Big Data Analysis of the Dynamic Relationship between Stock Prices and Business Cycles via Bayesian Methods

Koki Kyo

Abstract—**In previous study, we proposed a Bayesian modeling technique to decompose a daily time series of Nikkei Stock Average (NSA) into three components which include a trend component, and analyzed the behavior of each estimated component. It was confirmed that there is time-varying correlation between the trend component and the coincident Composite Index in Japan (CIJ). In this paper, as an extension of the previous study we analyze the dynamic relationship between the trend component in the NSA and the CIJ using a regression model with a time-varying coefficient and a lag parameter. The regression model is constructed using the NSA as the dependent variable and the CIJ as the explanatory variable. Bayesian smoothness prior technique is applied to estimate the time-varying coefficient. Moreover, we explain the dynamic relationship between business cycles and stock prices based on the estimates of the time-varying coefficient and the lag parameter. As an empirical example, we analyze the daily time series of NSA closing values from January 4, 1991, to March 30, 2018, together with the monthly CIJ data over the same period.**

Index Terms—**Bayesian modeling, state space model, big data analysis, daily stock price data, business cycles in Japan.**

I. Introduction

Many economic and social events, especially, business cycles can drive stock prices up and down. Conversely, we can think of stock prices as reflecting economic and social changes. Thus, we can better understand the impact of business cycles, economic and social events by analyzing movements in stock prices. From this perspective, stock prices data contain a lot of information about economic and social changes, and the ongoing streams of stock price data make good candidates for big data analysis. Thus, it is valuable to develop an approach to analyze the impact of economic and social changes on stock prices.

Chen examined whether changes in the business cycle could be predicted using a production asset pricing model [1]. Wang analyzed the impact of sudden events on China's automobile industry stock price using the incident research method [2]. In [3] the impact of public announcements on stock prices is investigated based on the event study method. Abhyankar et al. examined the relationship between oil price shocks and the Japanese stock market, found a positive correlation between oil price shocks arising from changes in aggregate global demand and stock market returns [4]. However, a common shortcoming of the above studies is the lack of analysis based on statistical modeling.

Recently, approaches based on time series modeling through the autoregressive integrated moving average in [5] have been used to forecast stock prices [6], [7]. Against these approaches, more advanced analysis of stock prices has used state space modeling methods (for example, [8]). In our previous study [9], we analyzed Nikkei Stock Average (NSA) data using a novel technique in which the state space model is constructed by combining the seasonally adjusted model [10] with the time-varying coefficient autoregressive model [11]. The objective data analyzed in [9] are daily time series for the closing price from January 4, 2000, to November 28, 2017. First, we decomposed the time series into a long-term tendency (called the trend component), short-term variation (called the cyclical component), and error term (called the irregular component) using a Bayesian linear modeling method. The behavior of each component was then examined in relation to economic and social events together with variations in the economic and social environment.

In our previous study [9], the trend component in the NSA time series was found to contain information about the business cycles. In [12], we analyzed the dynamic relationship between the NSA and the coincident Composite Index in Japan (CIJ), which indicates the business cycles, by introducing a time-varying coefficient regression method proposed by [13]. In this paper, as an extension of the previous studies [9] and [12], we analyze the dynamic relationship between the trend component in the NSA and the CIJ using a regression model, which is constructed using the trend in NSA as the dependent variable and the CIJ as the explanatory variable. This regression model is similar to that in the previous study [12] in terms of having a time-varying coefficient, but we also introduce a lag parameter to express the lead-lag relation between the NSA and the ICJ. Moreover, to obtain robust estimate of the time-varying coefficient, we apply the Bayesian smoothness prior technique introduced in [11]. The objective data are daily time series for the NSA closing price in the period from January 4, 1991, to March 30, 2017 together with monthly time series data for CIJ.

The rest of this paper is organized as follows. In Section II, we give a review of our previous studies in [9] and [12]. In Section III, we present the models and methods for parameter estimation. In Section IV, we analyze the estimated results. Finally, we offer some conclusions in Section V.

II. Review of Our Previous Studies

In this section, we give a review of our previous studies [9] and [12].

First, in [9] for the daily time series (in logarithmic scale)

Manuscript received May 3, 2018; revised August 2, 2018. This work is supported in part by a Grant-in-Aid for Scientific Research (C) (16K03591) from the Japan Society for the Promotion of Science.

Koki Kyo is with Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro, Hokkaido 080-8555, Japan (e-mail: kyo@obihiro.ac.jp).

doi: 10.18178/ijtef.2018.9.6.620

y_n of the NSA, we constructed a set of statistical models as follows:

$$y_n = t_n + r_n + w_n, \tag{1}$$

$$t_n = 2t_{n-1} - t_{n-2} + v_{n1}, \tag{2}$$

$$r_n = \sum_{i=1}^{q} \alpha_i(n) r_{n-i} + v_{n2}, \tag{3}$$

where t_n and r_n are the trend and the cyclical components in the time series y_n, respectively, w_n ~ $N(0,\sigma^2)$ is the observation noise or irregular component, and v_{n1} ~ $N(0,\tau_1^2)$ together with v_{n2} ~ $N(0,\tau_2^2)$ are system noises. It is assumed that w_n, v_{n1} and v_{n2} are independent of one another. Also, q represents the order of model for r_n, $\alpha_1(n), \ldots, \alpha_q(n)$ are coefficients that vary over time.

