidden Markov Model Structure: $\lambda = (\pi, \mathbf{A}, \mathbf{B})$.

The elements $\pi, \mathbf{A}$ are related to the hidden Markov Chain. The only assumption that will be imposed is that $A_{ii} = 0$, that is, the chain can not return immediately to the previous regime. Regarding $\mathbf{B}$, the observation distributions space, a Mixture of univariate Normals is assumed for each to have generated the observations under each regime. The number of the mixtures is given a priori and it is assumed to be the same across all regimes.[3] Hence, for $b_j \in \mathbf{B}$, we have:

$$b_j(y_t) = \sum_{m=1}^{M} c_{jm} f\left(y_t; \mu_{jm}, \sigma_{jm}^2\right) \tag{1}$$

with

$$f\left(y_t; \mu_{jm}, \sigma_{jm}^2\right) = \frac{1}{\sqrt{2\pi}\sigma_{jm}} exp\left[-\frac{\left(y_t - \mu_{jm}\right)^2}{2\sigma_{jm}^2}\right] \tag{2}$$

The assumption of a mixture of Normals for the observation distributions increases the flexibility of the model since it captures more of the dynamics for the observation processes. It is well known, see McLachlan and Peel (2000), that such a mixture can replicate a large variety of distributional characteristics which are observed in financial data, such as fat tails and asymmetries. Moreover, Bilmes (2006) argues that an improvement of the fit of a Hidden Markov Model can be accomplished by either an increase in the number of regimes assumed, or by considering distributions for the observed values that are very flexible, such as finite mixtures of Normal distributions with many mixtures.

(II) Duration Equation Specification: $\mathbf{D}$

[3]It is known that nonidentifiability can emerge in finite mixture models due to overfitting. The employment of an extra mixture component can be dealt with in two ways: either set one of the mixing components to zero or take two component densities to be the same. In this case, setting the number of mixtures provides an upper bound to the true number of mixture components. Regardless, if the purpose if to use the mixture as flexible parameterization of a density function, this is not a problem (Geweke 2007).

The time that the Markov Chain remains at each regime, τ, is given by parametric relations, with coefficients that are different across regimes. The parameters are collectively summarized as $\mathbf{D} \triangleq \{\kappa_j, \zeta_j\}_{j=1}^{n}$. The general form of the model, for regime j, is:

$$\log \tau = \kappa_j' Z + \zeta_j W \tag{3}$$

where Z is the vector of the exogenous variables (of dimension k), including an intercept term. If the innovations are assumed to follow a standard Normal distribution, the density function for the duration τ is:

$$d_j(\tau; Z) = \frac{1}{\tau} \phi \left(\frac{\ln \tau - \kappa_j' Z}{\zeta_j} \right). \tag{4}$$

Assuming that the regime duration is determined as soon as the process enters that particular state, the values of the variables should correspond with the time the transition to this regime occurred. To express it in an alternative way, if the Markov Chain (X) has just entered into regime i at time t ($X_t = i$), then the amount of time that it will stay there, τ, will depend on the values of the exogenous variables Z at that particular time t, i.e. Z_t.

2.2 Model estimation

It was mentioned that the HMM parameters, $\lambda \triangleq (\pi, \mathbf{A}, \mathbf{B})$, are usually estimated with the Baum-Welsh algorithm. The core of the Baum-Welsh algorithm is the recursive estimation of the forward and backward probabilities (α and β respectively), which are subsequently used to determine the quantities of interest. The estimation of the proposed model is based on an extension of Levinson(1986) methodology. Although that paper modeled durations according to a gamma distribution,, the recursive equations for the $\pi, \mathbf{A}$ parameters remain the same as in the parametric specification approach of this paper. Denote the observations random vector by $Y_{1:T}$, and by $X_{1:T}$ the sequence of the unobserved regimes that can take n discrete values $X_t = i, i = 1, \ldots, n$ (the realizations of these random variables are denoted by $y_{1:T}, x_{1:T}$ respectively).

2.2.1 Estimation of $\pi, \mathbf{A}$.

The Forward and Backward Recursions required for the calculation of the alpha and beta probabilities are:

- Forward recursion:

$$\alpha_1(j) = \pi_j b_j(y_1) d_j(1; z_1) \tag{5}$$

$$\alpha_t(j) = \sum_{\tau \le t} \sum_{i=1, i \ne j}^{n} \alpha_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \prod_{\theta=1}^{\tau} b_j\left(y_{t-\tau+\theta}\right) \tag{6}$$

- Backward recursion:

$$\beta_T(j) = 1 \tag{7}$$

$$\beta_t(i) = \sum_{\tau \le T-t} \sum_{j=1, j \ne i}^{n} A_{ij} d_j(\tau; z_t) \prod_{\theta=1}^{\tau} b_j\left(y_{t+\theta}\right) \beta_{t+\tau}(j) \tag{8}$$

It is apparent that the main difference between these recursions and those for the simpler Hidden Markov Model is the incorporation of the time that the chain remains in the regime, which is represented by the duration probability ($d_j(\tau; Z)$) and the product of the values of the observation distribution for the data within the particular regime.

The likelihood that the observed sequence is generated according to the model is then given by:

$$\mathcal{L}(\lambda \mid y_{1:T}) = \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \sum_{\tau \leq T-t} \alpha_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \prod_{\theta=1}^{\tau} b_j\left(y_{t-\tau+\theta}\right) \beta_t(j) \tag{9}$$

The following quantity, corresponding to the probability that the chain is in regime i at time t, can also be determined:

$$\gamma_t(i) = P\left(X_t = i \mid y_{1:T}, \lambda\right) = \frac{\alpha_t(i)\beta_t(i)}{\mathcal{L}\left(Y_{1:T} \mid \lambda\right)} = \frac{\alpha_i(t)\beta_i(t)}{\sum_{j=1}^{n} \beta_t(j)\alpha_t(j)} \tag{10}$$

Using the above relations, the estimator for the elements of the transition matrix $\mathbf{A}$ is:

$$\hat{A}_{ij} = \frac{\sum_{t=1}^{T-1} \sum_{\tau \leq t} \alpha_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \prod_{\theta=1}^{\tau} b_j\left(y_{t-\tau+\theta}\right) \beta_t(j)}{\sum_{t=1}^{T-1} \alpha_t(i)\beta_t(i)} \tag{11}$$

2.2.2 Estimation of B.

The space $\mathbf{B}$ contains the observation distributions that are assumed to be mixtures of normals. Before proceeding to the recursive formulae for the parameters, the following quantity must be introduced:

$$\gamma_t(j,k) = \left[\frac{\alpha_j(t)\beta_j(t)}{\mathcal{L}\left(Y_{1:T} \mid \lambda\right)}\right] \left[\frac{c_{jk} f\left(y_t; \mu_{jk}, \Sigma_{jk}\right)}{\sum_{m=1}^{M} c_{jm} f\left(y_t; \mu_{jm}, \Sigma_{jm}\right)}\right] \tag{12}$$

which is the probability of being at regime j at time t with the k^{th} mixture accounting for y_t. The term $\gamma_t(j,k)$ generalizes to the standard $\gamma_t(j)$ in the case of a simple mixture (or a discrete density).

