Economic Modelling 44 (2015) 54–58

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Bull, bear or any other states in US stock market?

Yu Jiang [1], Xianming Fang [*]

Department of Finance and Insurance, Nanjing University, Nanjing, Jiangsu 210093, China

ARTICLE INFO

Article history:
Accepted 19 September 2014
Available online 23 October 2014

Keywords:
Markov switching model
Optimal number of states
S&P 500 returns

ABSTRACT

A stock market is traditionally considered to shift between bull and bear markets, reflecting the states of high mean and low mean in stock returns, respectively. In this paper, we attempt to detect more different states in a stock market by applying a Bayesian Markov switching model, where the optimal number of states is determined according to the marginal likelihoods. An application to US stock market indicates that there exist four distinguishable states and each state represents different characteristics of US stock market.

1. Introduction

Traditionally speaking, a stock market trend is usually considered to switch between two states, bull and bear markets. The terms bull market and bear market describe upward and downward trends of stock index or positive and negative stock index returns over a period of time, respectively. Since the switching between bull and bear markets is similar to the switching of GDP growth between expansions and contractions, methodologies that originally developed to identify the business cycle were naturally applied to identify the bull market and the bear market. There exist two main categories of methods in general, non-parametric method and parametric method.

Non-parametric method directly deals with the time series data and attempts to determine the time points of peak (market top) and trough (market bottom) in the business cycle (stock index). An early study by Bry and Boschan (1971) developed a criterion to detect peaks and troughs and the method was applied by NBER to study the US business cycle. Harding and Pagan (2002) made adjustments to this method. Applications of non-parametric method to identify the bull market and the bear market include those of Pagan and Sossounov (2003) and Edvards et al. (2003).

Parametric method develops econometric models to quantitatively study the time series. Among many models, a widely and frequently used model is the Markov switching model that allows parameter values to vary across states and models the switching mechanism between states by a first-order Markov process. The expansion (bull market) and contraction (bear market) in business cycle (stock market) exactly represent the two states of high mean and low mean in GDP growth (stock returns), respectively. Hamilton (1989) introduced a 2-state Markov switching model that allows switching in the mean parameter of GDP growth and applied it to identify the US business cycle. Maheu and McCurdy (2000) applied 2-state Markov switching models to identify bull and bear markets in US stock returns.

It is known that stock returns will usually exhibit different characteristics in the volatility besides in the mean. Non-parametric method is not applicable to identify different states in the volatility of stock returns; however, the Markov switching model provides more flexibility in this aspect. Hardy (2001) introduced a Markov switching model that allows both mean and volatility parameters to change and identified states of high volatility and low volatility in US and Canada stock markets. A more complicated model that incorporates the Markov switching model and the autoregressive conditional heteroskedasticity model was developed by Hamilton and Susmel (1994). Applications to stock returns include those of Hamilton and Lin (1996), Li and Lin (2003), and Henneke et al. (2011).

Applications of the Markov switching model to stock returns have indicated that both the mean and volatility parameters of stock returns might exhibit different states, which implies that a stock market trend can be classified into more states rather than the traditional classification of bull and bear markets. However, a critical issue needs more considerations: what is the optimal number of states? Existing literature usually specifies the number of states ex ante such as 2 states (e. g. Maheu and McCurdy, 2000) or 3 states (e. g. Henneke et al., 2011), or selects between 2 and 3 states according to AIC or likelihood ratio test (e. g. Hardy, 2001). We think it is unreasonable to fix the number of states ex ante or to just select between 2 and 3 states. A wider range for the number of states should be inspected and the optimal number of states should be determined as an issue of model selection.