When the values for the time-varying coefficients, $\alpha_1(n), \ldots, \alpha_q(n)$, are given, the models in (1) - (3) form a set of Bayesian linear models in which the model in (1) defines the likelihood and the models in (2) and (3) define the priors for t_n and r_n, respectively. More specifically, the model in (2) forms a second order smoothness prior for the trend component t_n, and the model in (3) expresses an autoregressive (AR) model for the cyclical component r_n. We call this set of Bayesian linear models *model A*.

Features of the model A is as follows: (a) The time series y_n is structured into several components, so we can obtain various signals from the movements of each component. (b) The model in (2) expresses smoothness of the trend component over time. (c) The coefficients $\alpha_1(n), \ldots, \alpha_q(n)$ in the cyclical model vary over time so they are called time-varying AR coefficients, thus we can analyze the dynamics of short-term fluctuations in the time series of the NSA.

The model A can be also considered a special form of the seasonal adjustment models, in which a seasonal component is omitted and a cyclical component is added (see [14] and [15] for details of seasonal adjustment models). In this model, the trend component t_n captures the long-term tendency, the cyclical component r_n expresses short-term variation of the time series y_n, and the irregular component w_n denotes an error term in expressing the variation of y_n by t_n and r_n. Thus, we can search for signals of social or economic changes from the behavior of each component.

There remains another problem, i.e., there are many unknown values for the time-varying AR coefficients, $\alpha_1(n), \ldots, \alpha_q(n)$, to be estimated. For the estimation of the time-varying coefficients, we introduced a set of first order smoothness priors as follows:

$$\alpha_i(n) = \alpha_i(n-1) + \eta_{ni}, \quad (i = 1, \ldots, q) \tag{4}$$

where η_{ni} ~ $N(0,\delta^2)$ represents system noise. We assume that η_{ni} and η_{mj} are independent of each other for $m \neq n$ or $j \neq i$.

If the values for time series of the cyclical component r_n are given, then we can consider the cyclical model in (3) and the model in (4) as a set of Bayesian linear models in which the model in (3) defines the likelihood and the model in (4) plays a role of priors for the time-varying AR coefficients. We call a set of Bayesian linear models in (3) and (4) model B. Thus, the models in (1)-(4) form a set of hierarchical Bayesian linear models.

In [9], the summary for model estimation was given as follows. The process of estimation starts by setting all values for the time-varying AR coefficients to zero. In the first step, a set of state space representations is constructed for the model A. So the estimates for the constant parameters q, σ^2, τ_1^2 and τ_2^2 was obtained by the maximum likelihood method, and the components t_n and r_n were estimated using a Kalman filter and fixed-interval smoothing algorithms based on the state space representation. In the second step, another state space representation for the model B was constructed using the estimate for the cyclical component, which are obtained from the estimation of the model A. Then, the parameter δ^2 was estimated similarly by maximum likelihood method, and estimates for the AR time-varying coefficients, $\alpha_1(n), \ldots, \alpha_q(n)$, are obtained using a Kalman filter and fixed-interval smoothing algorithms based on the state space representation for the model B. So, the process of estimation which repeated between the first and the second steps proceeds until the results converge.

Moreover, in [12] to investigate the dynamic relationship between the NSA and the CIJ, we considered a regression model as follows:

$$\tilde{y}_n = a_n x_n + e_n, \tag{5}$$

where $\tilde{y}_n$ and x_n are normalized time series for the NSA and the ICJ respectively, a_n is the time-varying regression coefficient, and e_n is the error term which is assumed to be normally distributed. We applied a second order smoothness prior for the time-varying coefficient a_n as

$$a_n = 2a_{n-1} - a_{n-2} + \zeta_n \tag{6}$$

with ζ_n being the system noise which is assumed to be normally distributed and independent of e_n.

The models in (5) and (6) form a set of Bayesian linear model for the time-varying coefficient a_n, so it can be expressed in a state space representation. Thus, the Kalman filter algorithm can be used for parameter estimation (See [12] for the details).

III. Models and Parameter Estimation

A. Models

Following the methodology introduced in [12], we analyze the dynamic relationship between the trend component in the NSA and the CIJ based on the results that are obtained by use of the method in [9]. For necessity of modeling, we arrange monthly time series for the CIJ into a form of daily time series. Let $\{\tilde{x}_m; m = 1, 2, \ldots, M\}$ be a set of monthly time series data for the CIJ with M being the length. To make a one-to-one correspondence with $\{\hat{t}_n; n = 1, 2, \ldots, N\}$, which denotes a set of estimates for daily time series of the trend component in the CIJ, with N being the length of the time series y_n. We expand the set $\{\tilde{x}_m; m = 1, 2, \ldots, M\}$ into a set of daily time series data $\{x_n^*; n = 1, 2, \ldots, N\}$ by setting $x_n^* = \tilde{x}_m$ if and only if the n-th day is in the m-th month for $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$.