Moreover, the following weighting function, the ratio of the probability that the $Y_{t-\tau+\theta}$ observation is generated by the k^{th} element of the j^{th} mixture distribution to the overall probability that it is generated by the j^{th} mixture distribution, is required for the estimating equations of the means and variances of the mixtures:

$$w_\theta = \left[\frac{c_{jk} f\left(y_{t-\tau+\theta} \mid \mu_{jk}, \sigma_{jk}^2\right)}{\sum_{m=1}^{M} c_{jm} f\left(y_{t-\tau+\theta} \mid \mu_{jm}, \sigma_{jm}^2\right)}\right]. \tag{13}$$

(a) Mixing probabilities:

$$\hat{c}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(j,k)}{\sum_{t=1}^{T} \sum_{m=1}^{M} \gamma_t(j,m)} \tag{14}$$

(b) Mean of a mixing distribution:

$$\hat{\mu}_{jk} = \frac{\sum_{t=1}^{T-1} \sum_{\tau \leq t} \alpha_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \prod_{\theta=1}^{\tau} b_j\left(y_{t-\tau+\theta}\right) \beta_t(j) \sum_{\theta=1}^{\tau} w_\theta y_{t-\tau+\theta}}{\sum_{t=1}^{T-1} \sum_{\tau \leq t} \alpha_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \prod_{\theta=1}^{\tau} b_j\left(y_{t-\tau+\theta}\right) \beta_t(j) \tau \sum_{\theta=1}^{\tau} w_\theta} \tag{15}$$

(c) Variance of mixing distribution:

$$\hat{\sigma}_{jk}^2 = \frac{\sum_{t=1}^{T-1}\sum_{\tau\le t}\alpha_{t-\tau}(i)A_{ij}d_j(\tau;z_{t-\tau})\prod_{\theta=1}^{\tau}b_j\left(y_{t-\tau+\theta}\right)\beta_t(j)\sum_{\theta=1}^{\tau}w_\theta\left(Y_{t-\tau+\theta}-\mu_{jk}\right)^2}{\sum_{t=1}^{T-1}\sum_{\tau\le t}\alpha_{t-\tau}(i)A_{ij}d_j(\tau)\prod_{\theta=1}^{\tau}b_j\left(y_{t-\tau+\theta}\right)\beta_t(j)\tau\sum_{\theta=1}^{\tau}w_\theta} \tag{16}$$

Compared with Hidden Markov Models without explicit duration specification, the difference in the estimation formulae lies in considering the sum of the observations only for the period that the hidden chain stays in the particular regime. This is required for the correct determination of the mean; if the state does not change, more observations are generated according to its observation distribution and thus should be incorporated in the mean and variance estimates. Given the mixture approach, the observations also need to be weighted according to their probability of belonging to the particular element mixture (w_θ). The existence of the duration τ is included in the denominator (the number of the extra observations) in order to average over the extra observations due to the chain remaining in the same state. The number of these observations should also be weighted accordingly using the weighting function w_θ, so as to find the effective number of observations, within each duration τ, that have been generated by the particular element of the mixture distribution.

2.2.3 Estimation of the parameters for the duration model $d_j(\tau)$.

Imposing the bull market regime to be the regime with the higher mean, identification of the remaining parameters is assured. The parameters to be estimated are the regression coefficients κ_j and the standard deviation ζ_j. The estimating equations will be derived by taking the derivatives of the complete likelihood ($\mathcal{L}\left(Y_{1:T}\mid\lambda\right)$ with respect to each of the parameters and setting them equal to zero. Thus,

$$\frac{\vartheta\mathcal{L}}{\vartheta\kappa_j} = 0 \tag{17}$$

$$\frac{\vartheta\mathcal{L}}{\vartheta\zeta_j} = 0 \tag{18}$$

which will yield, using (9), the following relations:

$$\sum_{t=1}^{T}\sum_{\tau\le t}\sum_{i=1,i\ne j}^{n}\alpha_{t-\tau}(i)A_{ij}\frac{\vartheta d_j(\tau;z_{t-\tau})}{\vartheta\kappa_j}\prod_{\theta=1}^{\tau}b_j\left(y_{t-\tau+\theta}\right)\beta_t(j) = 0 \tag{19}$$

$$\sum_{t=1}^{T}\sum_{\tau\le t}\sum_{i=1,i\ne j}^{n}\alpha_{t-\tau}(i)A_{ij}\frac{\vartheta d_j(\tau;z_{t-\tau})}{\vartheta\zeta_j}\prod_{\theta=1}^{\tau}b_j\left(y_{t-\tau+\theta}\right)\beta_t(j) = 0 \tag{20}$$

Notice that the second summation over the states disappears; the duration model for each regime j depends only on the parameters κ_j,ζ_j. The summation over time is due to fact that the derivative over a product is taken. Thus, in order to write the estimating equations for the parameters, we need to write down the partial derivatives of the duration distributions with respect to those, i.e. $\frac{\vartheta d_j(\tau;Z)}{\vartheta\kappa_j}$ and $\frac{\vartheta d_j(\tau;Z)}{\vartheta\zeta_j}$:

$$\frac{\vartheta d_j(\tau;Z)}{\vartheta \kappa_j} = \frac{\ln\tau - \kappa_j' Z}{\zeta_j^2}\frac{1}{\tau}\phi\left(\frac{\ln\tau - \kappa_j' Z}{\zeta_j}\right) Z \tag{21}$$

$$\frac{\vartheta d_j(\tau;Z)}{\vartheta \zeta_j} = \frac{\left(\ln\tau - \kappa_j' Z\right)^2}{\zeta_j^3}\frac{1}{\tau}\phi\left(\frac{\ln\tau - \kappa_j' Z}{\zeta_j}\right). \tag{22}$$

We notice that the parameters of interest can not be isolated from the derivative functions so as to obtain the estimating equations as a part of the Baum-Welsh iterative procedure. Thus, a numerical estimation will be necessary so as to retrieve the parameter values κ_j, ζ_j and use them as inputs to the iterative algorithm. This is the reason that it is not plausible to use Yu and Kobayashi (2003) efficient algorithm for estimating an explicit duration Hidden Markov Model. In their paper, a more efficient algorithm to estimate duration HMM models is proposed but it requires the duration probabilities to be discrete and to be estimated as part of the iterative procedure.

2.2.4 Numerical Instability

Although the recursive formulae given in the previous sections are correct, they suffer from the numerical problem that $\alpha_t(i), \beta_t(i)$ tend to zero as t increases. The reason is that in the iterations presented, the quantities are recursively multiplied with lower than unity variables. If the sample size is large, then this quantity will tend to zero. Thus, it is necessary to introduce adjustments to the formulae to keep these probabilities within the limited dynamic range of the computer.

