



Can Markov switching model generate long memory?



Changryong Baek^{a,*}, Natércia Fortuna^{b,1}, Vladas Pipiras^{c,2}

^a Sungkyunkwan University, South Korea

^b CEF.UP - Centre for Economics and Finance at University of Porto, Universidade do Porto, Portugal

^c University of North Carolina, United States

HIGHLIGHTS

- The relationship between Markov switching (MS) model and long memory is reexamined.
- Common spectral estimators of the long memory parameter are found to be extremely biased for the MS model.
- This is explained by analyzing the expected periodogram and the variance of the partial sum process of the MS model.

ARTICLE INFO

Article history:

Received 26 February 2014

Received in revised form

16 April 2014

Accepted 26 April 2014

Available online 5 May 2014

JEL classification:

C22

C13

C14

C50

Keywords:

Markov switching model

Long memory

Changes in mean

ABSTRACT

In an influential work by Diebold and Inoue (2001), the Markov switching model was shown to exhibit long memory, in terms of the behavior of the second moments of partial sums. The relationship between the Markov switching model and long memory is reexamined here. Common estimators of the long memory parameter are found to be extremely biased when applied to the data generated by the Markov switching model. An explanation for these findings is provided.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In a seminal work, Diebold and Inoue (2001) related several forms of changes in regime to long memory. One of several fundamental models studied in their work was the Markov switching (MS) model, defined as

$$X_k^T = \mu_{\delta_k}^T + \epsilon_k, \quad k = 1, \dots, T, \quad (1.1)$$

* Correspondence to: Department of Statistics, Sungkyunkwan University, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul, 110-745, South Korea. Tel.: +82 2 760 0602; fax: +82 502 302 0522.

E-mail addresses: crbaek@skku.edu, crbaek@gmail.com (C. Baek), nfortuna@fep.up.pt (N. Fortuna), pipiras@email.unc.edu (V. Pipiras).

¹ Address: Faculdade de Economia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal.

² Address: Department of Statistics and Operations Research, UNC at Chapel Hill, CB#3260, Hanes Hall, Chapel Hill, NC 27599, USA.

where $\{\epsilon_k\}$ is a sequence of i.i.d. random variables with mean 0 and variance σ_ϵ^2 , $\{S_k^T\}_{k=1,\dots,T}$ is a stationary Markov chain, taking the value 0 or 1, with the transition probabilities $p_{ij} = p_{ij}(T)$, $i, j = 0, 1$, depending on T , and $\mu_0 \neq \mu_1$ are two different mean levels.

Supposing

$$p_{ij} = p_{ij}(T) = 1 - \frac{c_j}{T^{\delta_j}}, \quad j = 0, 1 \quad (1.2)$$

with $0 < c_0, c_1 < 1$ and $0 < \delta_0, \delta_1 < 1$, Diebold and Inoue (2001, Proposition 3), showed that

$$\text{Var}(S_T) \sim CT^{2d+1}, \quad \text{as } T \rightarrow \infty, \quad (1.3)$$

where $S_T = X_1^T + \dots + X_T^T$ is the partial sum of the series X_k^T , and $d = (\min\{\delta_0, \delta_1\} - |\delta_0 - \delta_1|)/2 \in (0, 1/2)$.

The behavior (1.3) of the variances of the partial sums is consistent with that of long memory time series, with the long memory parameter $d \in (0, 1/2)$ (e.g. Robinson, 2003; Beran et al., 2013). The MS model thus seemingly suggests a way to generate (induce)

long memory, providing another physical mechanism for long memory akin to ON/OFF models (Willinger et al., 1997; Mikosch et al., 2002; Smith, 2005) or aggregation of short memory models (Granger, 1980).

In this note, we reexamine the relationship between the MS model and long memory. We find through numerical simulations that the spectral (GPH or local Whittle) estimation of the long memory parameter d in the MS model is extremely biased and very sensitive to the tuning parameter (the number of low frequencies used in estimation). This occurs because the population quantities (the log of the expected periodogram) deviate very significantly from what would be expected under long memory (see Section 2).