Let $\bar{t}$ be the average of the elements in the set $\{\hat{t}_n; n = 1, 2, \ldots, N\}$, $\bar{x}$ and s be the average and the standard deviation of the elements in the set $\{x_n^*; n = 1, 2, \ldots, N\}$ respectively. Then, we transform the data $\{x_n^*; n = 1, 2, \ldots, N\}$ into $\{x_n; n = 1, 2, \ldots, N\}$ by

$$x_n = \frac{x_n^* - \bar{x}}{s},$$

and transform the data $\{\hat{t}; n = 1, 2, \ldots, N\}$ into $\{y_n^*; n = 1, 2, \ldots, N\}$ by

$$y_n^* = \hat{t}_n - \bar{t}.$$

To analyze the dynamic relationship between the trend of the NSA and the ICJ, we introduce a dynamic regression model as follows:

$$y_n^* = \beta_n x_{n+L} + \varepsilon_n, \tag{7}$$

where β_n is the time-varying coefficient that comprises a daily time series, and L denotes a lag parameter, $\varepsilon_n \sim \mathrm{N}(0, \lambda^2)$ is the observation noise with λ^2 being the unknown variance.

The lag L and the time-varying coefficient β_n are two important parameters. From the value of L we can see the lead-lag relationship between the NSA and the CIJ in which the case where $L > 0$ implies that the CIJ lags the NSA and the case where $L < 0$ implies that the CIJ precedes NSA. Moreover, from the estimate of the time-varying coefficient β_n we can examine the dynamic relationship between the NSA and the CIJ.

Moreover, similarly to the treatment for the time-varying coefficient a_n, we introduce a second order smoothness prior for the time-varying dependence coefficient β_n as follows:

$$\beta_n = 2\beta_{n-1} - \beta_{n-2} + \psi_n. \tag{8}$$

Here, $\psi_n \sim \mathrm{N}(0, \phi^2)$ represents system noise, where ϕ^2 denotes the unknown variance. It is assumed that ψ_n is independent of ε_n.

B. Estimating Time-Varying Coefficient

If we set

$$\mathbf{z}_n = \begin{bmatrix} \beta_n \\ \beta_{n-1} \end{bmatrix}, \ \mathbf{F} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \ \mathbf{G} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$\mathbf{H}_n = \begin{bmatrix} x_{n+L} & 0 \end{bmatrix}, \ \mathbf{Q} = \mathrm{E}\{\psi_n^2\},$$

then, the models in (7) and (8) can be expressed by a state space representation as

$$\boldsymbol{z}_n = \boldsymbol{F}\boldsymbol{z}_{n-1} + \boldsymbol{G}\psi_n, \tag{9}$$

$$y_n^* = \boldsymbol{H}_n\boldsymbol{z}_n + \varepsilon_n. \tag{10}$$

In the state space representation comprising (9) and (10), the time-varying coefficient β_n is included in the state vector $\boldsymbol{z}_n$. So, the estimate for β_n can be obtained from the estimate of $\boldsymbol{z}_n$. Moreover, the parameters, λ^2 and ϕ^2, which are called hyperparameters, can be estimated using the maximum likelihood method.

Let $\boldsymbol{z}_0$ denote the initial value of the state and let $Y_1^{(k)}$ denote a set of estimates for y_n^* up to the k-th day. Assume that $\boldsymbol{z}_0 \sim \mathrm{N}(\boldsymbol{z}_{0|0}, \boldsymbol{C}_{0|0})$. Because the distribution density $f(\boldsymbol{z}_n | Y_1^{(k)})$ for the state $\boldsymbol{z}_n$ conditional on $Y_1^{(k)}$ is Gaussian, it is only necessary to obtain the mean $\boldsymbol{z}_{n|k}$ and the covariance matrix $\boldsymbol{C}_{n|k}$ of $\boldsymbol{z}_n$ with respect to $f(\boldsymbol{z}_n | Y_1^{(k)})$.

Given the values of L, λ^2 and ϕ^2, the initial distribution $\mathrm{N}(\boldsymbol{z}_{0|0}, \boldsymbol{C}_{0|0})$, and a set of estimates for y_n^* up to time point N, the means and covariance matrices in the predictive distribution and filter distribution for the state $\boldsymbol{z}_n$ can be obtained using the Kalman filter for $n = 1,2,\cdots,N$ (see for example, [15], [16]):

[**Prediction**]

$$\boldsymbol{z}_{n|n-1} = \boldsymbol{F}\boldsymbol{z}_{n-1|n-1},$$

$$\boldsymbol{C}_{n|n-1} = \boldsymbol{F}\boldsymbol{C}_{n-1|n-1}\boldsymbol{F}^{\mathrm{t}} + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^{\mathrm{t}}.$$