This paper follows the Devijver and Dekesel (1988) approach to handling this issue. They proposed to replace the joint probabilities in the definitions of the forward and backward probabilities with the *a posteriori*, probabilities $(\alpha_t^{'}(i), \beta_t^{'}(i),)$ i.e

$$\alpha_t^{'}(i) = P\left(X_t = i \mid y_{1:T}\right) \tag{23}$$

$$\beta_t^{'}(i) = \frac{P\left(y_{t+1:T} \mid X_t = i\right)}{P\left(y_{t+1:T} \mid y_{1:T}\right)} \tag{24}$$

This replacement leads to the following new recursive relations:

$$\alpha_t^{'}(i) = \frac{\alpha_t(i)}{\sum_{j=1}^{n} \alpha_t(j)} \tag{25}$$

$$\beta_t^{'}(i) = \frac{\beta_t(i)}{\sum_{j=1}^{n} \alpha_t(j)} \tag{26}$$

These adjusted forward and backward probabilities should be used in the forward and backward recursions. This weighting scheme results in the following changes in the recursions:

- Forward recursion:

$$\alpha_t(j) = \sum_{\tau \le t} \sum_{i=1, i\neq j}^{n} \alpha_{t-\tau}^{'}(i) A_{ij} d_j(\tau; z_{t-\tau}) \left[\prod_{\theta=1}^{\tau-1} q_{t-\tau+\theta} b_j\left(y_{t-\tau+\theta}\right)\right] b_j\left(y_t\right) \tag{27}$$

- Backward recursion:

$$\beta_t(i) = \sum_{\tau \leq T-t} \sum_{j=1, j \neq i}^{n} A_{ij} d_j(\tau; z_t) \left[\prod_{\theta=1}^{\tau-1} q_{t+\theta} b_j\left(y_{t+\theta}\right) \right] b_j\left(y_{t+\tau}\right) \beta'_{t+\tau}(j) \tag{28}$$

where:

$$q_t = \left[\sum_{j=1}^{n} \alpha_t(j) \right]^{-1} \tag{29}$$

The adjusted forward and backward probabilities are subsequently used in the updating functions of the model parameters. Once more, the weighting scheme needs to be considered. For example, the updating equation for the transition probabilities becomes:

$$\hat{A}_{ij} = \frac{\sum_{t=1}^{T-1} \sum_{\tau \leq t} \alpha'_{t-\tau}(i) A_{ij} d_j(\tau; z_{t-\tau}) \left[\prod_{\theta=1}^{\tau-1} q_{t-\tau+\theta} b_j\left(y_{t-\tau+\theta}\right) \right] b_j\left(y_t\right) \beta'_t(j)}{\sum_{t=1}^{T-1} \alpha'_t(i) \beta'_t(i) / q_t} \tag{30}$$

2.3 Computation of standard errors and Recovering the sequence of states

The standard errors of parameter estimates are obtained using the Dietz and Böhning (1996) approach [4]. This approach is based on the result that in large samples from regular models for which the log likelihood is quadratic in the parameters, the likelihood ratio and the Wald test for the significance of an individual parameter are equivalent, implying that the deviance change, which is twice the change in log likelihood on omitting one variable, say λ_i, is equal to the square of the t-statistic. Thus the standard error can be calculated as the absolute value of the parameter estimate divided by the square root of the deviance change,

$$s.e.(\lambda_i) = \frac{\mid \lambda_i \mid}{\sqrt{2\left(l_\lambda - l_{\lambda_{\neg i}}\right)}}, \tag{31}$$

where l_λ is the unrestricted log-likelihood, and $l_{\lambda_{\neg i}}$ is the log-likelihood for the model with the parameter λ_i being equal to zero.

It is important to emphasize that estimates of the covariance matrix of the MLE based on expected or observed information matrices are guaranteed to be valid inferentially only asymptotically.

The final step in model estimation is to reconstruct the sequence of the unobserved regimes. It should be clear that for all but the case of degenerate models, there is no "correct" sequence to be found. Hence, for practical situations, an optimality criterion is usually employed to solve this problem (there are several reasonable optimality criteria that can be imposed).In this paper, the identification of the bull and bear markets is done by what is known as Bayesian Segmentation. This approach assigns data points to regimes according to probabilistic arguments for the likelihood of data point(s) belonging to a particular regime.

One potential way to do this is by employing the Maximum Posterior Mode (MPM) approach. This approach reconstructs the regime sequence by allocating at each time point the regime that it is more probable, given the observations and the model, i.e.

[4] Even though there exists other potential methods, Aittokallio et al. (1999),Aittokallio and Uusipaikka (2000) and Visser et al (2000), for estimating the standard errors, their application for the model proposed in this paper is computationally prohibiting.

$$\hat{x}_t = \arg\max_{i=1,\ldots,n} \gamma_t(i), \qquad 1 \leq t \leq T \tag{32}$$

Although this criterion maximizes the expected number of correctly identified regimes, there could be some problems with the resulting state sequence (the "optimal" state sequence may, in fact, not even be a valid state sequence). This occurs because MPM determines the most likely regime at every instant, without regard to the probability of sequences of regimes. Nevertheless, since the quantity $\gamma_t(i)$ is readily available by being calculated at each iteration step, the MPM criterion is simpler to be employed.

3 Data

Monthly returns from 1953:04 to 2007:08, from the Center for Research in Security Prices (CRSP) value-weighted portfolio for the New York Stock/American Exchange, are considered for the application of the model, a total of 653 observations.

The short term interest rate, defined as the three-month treasury bill secondary market rate, and the interest rate spread, defined as the difference between the ten-year treasury constant maturity rate and the short-term interest rate, are used as the explanatory variables in the duration function specification.[5] Interest rates have been widely documented to closely track the state of the business cycle and appear to be a key determinant of stock returns at the monthly horizon (Fama and French 1989, Whitelaw 1994, Pesaran and Timmermann 1995).

The use of short-term interest rates is an attempt to capture short-run economic outlook information in state durations. On the other hand, interest rate spread, which can be interpreted as the expectation of future rates corresponding to the period of two maturities, has been proven to have a considerable predictive ability for future economic activity, especially for longer forecasting horizons (Estrella and Hardouvelis 1991, Estrella and Mishkin 1995). This permits us to incorporate a determinant in the state duration function that contains information regarding longer periods in the future.

Figure 1 plots the monthly series for the stock market returns, and Figure 2 the interest rates and the interest rate spread. Table 1 provides the descriptive statistics for all variables into consideration. [6] Focusing on the distributional properties of the returns, we confirm stylized facts about stock index returns: a) there exists asymmetry, and b) there is excess kurtosis, compared with the normal distribution. These results indicate, as has often been observed, that stock market returns deviate from normality; the Jarque-Bera test is also indicative of this observation. Therefore, the use of mixtures of normal distributions for the analysis for the observations' distributions is not unreasonable [7].