In this paper, we attempt to identify distinguishable states in the US stock market using a Markov switching model. Our model focuses on the changes in the mean and volatility of stock returns and the optimal number of states in the US stock market is determined by comparing

* Corresponding author. Tel.: +86 25 83621165.
E-mail addresses: yujiang@nju.edu.cn (Y. Jiang), fxmfxm@nju.edu.cn (X. Fang).
[1] Tel.: +86 25 83621368.

http://dx.doi.org/10.1016/j.econmod.2014.09.020

models with various numbers of states. Generally speaking, the Markov switching model can be estimated using classical approach based on maximum likelihood estimation or Bayesian approach. Hamilton (2005) pointed out that classical approach may have difficulties in estimating parameters accurately in the case that the number of states is large and usually is applied to estimate the model of 2 or 3 states. Therefore, previous studies based on classical approaches can only identify 2 or 3 states in the US stock market (e.g. Hardy, 2001; Maheu and McCurdy, 2000). Moreover, the classical approach usually applies likelihood ratio test to determine the number of states, but almost all proposed tests have more or less deficiencies (Hamilton, 2005). Compared with the classical approach, the Bayesian approach turns out to be greatly facilitated by Markov chain Monte Carlo (MCMC) methods. The MCMC algorithm for Markov switching model, first developed by Albert and Chib (1993), is simple and straightforward in estimating parameters using the conjugate prior distributions, and a more efficient algorithm was introduced by Chib (1996). Furthermore, the Bayesian approach provides a standard and straightforward criterion for model selection, that is, the model with the optimal number of states should have the highest value of marginal likelihood. Considering the advantages of Bayesian approach in model estimation and model selection, we choose to estimate model parameters using the Bayesian approach.

We adopt uninformative and exchangeable prior distributions for parameters in the Markov switching model, which means that the identification of states would be completely determined by the data. We apply models with different numbers of states to monthly S&P 500 returns. Results indicate that the optimal number of states in the US stock market is 4. According to the posterior information of parameters, we analyze characteristics of the 4 states in the US stock market.

The remainder of this paper is organized as follows. Section 2 provides details of the Markov switching model. The model is applied to the US stock market in Section 3. Section 4 concludes.

2. The Markov switching model

Traditional models for stock prices assume that prices follow a geometric Brownian motion, which implies that the logarithm of stock prices follow a Brownian motion. If S_t is the stock price at time t, then the logarithmic stock return at time t, denoted by y_t, is normally distributed:

$$y_t = \log\left(\frac{S_t}{S_{t-1}}\right) \sim \mathrm{N}\left(\mu, \sigma^2\right), \tag{1}$$

where μ is the mean parameter and σ is the volatility parameter. The constant parameter model in Eq. (1) indicates that stock returns always follow one certain distribution, which could provide reasonable approximations for stock returns in short term but fail to capture the evolvement of stock returns in long term.

In a Markov switching model, the long term evolvement of stock returns is assumed to switch between m states. In different states, the distribution of stock returns has different parameters. For each stock return y_t $(t = 1, \cdots, T)$, a latent state variable s_t is introduced to indicate the state to which stock return y_t $(t = 1, \cdots, T)$ belongs. If $s_t = i$ $(i = 1, \cdots, m)$, then y_t is in state i and

$$y_t | s_t = i \sim \mathrm{N}\left(\mu_i, \sigma_i^2\right) \tag{2}$$

where μ_i is the mean parameter and σ_i is the volatility parameter of state i. The switching mechanism between states is governed by a stationary discrete Markov process with

$$\Pr\left(s_{t+1} = j | s_t = i\right) = p_{ij} \ \ (i, j = 1, \cdots, m), \tag{3}$$

where p_{ij} denotes the transition probability from state i at time t to state j at time $t + 1$.

The commonly used conjugate prior distributions are specified for state-specific parameters. For $i = 1, \cdots, m$, the prior distributions for μ_i and σ_i^2 are

$$\mu_i \sim \mathrm{N}\left(\underline{\mu}, \underline{\sigma}^2\right), \tag{4}$$

$$\sigma_i^2 \sim \mathrm{IG}\left(\underline{\alpha}, \underline{\beta}\right), \tag{5}$$

and the prior distribution for the i-th row of the transition probability matrix is

$$\boldsymbol{p}_i = \left(p_{i1}, \cdots, p_{im}\right)' \sim \mathrm{Dirichlet}\left(\underline{a}_{i1}, \cdots, \underline{a}_{im}\right), \tag{6}$$

where $\underline{\mu}, \underline{\sigma}^2, \underline{\alpha}, \underline{\beta}$ and $\underline{a}_{ij}$ $(i, j = 1, \cdots, m)$ are hyperparameters. The prior distributions given in Eqs. (4)–(6) are set to be symmetric and exchangeable across states. There are two reasons for doing this. First, by doing this, we do not impose any specified prior information on parameters and the identification of states would be completely determined by the data according to the posterior information of parameters. Second, the use of asymmetric priors may cause the reduction in the marginal likelihood.