[**Filter**]

$$\boldsymbol{K}_n = \boldsymbol{C}_{n-1|n-1}\boldsymbol{H}_n^{\mathrm{t}}\left(\boldsymbol{H}_n\boldsymbol{C}_{n-1|n-1}\boldsymbol{H}_n^{\mathrm{t}} + \lambda^2\right)^{-1},$$

$$\boldsymbol{z}_{n|n} = \boldsymbol{z}_{n|n-1} + \boldsymbol{K}_n\left(y_n^* - \boldsymbol{H}_n\boldsymbol{z}_{n|n-1}\right),$$

$$\boldsymbol{C}_{n|n} = (\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H}_n)\boldsymbol{C}_{n|n-1}.$$

Based on the results of the Kalman filter, we can obtain the final estimate for $\boldsymbol{z}_n$ using the fixed-interval smoothing for $n = N-1,\ N-2,\ \cdots, 1$ as follows:

[**Fixed-Interval Smoothing**]

$$\boldsymbol{A}_n = \boldsymbol{C}_{n|n}\boldsymbol{F}^{\mathrm{t}}\boldsymbol{C}_{n+1|n}^{-1},$$

$$\boldsymbol{z}_{n|N} = \boldsymbol{z}_{n|n} + \boldsymbol{A}_n\left(\boldsymbol{z}_{n+1|N} - \boldsymbol{z}_{n+1|n}\right),$$

$$\boldsymbol{C}_{n|N} = \boldsymbol{C}_{n|n} + \boldsymbol{A}_n\left(\boldsymbol{C}_{n+1|N} - \boldsymbol{C}_{n+1|n}\right)\boldsymbol{A}_n^{\mathrm{t}}.$$

Then, the posterior distribution of $\boldsymbol{z}_n$ is given by $\boldsymbol{z}_{n|N}$ and $\boldsymbol{C}_{n|N}$. Subsequently, the estimate $\hat{\beta}_n$ for the time-varying coefficient β_n can be obtained because the state space representation described by (9) and (10) incorporates β_n in the state vector $\boldsymbol{z}_n$.

Incidentally, from (7) the predicted value $\hat{y}_n^*$ of y_n^* can be calculate as follows:

$$\hat{y}_n^* = \hat{\beta}_n x_{n+L} \tag{11}$$

C. Estimating Hyperparameters

Given the time series data $Y_1^{(N)} = \{y_1^*,\ y_2^*,\ \cdots,\ y_N^*\}$ and the corresponding time series data $\{x_1,\ x_2,\ \cdots,\ x_N\}$, a likelihood function for the hyperparameters λ^2 and ϕ^2 and the parameter L is given by:

$$f\left(Y_1^{(N)}|\lambda^2,\phi^2,L\right) = \prod_{n=1}^{N} f_n(y_n^*|\lambda^2,\phi^2,L),$$

where $f_n(y_n^*|\lambda^2,\phi^2,L)$ is the density function of y_n^*. By taking the logarithm of $f\left(Y_1^{(N)}|\lambda^2,\phi^2,L\right)$, the log-likelihood is defined as

$$l(\lambda^2,\phi^2,L) = \log f\left(Y_1^{(N)}|\lambda^2,\phi^2,L\right) = \sum_{n=1}^{N}\log f_n(y_n^*|\lambda^2,\phi^2,L). \tag{12}$$

Following [15], using the Kalman filter, the density function $f_n(y_n^*|\lambda^2,\phi^2,L)$ is a normal density given by:

$$f_n(y_n^*|\lambda^2,\phi^2,L) = \frac{1}{\sqrt{2\pi w_{n|n-1}}} \times \exp\left\{-\frac{\left(y_n^* - y_{n|n-1}\right)^2}{2w_{n|n-1}}\right\}, \tag{13}$$

where $y_{n|n-1}$ is the one-step-ahead prediction for y_n^* and $w_{n|n-1}$ is the variance of the predictive error, respectively given by

$$y_{n|n-1} = \boldsymbol{H}_n\boldsymbol{z}_{n|n-1},$$

$$w_{n|n-1} = \boldsymbol{H}_n\boldsymbol{C}_{n|n-1}\boldsymbol{H}_n^{\mathrm{t}} + \lambda^2.$$

Moreover, for a fixed value of L, the estimates of the hyperparameters can be obtained using the maximum likelihood method, i.e., we can estimate the hyperparameters by maximizing the log-likelihood $l(\lambda^2,\phi^2,L)$ in (12) together with (13). In practice, when we substitute the new $\lambda^2 = 1$ into the Kalman filter algorithm outlined above, the estimate $\hat{\lambda}^2$ for λ^2 is obtained analytically by

$$\hat{\lambda}^2 = \frac{1}{N}\sum_{n=1}^{N}\frac{\left(y_n^* - y_{n|n-1}\right)^2}{w_{n|n-1}}. \tag{14}$$

Thus, an estimate $\hat{\phi}^2$ for ϕ^2 can be obtained by maximizing the log-likelihood $l(\hat{\lambda}^2,\phi^2,L)$ using (12) together with (14).

Information about the value of lag L is important for analyzing the lead-lag relationship between the NSA and the CIJ, and can be obtained from the maximum value of the marginal log-likelihood $l(\hat{\lambda}^2, \hat{\phi}^2, L)$.