4 Results

The duration HMM (DHMM) is estimated for the case of two regimes for the stock market, i.e. we attempt to identify the bull and bear market regimes. The number of mixtures for the observation distributions, M,

[5]Data for the interest rates were obtained by the Board of Governors Federal Reserve System (FRED II).

[6]Interest rates and interest rates spread are expressed in percentages.

[7]Of course, the deviation of stock market returns from normality could also be attributed to the existence of nonlinear dynamics.

was assumed to be equal to 3. This number is considered to be sufficient for the mixture to capture the existing asymmetry and fat tails in the data.

Three models are considered for comparison purposes. The first assumes an Autoregressive model of order one (AR(1)) for the stock market returns with the innovations following a GARCH(1,1) process. The second model considers that the conditional mean of the stock market returns is a linear function of the short-term interest rates and of the interest rate spread, whereas the conditional variance is model similarly to the AR case as a GARCH(1,1). These two time series models are not related to DHMM. For model 2, the explanatory variables are used for determining the stock market returns and not the durations of the regimes as in DHMM. Regardless, these two simple models can be set as a minimum benchmark for the model to compare with.

- Model 1:

$$r_t = c_0 + c_1 r_{t-1} + \epsilon_t, \qquad \epsilon_t = \sqrt{h_t} v_t \tag{33}$$

$$h_t = b_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}, \qquad v_t \sim i.i.d(0,1) \tag{34}$$

- Model 2:

$$r_t = c_0 + c_1 i_t + c_2 s_t + \epsilon_t, \qquad \epsilon_t = \sqrt{h_t} v_t \tag{35}$$

$$h_t = b_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}, \qquad v_t \sim i.i.d(0,1) \tag{36}$$

where r_t is the stock market return, i_t is the T-Bill rate (in %), and s_t is the interest rate spread (in %).

The third model is a standard Markov Regime Switching model (MRS), with time-homogeneous probability transition matrix. Two regimes are assumed, and the bull market is assumed to be the identified regime with the higher mean. In this case, results are more comparable: in terms of the distributional characteristics of the regimes, of the reconstructed regime sequence, and of the implied expected durations.

The results from the estimation of the duration HMM and of the alternative 3 models are presented in Tables 2 to 6, and they can be summarized as follows:

- The distributional characteristics for each regime, as implied by the estimated mixtures, are reported in Table 4. The bull market is identified as the regime with the higher mean and the lower variance for the observations' distribution,i.e. regime 1; the bear market is identified as having a lower mean but a higher variance compared to the bull market, i.e. regime 2. The standard deviation of the bear market is almost twice the standard deviation of the bull market: bear markets are generally associated with higher variability. The results obtained by the DHMM are similar to those obtained by the MRS. The difference in the log-likelihood values suggests that the DHMM is preferable to the MRS model.

- The results for the elements of the mixtures are reported in Table 3. Although the assumption of a finite mixture of Normal distribution was made in order to allow capturing potential asymmetries or fat tail behavior of the data within each state, the implied distributional characteristics for each regime (Table 4) suggest that this was not necessary. The estimated mixtures do not exhibit asymmetries or fat tails; the Jarque-Bera statistic fails to reject normality at any conventional significance level. That is, the Hidden Markov structure itself seems to have captured adequately the asymmetries and fat-tails presented in financial data; these features follow from the estimated representation. Thus, simpler distributions, like simple Normals, could have been assumed for the observations' distributions. This is also in line with the existing Regime Switching literature where the data nonlinearity is mostly captured by the Regime Switching structure. Regardless, as it was suggested in previous sections, the value of M sets an upper bound for the number of mixtures, and overfitting is not a problem when the just density is of interest. However, it is true that this adds increased computational burden in the model. Regardless, it is the author's view that it is safer to always perform the estimation using a mixture. Furthermore, if estimation takes place using a single Normal distribution instead of a mixture, the results regarding the distributions of the regimes and the duration parameters are very similar.

- Estimation results for Models 1 and 2 are tabulated in Table 6. The explanatory variables entering the conditional mean specifications are found to be insignificant, but there is strong evidence in favor of the conditional variance specification. Regardless, the difference with the DHMM in terms of the log-likelihood is large, pointing to DHMM being a more favorable model than Model 1 or Model 2.

- The estimates for the duration parameters suggest adverse effects for the short-term interest rates and the interest rate spread. When the state enters a bull market regime, higher short-term interest rates will have a negative effect in the state duration. This may be explained by changes in portfolio holdings with shifts towards the risk-free asset when its returns are increasing. By contrast, a higher interest rate spread, which has been advocated as a signal for future economic activity, points to a larger duration for the bull market. The same reasoning applies to the effects of these variables in the bear market duration. When the future economic outlook is favorable, as suggested by the higher interest rate spread, there is a positive impact on the stock market in the long-run, leading to a reduction in the duration of the bear market regime.

- The estimated quantitative effects of the short-term interest rate and the interest rate spread on the duration of the different regimes are: a) for bull markets, an increase in the T-Bill rate of 1% will lead to a 21% reduction in the state duration, whereas a 1% increase in the spread will lead to a 24% increase in the duration, and b) for the bear markets,an increase in the T-Bill rate of 1% will lead to an approximately 10% increase in the duration, whereas a 1% increase in the spread will lead to a 6% decrease in the duration. For example, a 25 basis point reduction in the federal funds rate that will translate in an approximate reduction of 25 basis point in the T-Bill rate, all other things being equal, will have prolonged an expected bull market duration by a total of approximately 8.5% [8]. On the other hand, such a decrease will lead to a decrease by only 4% of the duration of a bear market. Thus, a decrease in the short-term interest rate has a more pronounced effect in the duration of the bull market compared to the bull market.

The reconstructed state sequence is depicted in Figure 3. The positive bar corresponds to the bull market regime, whereas the negative bars correspond to the bear market regime. In the same figure, two more depictions of the bull and bear markets exists. The first is the filtered probabilities of the bull market regime according to the Markov Regime Switching Model. A value higher than 0.5, which is the usual threshold in

[8] The total effect is the sum of 6%, which is the direct effect in the change in the short-term interest rates assuming the change in the FED rate pass completely in the 3-month T-Bill rate, and of 2.5%, which is the indirect change in the interest rate spread.

the literature, indicates the respective data point belongs to the bull market regime. The second corresponds to the Lunde and Timmermann (2004) data-based filtering for the identification of the bull and bear stock market. Following Maheu, McCurdy and Song (2010), this is employed so as to compare the reconstructed regime sequence of my model with the sequence provided by a data-based algorithm that does not allow for endogenous regime identification. The algorithm was implemented with a common threshold value for peaks and troughs of 0.2 (see the Appendix for a more detail description of their algorithm).