The conditional posterior distributions of μ_i $(i = 1, \cdots, m)$ is

$$\mu_i \left| \left(\sigma_i^2, \boldsymbol{s}, \boldsymbol{y}\right) \right. \sim \mathrm{N}\left(\frac{\sum_{t=1}^T I_{\{s_t=i\}} y_t/\sigma_i^2 + \underline{\mu}/\underline{\sigma}^2}{T_i/\sigma_i^2 + 1/\underline{\sigma}^2}, \left(T_i/\sigma_i^2 + 1/\underline{\sigma}^2\right)^{-1}\right), \tag{7}$$

where $T_i = \sum_{t=1}^T I_{\{s_t=i\}}$ is the number of observations in state i, $\boldsymbol{s} = (s_1, \cdots, s_T)'$ and $\boldsymbol{y} = (y_1, \cdots, y_T)'$. The conditional posterior distribution of σ_i^2 $(i = 1, \cdots, m)$ is

$$\sigma_i^2 | (\mu_i, \boldsymbol{s}, \boldsymbol{y}) \sim \mathrm{IG}\left(\underline{\alpha} + T_i/2, \underline{\beta} + \sum_{t=1}^T I_{\{s_t=j\}}\left(y_t - \mu_j\right)^2/2\right). \tag{8}$$

The conditional posterior distribution of $\boldsymbol{p}_i$ $(i = 1, \cdots, m)$ is

$$\boldsymbol{p}_i | \boldsymbol{s} \sim \mathrm{Dirichlet}\left(T_{i1} + \underline{a}_{i1}, \cdots, T_{im} + \underline{a}_{im}\right), \tag{9}$$

where T_{ij} denotes the number of one-step transitions from state i to state j. The conditional posterior distribution of $\boldsymbol{s}$ is non-standard and its density kernel is

$$f(\boldsymbol{s}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P}, \boldsymbol{y}) \propto \pi_{s_1} h_{s_1}^{1/2} \exp\left[-h_{s_1}\left(y_1 - \mu_{s_1}\right)^2/2\right] \prod_{t=2}^T p_{s_{t-1}s_t} h_{s_t}^{1/2} \exp\left[-h_{s_t}\left(y_t - \mu_{s_t}\right)^2/2\right], \tag{10}$$

where π_{s_1} is the unconditional steady-state probability, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)'$, $\boldsymbol{\sigma} = (\sigma_1^2, \cdots, \sigma_m^2)'$, and $\boldsymbol{P} = [p_{ij}]_{m \times m}$.

The posterior simulator is globally a Gibbs sampler with 4 blocks: $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P}, \boldsymbol{s}$. Draws of $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P}$ can be directly sampled from posterior distributions in Eqs.(7)–(9).[2] An efficient algorithm due to Chib (1996) samples $\boldsymbol{s}$ directly from Eq. (10), and also yields several important functions of interest as byproduct, such as the smoothed probabilities for states.

The evaluation of marginal likelihood is the key to determine the optimal number of states. Chib (1995) provided a method to approximate the marginal likelihood using Gibbs sampler. From the identity $f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P}|\boldsymbol{y})f(\boldsymbol{y}) = f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P})f(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{P})$, we have

$$f(\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{\mu}^*, \boldsymbol{\sigma}^*, \boldsymbol{P}^*) f(\boldsymbol{\mu}^*, \boldsymbol{\sigma}^*, \boldsymbol{P}^*)}{f(\boldsymbol{\mu}^*, \boldsymbol{\sigma}^*, \boldsymbol{P}^*|\boldsymbol{y})} \tag{11}$$

for any fixed $\boldsymbol{\mu}^*, \boldsymbol{\sigma}^*, \boldsymbol{P}^*$, where $\boldsymbol{\mu}^*, \boldsymbol{\sigma}^*, \boldsymbol{P}^*$ is usually set to be the mean or mode of posterior draws of $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$, $\boldsymbol{P}$ in order to improve the

[2] The Gibbs sampler may suffer from the label switching problem, and in this case we label the states by values of the mean parameter.