IV. Results and Analysis

Here, we use data on the closing values of the NSA as the object for the analysis. We analyze daily time series data covering the period from January 4, 1991 to March 30, 2018, which were obtained from the website of Yahoo Japan (http://info.finance.yahoo.co.jp/). Thus, we have $N = 9948$. Every day is included in the period of analysis; the values of the NSA on weekends and holidays are treated as missing values. We also use the data for the CIJ time series covering the same period (the CIJ data were supplied by the Cabinet Office, Government of Japan: (http://www.esri.cao.go.jp/jp/stat/di/di.html).

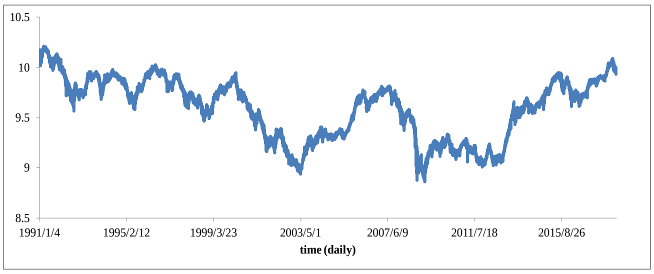

Fig. 1. Original time series data for the NSA.

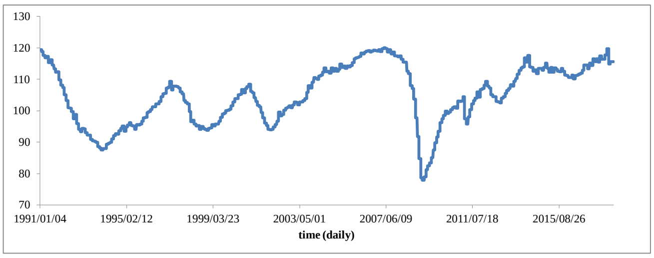

Fig. 2. Expanded original time series data for the CIJ.

Fig. 1 shows the original time series data for the NSA, and Fig. 2 shows the expanded original time series data for the CIJ.

From Fig. 1 and Fig. 2, we can see that the overall behavior of the NSA time series is very similar to that of the CIJ time series, but the similarity between the movements of these two time series differs in each short period. Broadly speaking, there may be a high degree of correlation between the two time series, but the pattern of the relationship changes over time. Thus, the relationship between the NSA and CIJ time series varies with time, and it is very interesting to analyze the dynamic relationship between the NSA and the CIJ.

Fig. 3 shows the time series of the estimates for the trend component in the NSA. From this figure, we see that the estimate for the trend component is smoother than the original time series but shows periodic movement with irregular cycles. We expected that the periodic movements might express natural business cycles, which prompted us to analyze the relationship between the estimate of the trend component and the CIJ, a Japanese business cycle index.

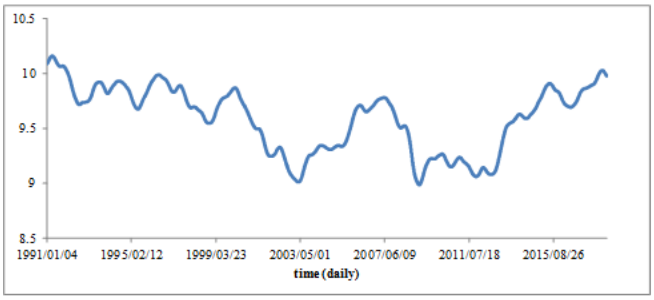

Fig. 3. Time series of the estimate for the trend component in the NSA.

The order of the AR model in (3) is determined as $q = 2$ based on the method of maximum likelihood. Fig. 4 shows the estimates for the time-varying AR coefficients (a) $\alpha_1(n)$ and (b) $\alpha_2(n)$ in the cyclical model, in which the behavior of the coefficients expresses structural change in the self-correlation of the cyclical component.

From Fig. 4, we can see, for example, an extreme fluctuation around October, 2008, which occurred shortly after the bankruptcy declaration of Lehman Brothers, and so on.

We give below the estimation results for the time-varying coefficient regression models in (7) and (8) obtained based on the estimates for the trend component in the NSA and data of the CIJ. We use a part of the estimates for the trend component in the period from October 1, 1992 to March 30, 2017 in order to ensure necessary values for the lag L. Thus, we have $N = 9312$ for the analysis below.

First, to determine the value of the lag L, we practice related calculation for each value of L on the interval from -600 to -200. Fig. 5 shows the distribution of marginal log-likelihood on this interval of the lag. The maximum value of the marginal log-likelihood is 7732.48, which is attained at $L = -457$. That is, the movement of the CIJ precedes that of NSA, it implies that the stock prices are as the effects of business cycles.

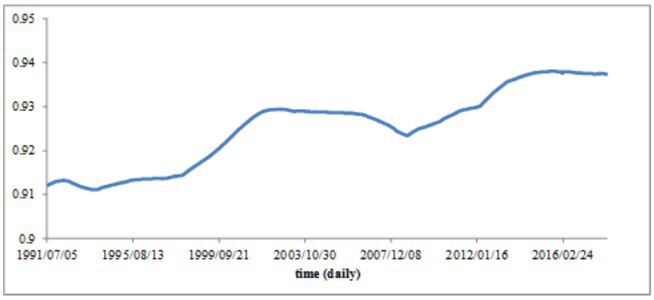

Fig. 4. (a) Time series of the estimate for alpha_1(n).