Given Figure 3, the following remarks can be made:

- The resulting bull and bear stock market regimes are in line with the general consensus. For example, it captures the bear market after the first oil crisis, the bull market after the beginning of 80's coinciding with President Reagan's administration and the subsequent economic growth. Moreover, the boom period for the stock market during the 90's is also well captured by the model. The bear stock market in the middle of the 80's contains the October 1987 stock market crash and the aftermath. However, it is interesting to notice that the stock market entered in a bear regime in the beginning of 1987, even before the crash. One more consideration should be noted. Even though the particular state sequence is one of many possible sequences (see discussion in section 2.5) and it depends on the choice of the Bayesian Segmentation method, it is still the outcome of the model's identification of the characteristics of the bull/bear markets. Different outcomes will have yielded different paths.

- Compared to the identified regimes obtained from the simple Markov Regime Switching model, there seem to be small difference. Regardless, the expected durations according to the MRS are 30 months for the bull market regime and 17 for the bear market, whereas for the case of the DHMM, it is inferred that durations were, on average, 42 and 25 months for the bull and the bear market respectively.

- Compared to the Lunde and Timmermann (2004) identified markets, the DHMM suggests that during the period under examination there were more switches between regimes. Furthermore,the switches according to the data-based rule seems to occur before the switches from either the DHMM or the MRS take place.

5 Conclusion

In this paper, I have proposed a Duration Hidden Markov Model structure for modeling processes with two or more regimes. The model combines a time-homogeneous Hidden Markov Model structure, which allows for the endogenous identification of the regimes, with an explicit functional specification for the duration of each regime. Durations are considered to be random variables that depend on a set of explanatory variables, whose values are available at the time that the transition from one regime to another takes place. The functional form for the duration specification does not alter across regimes, but the parameters are allowed to differ.

The advantages of the model compared with the Markov regime switching currently dominant in the literature in economics are: a) the direct modeling of the expected duration of the different regimes by the introduction of covariate-dependent regime durations in the Hidden Markov Model, b) the analysis of the impact of these covariates in regime's durations and the corresponding policy implications, and c) the consideration of a more general distributional form for the observations' distributions.

The issues of the estimation of the parameters and the derivation of their standard errors were discussed. The numerical instability issues were addressed by the adjustment of the forward and backward recursions. The procedure for obtaining forecasts by this model, which is based on simulating a number of paths, was developed. The model was subsequently applied in New York monthly stock market returns. The short-term interest rate and the interest rate spread, i.e. the difference between long-term interest rates and short-term interest were used as the explanatory variables for the regime's duration in order to capture the short-run and the long-run economic outlook.

In applying the model, I identified a bull market regime in which returns have a higher mean and lower variance, versus a bear market regime with lower mean returns and higher variance. Both the results from the reconstruction of the historical sequence of the states, and the estimated effect on duration of the explanatory variables were in accordance with intuition. An increase in the short-term interest rate decreases the duration of the bull market regime and increases the duration of the bear market regime. On the contrary, an increase in the interest rate spread leads to a higher duration for the bull market and a lower one for the bear market. Furthermore, the model appears to be a better fit for the data when compared with alternative models, including a simple Markov Regime Switching model. The results for the regimes durations not only suggest the ability of policy makers to influence the durations of the bull and bear market regimes by changing the short-term interest rates, which also indirectly affects the interest rate spread, but provide quantitative estimates of impacts. A reduction of 25 basis points in the federal funds rate, and in effect in the T-Bill rate, is estimated to increase by 8.5% the duration of a bull market and to reduce by 4 % the duration of a bear market. This feature of the model has significant policy implications for central bankers and investors as well.

This duration model suggests that the bull/bear regime durations can be quantified, and adjustments can be made by indirectly changing the short-term interest rates via the federal funds rate. Thus, if the concern is to prevent a prolonged period of bear markets, reduction of interest rates should be eminent so as to reduce the duration of the bear regime, and more should be done more aggressively compared to a situation that the prolonging of a bull market is desirable. Nevertheless, it should be noted that only the duration can be affected; if the economy is weak, as evidenced by a negative spread the duration of a bear market may still be considerable.

On the investors' side, the importance in determining the durations of the two regimes lies in the optimal timings of their investments. It should be expected that long or short positions on the market would be taken accordingly to anticipation of entering a bull or bear regimes. The timing of such actions depends upon the durations of the regimes. In anticipation of a reduction in the interest rate by the FED so as to reduce the duration of a bear market, investors can rebalance their positions by reducing the duration of their short positions or by entering long in the market.

References

Aittokallio T., V. Ahola, and E. Uusipaikka, (1999). Likelihood Based Inference in Hidden Markov Models, *Turku Centre for Computer Science*, TUCS Technical Report 260.

Aittokallio T., and E. Uusipaikka, (2000). Computation of Standard Errors for Maximum-Likelihood Estimates in Hidden Markov Models, *Turku Centre for Computer Science*, TUCS Technical Report 379.

Baum L.E, T. Petrie, G. Soules and N. Weiss, (1970). A Maximization Technique in the Statistical Analysis of Probabilistic Functions of Finite State Markov Chains,*Annals of Mathematical Statistics*, 41, 164-171.

Bilmes J., (2006). What HMMs Can Do, *EICE Transactions on Information and Systems*, E89, 869-89.

Blanchard O.J., and M. Watson, (1982). Bubbles, Rational Expectations and Financial Markets, *Crises in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*, ed. P. Watchel, Lexington Books.

Case K.E., J.M. Quigley, and R.J.Shiller, (2005). Comparing Wealth Effects: The Stock Market vs. the Housing Market, *Advances in Macroeconomics*, vol. 5, 1.

Cecchetti S.G., P.Lam, and N.C. Mark, (1990). Mean Reversion in Equilibrium Asset Prices, *American Economic Review*, 80, 398-418.

Cecchetti S.G., P. Lam, N.C. Mark, (2000). Asset Pricing Under Distorted Beliefs: Are Equity Returns Too Good To Be True?, *American Economic Review*, 90, 787-805.

Cohran S., and R. Defina, (1995). Duration Dependence in the US Stock Market Cycle: A Parametric Approach, *Applied Financial Economics*, 5, 309-318.

Dempster A. P., N.M Laird, and D.B. Rubin, (1977). Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society, Ser. B*, 39, 1-38.

Devijver P.A., and M. Dekesel, (1998). Champs aléatoires de Pickard et modelization d' images digitales, Traitement Du Signal, 5, 131-150.

Diebold F.X., and G.D. Rudebusch, (1990). A Nonparametric Investigation of Duration Dependence in the American Business Cycle, *The Journal of Political Economy*, 98, 596-616.

Diebold F.X., J-H. Lee, and G. Weinbach, (1994). Regime Switching with Time-Varying transition Probabilities, in *Nonstationary Time Series Analysis and Cointegration*, ed. C. Hargreaves (*Advanced Texts in Econometrics*, series eds. C.W.J Granger and G. Mizon), Oxford University Press.

Dietz E., and D. Böhning, (1996). Statistical Inference Based On a General Model of Unobserved Heterogeneity, in *Advances in GLIM and Statistical Modeling*, eds Fahrmeir L., F. Francis, R. Gilchrist, and G. Tutz, (*Lecture Notes in Statistics*, Springer, Berlin.