We shed further light on the relationship between the MS model and long memory by examining the variance of the partial sum process $S_T(t) = \sum_{k=1}^{\lfloor Tt \rfloor} (X_k^T - EX_k^T)$, $t \in [0, 1]$. As shown below (see Section 3), the behavior of this variance is more in line with short memory (associated with $d = 0$) and not long memory ($d \in (0, 1/2)$ as above). Since this behavior often underlies the estimation of the long (short) memory parameter, the MS model should thus be viewed as a short memory model which happens to have some long memory features as in (1.3). This also sheds light on the estimation results in the spectral domain.

Our findings show that the MS model should not be viewed as another physical mechanism to generate true long memory (see also Section 3 below). The MS model may nevertheless suggest long memory through common spectral estimators such as GPH or local Whittle. Diebold and Inoue (2001) showed this by focusing on testing for short memory (which can be rejected for the MS model). We differ from Diebold and Inoue (2001) by focusing on point estimates of the long memory parameter. This not only confirms the findings of Diebold and Inoue (2001) but also identifies the issue of a serious bias in estimating the expected long memory parameter.

The rest of the note is organized as follows. Estimation results are presented in Section 2 and further discussion can be found in Section 3. Section 4 contains some concluding remarks.

2. Estimation of long memory parameter

In this section, we consider the estimation of the long memory parameter in the MS model. We first consider the popular GPH and local Whittle (LW, in short) methods. Both methods are in the spectral domain, and use the periodogram $I_X(\omega_j)$ of the series $\{X_k^T\}$ at the Fourier frequencies $\omega_j = 2\pi j/T$, $j = 1, \dots, T$. The GPH estimator is obtained from the regression of $\log I_X(\omega_j)$ on $\log \omega_j$ and can be expressed as

$$\hat{d}_{gph} = \sum_{j=1}^m a_j \log I_X(\omega_j),$$

where m is the number of low frequencies used in estimation and the regression weights $a_j = (z_j - \bar{z}) / \sum_{j=1}^m (z_j - \bar{z})^2$ with $z_j = -2 \log \omega_j$, $\bar{z} = \sum_{j=1}^m z_j / m$ (Geweke and Porter-Hudak, 1983). The LW estimator, based on the Whittle approximation of the Gaussian log-likelihood, is defined as (Robinson, 1995)

$$\hat{d}_{lw} = \operatorname{argmin}_{d \in [\Theta_1, \Theta_2]} \left\{ \log \left(\frac{1}{m} \sum_{j=1}^m \omega_j^{2d} I_X(\omega_j) \right) - \frac{2d}{m} \sum_{j=1}^m \log \omega_j \right\}, \quad (2.1)$$

where $-0.5 < \Theta_1 < 0 < \Theta_2 < 1$ are fixed.

Fig. 1 depicts GPH and LW estimation for the MS model. We took $\mu_0 = 0$, $\mu_1 = 1$, $c_0 = c_1 = 0.9$, normal errors with $\sigma_\epsilon^2 = 1$, $T = 10^5$ and two choices of $\delta_0 = \delta_1 = 0.8$ and $\delta_0 = \delta_1 = 0.2$ corresponding to $\bar{d} = 0.4$ and $d = 0.1$, respectively. The top plots are the estimates $\hat{d} = \hat{d}(m)$ as functions of m when $d = 0.4$, for 100 realizations. The bottom plots correspond to $d = 0.1$. It is clearly

seen from Fig. 1 that both estimators are extremely biased for a wide range of low frequencies used in estimation even for such a large sample size. For larger $d = 0.4$, the estimators suggest long memory for a wide range of lower frequencies but the bias is close to zero only for a tiny subset of this range.