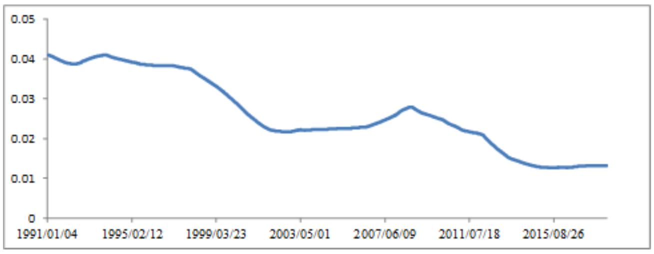

Fig. 4. (b) Time series of the estimate for alpha_2(n).

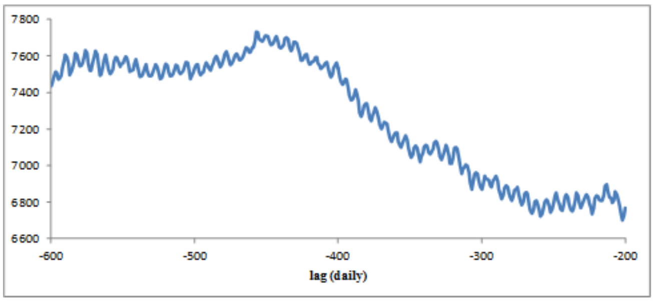

Fig. 5. Distribution of marginal log-likelihood on the lag.

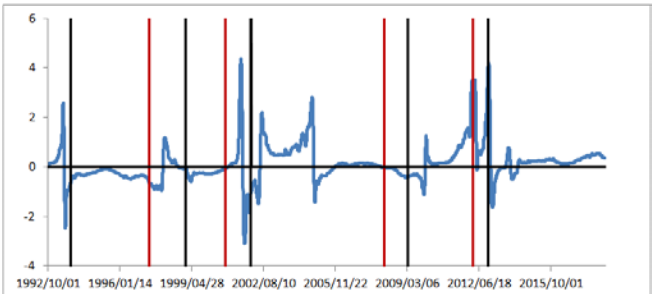

Fig. 6. Time series of estimated mean for the time-varying coefficient

Fig. 6 shows the estimated mean of the time-varying coefficient β_n. In this figure, the vertical lines indicate the turning points in the business cycles (the red and black lines indicate peaks and troughs, respectively). From Fig. 6, it can be seen that the mean of the coefficient varies smoothly over time, but contains a rapid change in each phase of the business cycle. Moreover, we can see that the sign of the time-varying coefficient switches at each point of this rapid change. In the recession phase, which corresponds to the period from a peak to the next trough in the business cycle, the time-varying coefficient turns from positive to negative with a rapid change point at the boundary; similarly, in the expansion phase, which corresponds to the period from a trough to the next peak in the business cycle, the time-varying coefficient turns from negative to positive.

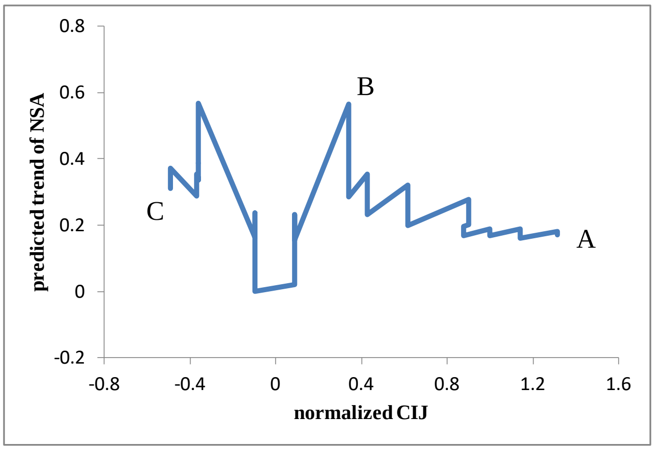

Fig. 7. (a) Scatter diagram from Oct. 1, 1992 to Oct. 14, 1993.

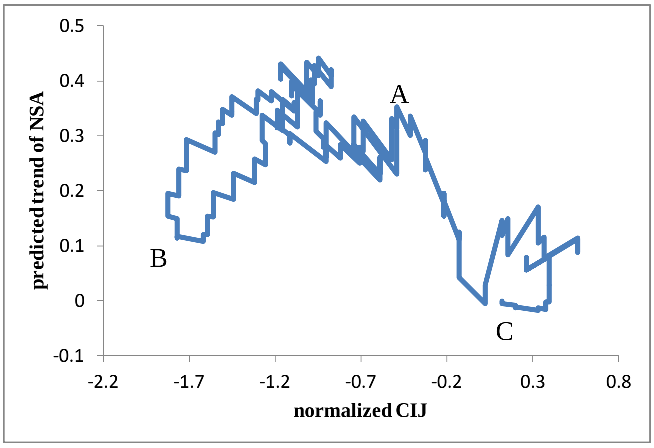

Fig. 7. (b) Scatter diagram from Oct. 15, 1993 to Jan. 14, 1999.