Dueker M.J., (1997). Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility, *Journal of Business and Economic Statistics*, 15, 26-34.

Durland J.M., and T. McCurdy, (1994). Duration-dependent transitions in a Markov model of U.S. GNP growth, *Journal of Business and Economic Statistics*, 12, 279-288.

Dynan K.E., and D.M. Maki, (2001). Does Stock Market Wealth Matter for Consumption?, *Finance and Economics Discussion Series*, Board of Governors of the Federal Reserve System, 23.

Engel C., (1994). Can the Markov Switching Model Forecast Exchange Rates?, *Journal of International Economics*, 36, 151-165.

Estrella A., and G.A. Hardouvelis, (1991). The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance*, 46, 555-576.

Estrella A., and F. S. Mishkin, (1995). Predicting U.S. Recessions: Financial Variables as Leading Indicators, *National Bureau of Economic Research*, Working Paper 5379.

Fama E.F., and K.R. French, (1989). Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics*, 25, 23-49.

Ferguson J.D., (1980). Variable Duration Models for Speech, *Proceedings of the Symposium on the Application of Hidden Markov Models to Text and Speech*, ed. J. D. Ferguson, 143-179, Princeton, New Jersey.

Filardo A.J., (1994). Business-Cycle Phases and Their Transitional Dynamics, *Journal of Business and Economic Statistics*, 12, 299-308.

Geweke J., (2007). Interpretation and Inference in Mixture Models: Simple MCMC works, *Computational Statistics & Data Analysis*, 51, 3529-3550.

Gordon S., and P. St-Amour, (2000). A Preference Regime Model of Bull and Bear Markets, *The American Economic Review*, 90,1019-1033.

Gray S., (1996). Modeling the Conditional Distribution of Interest Rates as a Regime Switching Process, *Journal of Financial Economics*, 42, 27-62.

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

Hamilton J. D., (1994). *Time Series Analysis*, Princeton: Princeton University Press.

Hamilton J. D., and R. Susmel (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics*, 64, 307-333.

Levinson S.E., (1986). Continuously Variable Duration Hidden Markov Models for Automatic Speech Recognition, *Computer Speech and Language*, 1, 29-45.

Lunde A. and A.Timmermann, (2004). Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets, *Journal of Business and Economic Statistics*, 22, 253-273.

McLachlan C. and D.Peel, (2000). *Finite Mixture Models*, New York: John Wiley and Sons.

McMillan D.G., (2001). Nonlinear Predictability of Stock Market Returns: Evidence from Nonparametric and Threshold Models,*International Review of Economics and Finance*, 10, 353-368.

Maheu J., and T. McCurdy, (2000). Identifying Bull and Bear Markets in Stock Returns, *Journal of Business and Economic Statistics*, 18, 100-112.

Maheu J.,T. McCurdy, and Y.Song (2010). Components of Bull and Bear Markets: Bull Corrections and Bear Rallies, Working Paper, University of Toronto.

Mitchell C.D., and L.H. Jamieson, (1993). Modeling Duration in a Hidden Markov Model with the Exponential Family, *Proceedings of the 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, April 1993, 331-334.

Neftçi S.N, (1984). Are Economic Time Series Asymmetric over the Business Cycle?, *The Journal of Political Economy*, 92, 307-328.

Pesaran M.H., and A. Timmermann, (1995). Predictability of Stock Returns: Robustness and Economic Significance, *Journal of Finance*, 50, 1201-1228.

Poterba J.M., (2000). Stock Market Wealth and Consumption, *The Journal of Economic Perspectives*, 14, 99-118.

Rabiner L.R., (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, 77, 257-286.

Russell M. J., and R.K. Moore, (1985). Explicit Modeling of State Occupancy in Hidden Markov Models for Automatic Speech Recognition, *Proceedings of ICASSP-85*, Tampa, Florida, March 1985, 5-8.

Rydén T., T. Teräsvirta, and S. Åsbrink, (1998). Stylized Facts of Daily Return Series and the Hidden Markov Model, *Journal of Applied Econometrics*, 13, 217-244.

Schaller H., and S. van Norden, (1997). Regime Switching in Stock Market Returns, *Applied Financial Economics*, 7, 177-191.

Scheinkman J.A., and B. LeBaron, (1989). Nonlinear Dynamics and Stock Returns, *The Journal of Business*, 62, 311-337.

Tong H., (1990). *Non-linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.

Visser I., M.E.J. Raijmakers, and P.C.M. Molenaar, (2000). Confidence Intervals for Hidden Markov Model Parameters, *British Journal of Mathematical and Statistical Psychology*, 53, 317-327.

Whitelaw R.F., (1994). Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns, *Journal of Finance*, 49, 515-541.

Yu S-Z., and H. Kobayashi, (2003). An Efficient Forward-Backward Algorithm for an Explicit-Duration Hidden Markov Model, *IEEE Signal Processing Letters*, 10, 11-14.

A Lunde Timmermann (2004) data-based algorithm

This algorithm identifies bull and bear markets using a cumulative return threshold to locate peaks and troughs moving forward. Maheu, McCurdy and Song (2010) provide a more detailed exposition of the algorithm compared to the one provided in the original paper. According to their description, the algorithm can be summarized as follows:

Begin by defining a binary market indicator variable I_t. This variable will take the value 1 if the stock market is in a bull state at time t, and 0 if it is in a bear state. The stock price at the end of period t is denoted by P_t. Denote by λ_1 and λ_2 two scalars defining the thresholds of the movements in stock prices: the first triggers a switch from a bear to a bull market, and the second triggers a switch from a bull to a bear market.

Now, suppose that the stock price is at a local maximum at time t_0, in which case it is set that $P_{t0}^{max} = P_{t0}$.

1. Define the following stopping-time variables associated with a bull market as

$$\tau_{max}\left(P_{t_0}^{max}, t_0 \mid I_{t_0} = 1\right) = \inf\{t_0 + \tau : P_{t_0+\tau} \geq P_{t_0}^{max}\} \tag{37}$$

$$\tau_{min}\left(P_{t_0}^{max}, t_0 \mid I_{t_0} = 1\right) = \inf\{t_0 + \tau : P_{t_0+\tau} \leq (1-\lambda_2)\, P_{t_0}^{max}\} \tag{38}$$

2. Then $min\left(\tau_{max}, \tau_{min}\right)$ is the first time that the price process crosses one of the barriers $\{P_{t_0}^{max}, (1-\lambda_2)\, P_{t_0}^{max}\}$. Consider two cases:

   - If $\tau_{max} < \tau_{min}$ bull market continues, update the local maximum to $P_{t_0+\tau_{max}}^{max} = P_{t_0+\tau_{max}}$, and set $I_{t_0+1} = \cdots = I_{t_0+\tau_{max}} = 1$. Go to step 1 directly above.