The difficulties with estimation can be explained further through the following argument. By using Proposition 3 in Diebold and Inoue (2001), which provides the autocovariance function of the MS model $\{X_k^T\}$, it can be shown that

$$\begin{aligned} \mathbb{E}I_X(\omega_j) &= \frac{1}{2\pi} \left\{ \sigma_\epsilon^2 + \frac{(1-p_{00})(1-p_{11})}{(2-p_{00}-p_{11})^2} (\mu_0 - \mu_1)^2 \right. \\ &\quad \times \left. \left\{ 1 + 2 \sum_{k=1}^{T-1} \left(1 - \frac{|k|}{T} \right) \lambda^k \cos(k\omega_j) \right\} \right\} \\ &= \frac{1}{2\pi} \left\{ \sigma_\epsilon^2 + \frac{(1-p_{00})(1-p_{11})}{(2-p_{00}-p_{11})^2} (\mu_0 - \mu_1)^2 \right. \\ &\quad \times \left\{ 1 + 2\Re \left(\frac{\lambda e^{i\omega_j}}{1 - \lambda e^{i\omega_j}} \right) \right. \right. \\ &\quad \left. \left. + \frac{1}{T} \frac{(\lambda e^{i\omega_j})^{T+1} - \lambda e^{i\omega_j}}{(1 - \lambda e^{i\omega_j})^2} \right\} \right\}, \quad (2.2) \end{aligned}$$

where \Re denotes the real part. Fig. 2 shows the log–log plot of the expected periodogram (2.2) in plus marker, an average of 100 periodograms from simulated data in dotted line, and the slope with the long memory parameter d in dashed–dotted line, against the Fourier frequencies. Note the good agreement between the expected and averaged periodograms. But also note that they deviate very significantly from the line of the expected slope determined by d .

Diebold and Inoue (2001) have also reported some numerical simulations for the MS model (see also Yu, 2009). In figures 8 and 9, they consider the GPH estimates for the choice of $m = T^{0.5}$, finding them in the interval $(0, 0.5)$. Although these findings are consistent with the ones reported here (Fig. 1), we go a step further by focusing on the bias of the GPH estimator.

3. Behavior of variance of partial sum process

A disagreement of the MS model with the expected long memory behavior also manifests itself in the following more basic and fundamental way. If $S_T(t) = \sum_{k=1}^{\lfloor Tt \rfloor} (X_k^T - EX_k^T)$, $t \in [0, 1]$, is the partial sum process of the MS model, then Proposition 3 in Diebold and Inoue (2001) can be used to show that

$$\begin{aligned} \operatorname{Var}(S_T(t)) &= \frac{(1-p_{00})(1-p_{11})}{(2-p_{00}-p_{11})^2} (\mu_0 - \mu_1)^2 [Tt] \\ &\quad \times \left\{ \frac{2\lambda}{1-\lambda} + 1 + \frac{2(\lambda^{\lfloor Tt \rfloor + 1} - \lambda)}{[Tt](1-\lambda)^2} \right\} + \sigma_\epsilon^2 [Tt]. \quad (3.1) \end{aligned}$$

This further leads to

$$v_T(t) := \operatorname{Var} \left(\frac{S_T(t)}{T^{d+1/2}} \right) \sim Ct, \quad (3.2)$$

as $T \rightarrow \infty$ (see Baek et al., 2014 for details).

In the case of a true long memory time series, the right-hand side of (3.2) is Ct^{2d+1} . The behavior (3.2), on the other hand, is associated with short memory and $d = 0$. (As shown in Baek et al., 2014, $S_T(t)/T^{d+1/2}$ also converges in the sense of finite-dimensional distributions to Brownian motion.) Fig. 3 presents the log–log plot of $v_T(t)$ as a function of t , together with the expected slope equal to 1 ($d = 0$). Note the good agreement between $\log v_T(t)$ and the expected slope.

In fact, the relation (3.2) with t^{2d+1} on its right-hand side gives rise to the so-called aggregated variance estimator of the long

Download English Version:

<https://daneshyari.com/en/article/5059452>

Download Persian Version:

<https://daneshyari.com/article/5059452>

[Daneshyari.com](https://daneshyari.com)