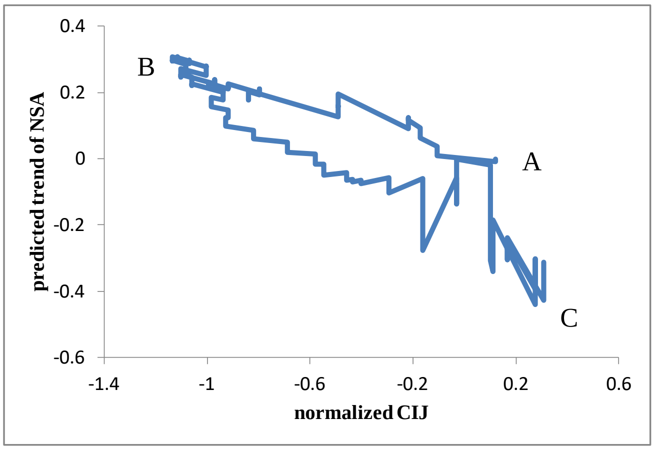

Fig. 7. (c) Scatter diagram from Jan. 15, 1999 to Jan. 14, 2002.

To understand the dynamic relationship between the NSA and the CIJ more readily, we indicate the dynamic relationship between the normalized CIJ and the predicted trend of NSA, which is obtained using (11), by line graph of scatter diagram. Fig. 7 shows the scatter diagrams for each business cycle in Japan: (a) latter period of the eleventh cycle (from Oct. 1, 1992 to Oct. 14, 1993), (b) the twelfth cycle (from Oct. 15, 1993 to Jan. 14, 1999), (c) the thirteenth cycle (from Jan. 15, 1999 to Jan. 14, 2002), (d) the fourteenth cycle (from Jan. 15, 2002 to Mar. 14, 2009), (e) the fifteenth cycle (from Mar. 15, 2009 to Nov. 14, 2012), (f) previous period of the sixteenth cycle (from Nov. 15, 2012 to Mac. 30, 2017). Note that in each cycle, route of the scatter diagram starts from the point A, and reaches point C via point B.

From Fig. 7, it can be seen that in almost panels each route of the scatter diagram forms an irregular ellipse except the panels (a) and (f) which show only a part of business cycle respectively.

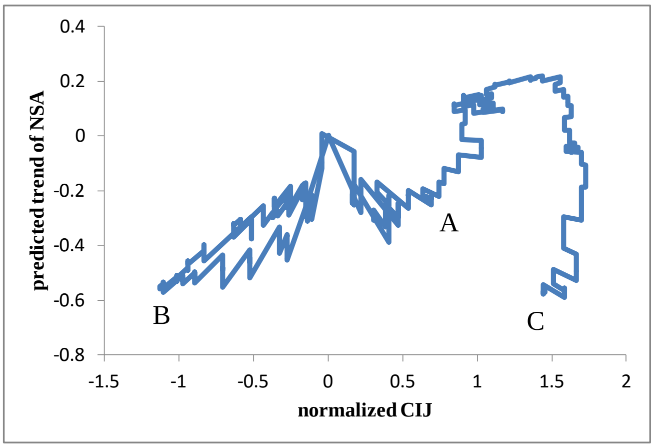

Fig. 7. (d) Scatter diagram from Jan. 15, 2002 to Mar. 14, 2009.

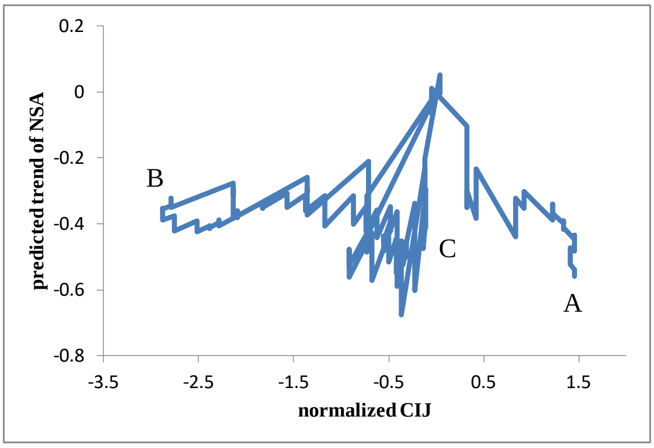

Fig. 7. (e) Scatter diagram from Mar. 15, 2009 to Nov. 14, 2012.

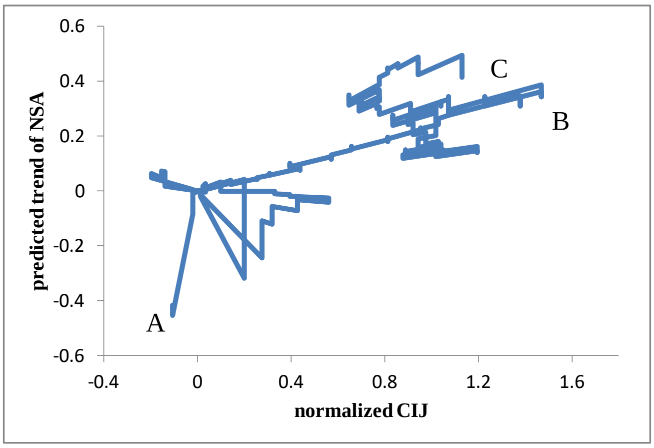

Fig. 7. (f) Scatter diagram from Nov. 15, 2012 to Mar. 30, 2017.