   - If $\tau_{max} > \tau_{min}$, so that the stock price at $t_0 + \tau_{min}$ has declined by a fraction λ_2 since its local peak, a trough is found at time $t_0 + \tau_{min}$: a bear market occurred from time $t_0 + 1$ to $t_0 + \tau_{min}$, i.e. $I_{t_0+1} = \cdots = I_{t_0+\tau_{min}} = 0$. Record the value $P_{t_0+\tau_{min}}^{min} = P_{t_0+\tau_{min}}$, and mark the original time t_0 as one peak. Go to step 1 for the bear market.

Now, suppose t_0 was a local minimum $P_{t0}^{min} = P_{t0}$.

1. Similarly to before, stopping-time variables associated with a bull market can be defined according to

$$\tau_{min}\left(P_{t_0}^{min}, t_0 \mid I_{t_0} = 0\right) = \inf\{t_0 + \tau : P_{t_0+\tau} \leq P_{t_0}^{min}\} \tag{39}$$

$$\tau_{max}\left(P_{t_0}^{min}, t_0 \mid I_{t_0} = 0\right) = \inf\{t_0 + \tau : P_{t_0+\tau} \geq (1+\lambda_1)\, P_{t_0}^{min}\} \tag{40}$$

2. As before, two cases can be considered:

   - If $\tau_{min} < \tau_{max}$ bear market continues, update the trough point forward, and update the local minimum to $P_{t_0+\tau_{min}}^{min} = P_{t_0+\tau_{min}}$, and set $I_{t_0+1} = \cdots = I_{t_0+\tau_{min}} = 1$. Go to 1 directly above.

- If $\tau_{min} > \tau_{max}$ a peak is found at time $t_0 + \tau_{max}$: a bull market occurred from time $t_0 + 1$ to $t_0 + \tau_{max}$, i.e. $I_{t_0+1} = \cdots = I_{t_0+\tau_{max}} = 1$. Record the value $P^{max}_{t_0+\tau_{max}} = P_{t_0+\tau_{max}}$, and mark the original time t_0 as a trough. Go to step 1 for the bull market.

The process is repeated until the last data point. This definition of bull and bear states partitions the data on stock prices into mutually exclusive and exhaustive bull and bear market subsets based on the sequences of first passage. A range of values for λ_1 and λ_2 can be considered: smaller values will lead to more bull and bear market regimes, but if they get too small, they will tend to capture short-term dynamics in stock price movements (see Lunde and Timmermann 2004). A conventional value used in the financial press is 0.2 for both thresholds.

B Tables

B.1 Data Descriptive Statistics

	Stock Returns % (%)	**T-Bill Rate % (3 month)**	**Contract Rate % (10 year)**	**Spread %**
mean	0.996	5.124	6.475	1.351
standard deviation	4.215	2.799	2.683	1.186
skewness	-0.478	1.106	0.932	-0.044
kurtosis	5.066	6.817	3.610	-2.812
min	-22.534	0.640	2.290	-2.650
max	16.558	16.300	15.320	4.420
J-B stat	140.956	223.114	104.668	1.169

Table 1: **Descriptive Statistics**

B.2 Estimation Results

Model	DHMM		MRS	
Regime	**1**	**2**	**1**	**2**
π	0.6475	0.3525	0.6392	0.3608
A				
1	0	1	0.9671	0.0329
2	1	0	0.0583	0.9417

Table 2: **Markov Chain Results: DHMM vs Markov Regime Switching**

Regime mixture	1	2	1	2	1	2
	Mixing Prob.		**Means**		**Standard Deviations**	
1	0.276126	0.149811	2.076026	5.460774	3.804402	0.843793
2	0.190312	0.777897	2.975791	-1.08153	1.439625	5.191445
3	0.533562	0.072291	0.241948	9.492694	2.718696	4.932934

Table 3: **DHMM: Estimates of Finite Mixture of Normals**

Regime	Mean	Standard Deviation	Skewness	Kurtosis	J-B test*	Log-Likelihood
			Duration HMM			
1	1.2687	3.1035	0.0123	0.3764	0.2856	1172.114
2	0.6630	5.8552	0.0415	0.0376	0.9609	
			Markov Regime Switching			
1	1.3766	3.0843				1170.086
2	0.3410	5.6159				
	*: p-value					

Table 4: **DHMM vs MRS: Characteristics of Bull & Bear Markets**

Regime parameter	1 Coefficient	1 Std. Error	2 Coefficient	2 Std. Error
κ_{j1}	4.0182	(0.8232)	2.647	(1.4291)
κ_{j2}	-0.21	(0.1301)	0.10599	(0.2079)
κ_{j3}	0.2405	(0.4017)	-0.06702	(0.6962)
ζ_j	0.9175		0.8266	

Table 5: **DHMM: Duration parameters results**

Model 1			Model 2		
	Coefficient	**Std. Error**		**Coefficient**	**Std. Error**
Mean Equation					
Const	0.00991	0.00154	**Const**	0.00971	0.00452
AR(1)	0.05170	0.04170	i_t	-0.06408	0.06178
			s_t	0.26085	0.14578
Variance Equation					
Const	7.32E-05	3.39E-05	**Const**	7.87E-05	3.57E-05
α_1	0.101	0.0268	α_1	0.0993	0.0265
β_1	0.863	0.0312	β_1	0.8606	0.03201
Log-Likel.	1162.971		**Log-Likel.**	1163.797	