V. Conclusion

In this paper, we expanded our previous researches for analyzing the dynamic relationship between stock prices and business cycles using Bayesian modeling techniques. First, we described a process for decomposing the daily time series of Nikkei Stock Average (NSA) into a trend component, a cyclical component, and an irregular component. Then, we analyzed the dynamic relationship between the estimated trend component in NSA and the coincident Composite Index in Japan (CIJ) by constructing a regression model with time-varying coefficient. In the regression model, we employed the NSA as the dependent variable, and used the lagged CIJ as the explanatory variable. Bayesian smoothness prior technique is applied to estimate the time-varying coefficient. As an empirical study, we analyzed the daily time series for the closing price of the NSA from January 4, 1991 to March 30, 2017, it yielded the following results:

1) The CIJ precedes NSA in 457 days (almost one and a half years), it implies that the stock prices are as the effects of business cycles.
2) The mean of the coefficient generally varies smoothly over time, but exhibits rapid changes in each phase of the business cycle. The sign of the time-varying coefficient switches when such a rapid change occurs.
3) In the recession phase, the time-varying coefficient turns from positive to negative, and in the expansion phase, the time-varying coefficient turns from negative to positive, with a point of rapid change at the boundary.
4) In almost business cycles routes of the scatter diagram between CIJ and the predicted trend of NSA form irregular ellipses.

REFERENCES

[1] M. H. Chen, "Stock returns and changes in the business cycle," *Asia Pacific Management Review*, vol. 10, pp. 321-327, September 2005.

[2] X. Wang, "An analysis of the impact of sudden events on China's automobile industry stock price," *Asian Journal of Social Science Studies*, vol. 2, pp. 140-147, March 2017.

[3] J. Stankeviciene and S. Akelaitis, "Impact of public announcements on stock prices: relation between values of stock prices and the price changes in Lithuanian stock market," *Procedia: Social and Behavior Sciences*, vol. 156, pp. 538-542, April 2014.

[4] A. Abhyankar, B. Xu, and J. Wang, "Oil prices shock and the stock market: evidence from Japan," *Energy Journal*, vol. 34, pp. 199-222, April 2013.

[5] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day, 1970.

[6] J. Zhang, R. Shan, and W. Su, "Applying time series analysis builds stock price forecast model," *Modern Applied Science*, vol. 3, pp. 152-157, March 2009.

[7] B. U. Devi, D. Sundar, and Dr. P. Alli, "An effective time series analysis for stock trend prediction using ARIMA model for Nifty Midcap-50," *International Journal of Data Mining & Knowledge Management Process*, vol.3, pp. 65-78, January 2013.

[8] N. S. Balke and M. E. Wohar, "Low-frequency movements in stack prices: a state-space decomposition," *Review of Economics and Statistics*, vol. 84, pp. 649-667, November 2002.

[9] K. Kyo, "A method for big data analysis of the impact of economic and social events on Japanese stock prices," in *Proc. of the 2018 International Conference on Big Data and Education*, New York: ACM, 2018.

[10] G. Kitagawa and W. Gersch, "A smoothness priors-state space modeling of time series with trend and seasonality," *Journal of the American Statistical Association*, vol. 79, pp. 378-389, June 1984.

[11] G. Kitagawa and W. Gersch, "A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series," *IEEE Transactions on Automatic Control*, Vol. AC-30, pp. 48-56, January 1985.

[12] K. Kyo, "Big data analysis of the dynamic effects of business cycles on stock prices in Japan," in *Proc. of the 15th International Symposium on Pervasive Systems, Algorithms and Networks*, IEEE Computer Society, 2018 (in press).

[13] X. Q. Jiang (K. Kyo), "A Bayesian method for the dynamic regression analysis," *Transactions of the Institute of Systems, Control and Information Engineers*, vol. 8, pp. 20-32, January 1995.

[14] G. Kitagawa and W. Gersch, *Smoothness Priors Analysis of Times Series*, New York: Springer-Verlag, 1996.

[15] G. Kitagawa, *Introduction to Time Series Modeling*, New York: CRC Press, 2010.

[16] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, New Jersey: Prentice-Hall, 1979.

Koki Kyo was born in Shenyang, China, on Sept. 27, 1957. He received his BE degree and ME degree in Management Engineering from Northeastern University, China, in 1982 and 1984 respectively, and PhD degree in Statistical Science from the Graduate University for Advanced Studies, Japan, in 1993. He is a Professor of Statistics at Obihiro University of Agriculture and Veterinary Medicine. His current research interests include Bayesian statistical modeling, estimation and identification of dynamic systems and Bayesian methods of economic analysis. Prof. Kyo is a member of the Japan Statistical Society, a member of the Institute of Systems, Control and Information Engineers, and many others.