Table 6: **Time Series Models Results**

C Figures

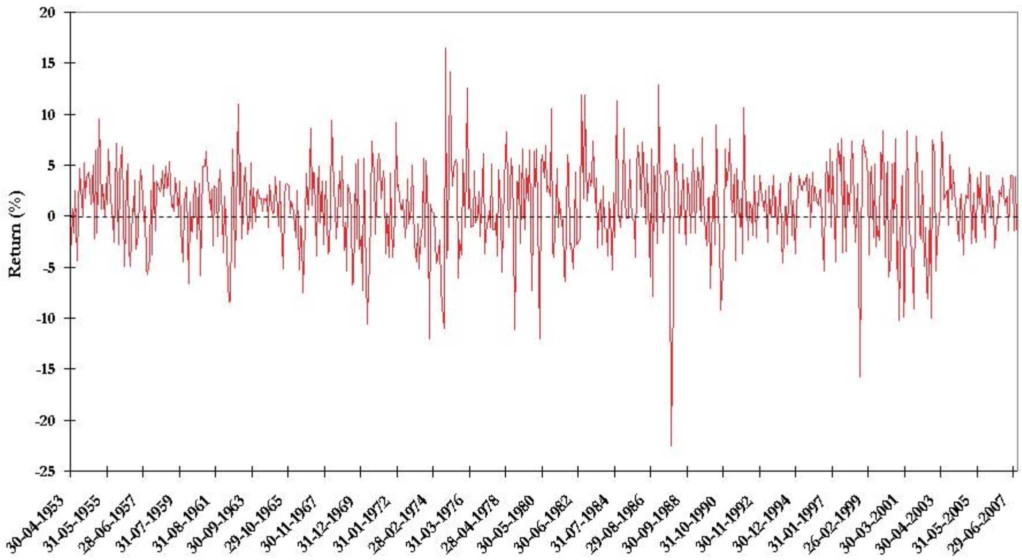

Figure 1: NYSE/AMEX Returns Plot

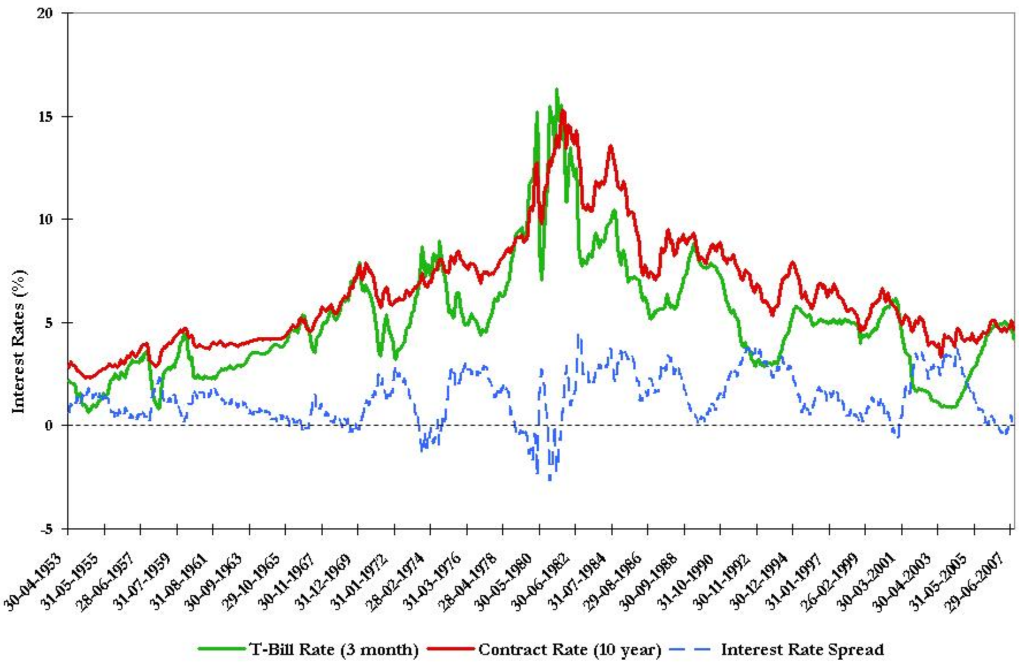

Figure 2: Interest Rates and Interest Rate Spread Plot

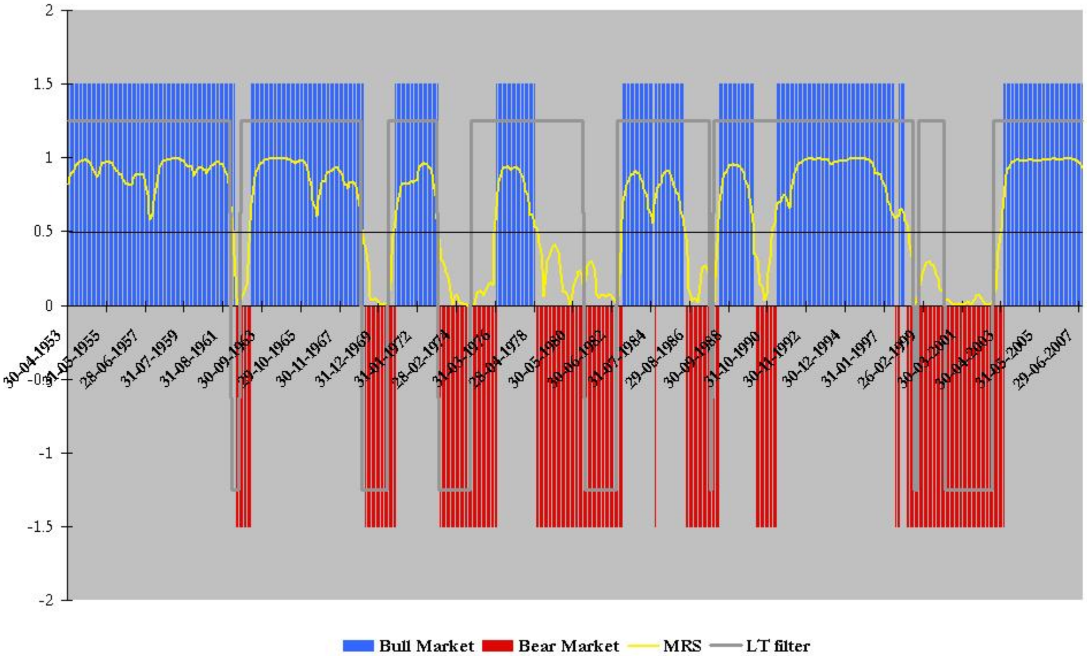

Figure 3: Reconstrunced States Sequences

Research Papers 2010

2010-37: Tom Engsted and Thomas Q. Pedersen: The log-linear return approximation, bubbles, and predictability

2010-38: Thomas Q. Pedersen: Predictable return distributions

2010-39: Rasmus Tangsgaard Varneskov: The Role of Dynamic Specification in Forecasting Volatility in the Presence of Jumps and Noisy High-Frequency Data

2010-40: Antonis Papapantoleon and David Skovmand: Picard Approximation of Stochastic Differential Equations and Application to Libor Models

2010-41: Ole E. Barndorff–Nielsen, Fred Espen Benth and Almut E. D. Veraart: Modelling electricity forward markets by ambit fields

2010-42: Robinson Kruse and Philipp Sibbertsen: Long memory and changing persistence

2010-43: Dennis Kristensen: Semi-Nonparametric Estimation and Misspecification Testing of Diffusion Models

2010-44: Jeroen V.K. Rombouts and Lars Stentoft: Option Pricing with Asymmetric Heteroskedastic Normal Mixture Models

2010-45: Rasmus Tangsgaard Varneskov and Valeri Voev: The Role of Realized Ex-post Covariance Measures and Dynamic Model Choice on the Quality of Covariance Forecasts

2010-46: Christian Bach and Stig Vinther Møller: Habit-based Asset Pricing with Limited Participation Consumption

2010-47: Christian M. Dahl, Hans Christian Kongsted and Anders Sørensen: ICT and Productivity Growth in the 1990's: Panel Data Evidence on Europe

2010-48: Christian M. Dahl and Emma M. Iglesias: Asymptotic normality of the QMLE in the level-effect ARCH model

2010-49: Christian D. Dick, Maik Schmeling and Andreas Schrimpf: Macro Expectations, Aggregate Uncertainty, and Expected Term Premia

2010-50: Bent Jesper Christensen and Petra Posedel: The Risk-Return Tradeoff and Leverage Effect in a Stochastic Volatility-in-Mean Model

2010-51: Christos Ntantamis: A Duration Hidden Markov Model for the Identification of Regimes in Stock Market Returns