96-F-11

Stationary and Non-stationary Simultaneous Switching Autoregressive Models with an Application to Financial Time Series

by

Naoto Kunitomo
University of Tokyo
and
Seisho Sato
Institute of Statistical Mathematics

June 1996

Stationary and Non-stationary Simultaneous Switching Autoregressive Models with an Application to Financial Time Series *

By Naoto Kunitomo †and Seisho Sato ‡

July 1995 (First Version)
May 1996 (Revised Version)

Abstract

A common observation among economists on many economic time series including major financial time series is the asymmetrical movement between the downward phase and the upward phase of their sample paths. Since this feature of time irreversibility cannot be described by the Gaussian ARMA, ARIMA, and ARCH time series models, we propose stationary and non-stationary Simultaneous Switching Autoregressive (SSAR) models, which are non-linear switching time series models. We discuss some properties of these time series models and the estimation method for their unknown parameters. The asymmetrical conditional heteroskedasticity can be easily incorporated into the SSAR models. We also report a simple empirical result on Nikkei 225 spot and futures indices by using a non-stationary SSAR model.

Key Words

Asymmetry, Non-linearity, Non-stationarity, Simultaneous Switching Autoregressive Model, Time Irreversibility, Conditional Heteroskedasticity, Financial Time Series

---

*This paper is a revised version of Discussion Paper No.95-F-13, Faculty of Economics, University of Tokyo. We thank Professor Peter Robinson for giving helpful comments on the first version of the paper. We also thank Dr. Kunio Okina and Professor Genshiro Kitagawa for useful discussions on the related topics. The research was conducted while the first author was a visiting research fellow at Institute of Monetary and Economic Studies, Bank of Japan.

†Faculty of Economics, University of Tokyo, Bunkyo-ku, Hongo 7-3-1, Tokyo 113, JAPAN.

‡Institute of Statistical Mathematics, Minato-ku, Minami-Azabu 4-6-7, Tokyo 106, JAPAN.

1. Introduction

In the past decade, several non-linear time series models have been proposed by statisticians and econometricians. Granger and Andersen (1978), for instance, introduced the bilinear time series models. In statistical time series analysis, Ozaki and Oda (1978), and Tong (1983) proposed the exponential autoregressive (EXPAR) model and the threshold autoregressive (TAR) model respectively. In particular, considerable attention has been paid to the TAR model in the past decade by statisticians and econometricians and several related applications have been reported. The statistical details of many non-linear time series models in statistical time series analysis have been discussed by Tong (1990). Several non-linear time series models have also been proposed for the econometric analyses of time series In particular, considerable attention has been focussed on the autoregressive conditional heteroskedasticity (ARCH) model, which was originally proposed by Engle (1982) and has been used in many empirical studies. Extensions of the standard ARCH model and related issues have been discussed by Nelson (1991) and Hamilton (1994).

In this paper we shall propose an alternative class of non-linear time series models, which we shall call the Simultaneous Switching Autoregressive (SSAR) time series model. This model is a kind of Markovian switching time series model with a quite distinctive structure of simultaneity. We propose this class of statistical models because we believe that the class of Gaussian Autoregressive Moving-average (ARMA) time series model and Gaussian Autoregressive Integrated Moving-average (ARIMA) time series model cannot describe one important aspect in many economic time series, that is, the asymmetrical movement in the upward phase (or regime) and in the downward phase (or regime). It has sometimes been argued that major economic time series display some kind of asymmetrical movements over various phases of the business cycle. In particular, a number of economists have observed the asymmetrical pattern in the upward phase and in the downward phase for major financial time series including stock prices. This feature of economic time series can be regarded as one form of the time irreversibility discussed in statistical time series analysis : see Chapter 4 of Tong (1990).

Earlier, we introduced the simple stationary SSAR time series model and discussed its statistical properties in some detail (Kunitomo and Sato (1996)). Let $\{y_t\}$ be a sequence of scalar time series satisfying

$$y_t = \begin{cases} Ay_{t-1} + \sigma_1 v_t & \text{if} \quad y_t \geq y_{t-1} \\ \\ By_{t-1} + \sigma_2 v_t & \text{if} \quad y_t < y_{t-1} \end{cases} , \tag{1.1}$$

where A, B, σ_i ($\sigma_i > 0, i = 1, 2$) are scalar unknown coefficients, and $\{v_t\}$ is a sequence of i.i.d. random variables followed by $N(0, 1)$. If we impose the coherency condition given by

(1.2) $$\frac{1-A}{\sigma_1} = \frac{1-B}{\sigma_2} = r \ ,$$

we have the Markovian representation

(1.3) $$y_t = y_{t-1} + [\sigma_1 1_{\{v_t \geq r y_{t-1}\}} + \sigma_2 1_{\{v_t < r y_{t-1}\}}][-r y_{t-1} + v_t] \ ,$$

where r is an unknown parameter and $1_{\{\cdot\}}$ is the indicator function. When $\sigma_1 = \sigma_2 = \sigma$, then this model becomes the standard $AR(1)$ model if we re-parametrize $A = B = 1 - \sigma r$. As we have shown, even this simplest univariate SSAR model, called SSAR(1), provides some explanations and descriptions of a very important aspect of the asymmetrical movement of time series in two different phases (Kunitomo and Sato (1996)). This characteristic of economic time series has been observed by a number of economists. However, as far as we are aware there has not been any useful time series model incorporating this feature explicitly in the econometric literature. The main point of our studies (Kunitomo and Sato (1994,1996)) was to link the stationary non-linear time series models to the disequilibrium econometric models. We also investigated the conditions for ergodicity and the basic properties of the stationary distribution in the stationary SSAR model.

This paper extends the basic SSAR model (denoted by $SSAR_m(p)$) discussed by us (Kunitomo and Sato (1996)) in two important directions for econometric applications. First, we shall allow the disturbance terms in the SSAR model to be auto-correlated and have a finite order moving-average (MA) structure. By this extension the SSAR model can exhibit more complicated patterns of auto-correlations among economic time series and their differenced data. Second, and more importantly, we shall consider a class of non-stationary SSAR models, which is one type of the $I(1)$ processes and hence useful for application to major financial time series. In the past analyses of financial time series data, the linear non-stationary time series models have often been used because the movements of most financial time series are usually too volatile as the realizations of stationary time series. We shall put forward one convincing economic reason why the non-stationary SSAR model introduced in this paper is interesting and useful in its applications to financial time series. Although it has been a fairly common observation among many economists that many financial time series including stock prices have asymmetrical movements between the upward phase and the downward phase, it is not possible to describe this kind of asymmetrical pattern by the standard linear non-stationary time series models including the ARIMA time series model and the standard ARCH model proposed by Engle (1982). The stationary and non-stationary SSAR models we shall propose have the property of asymmetrical movement of time series in the two phases. Hence they can easily be extended to handle the asymmetrical conditional heteroskedasticities. The non-stationary SSAR model could also be called a simultaneous switching integrated autoregressive (SSIAR) model, because it can be regarded as a simple non-linear extension of the standard ARIMA model.

In Section 2, we shall introduce the general SSAR model, which can be stationary or non-stationary, and discuss several important examples of its possible applications. We shall also investigate the basic properties of a non-stationary univariate SSAR model with time trend. Furthermore, we shall discuss some generalizations of the SSAR model and some implications for modelling asymmetrical conditional heteroskedasticities. Then in Section 3, we shall discuss one justification for the non-stationary SSAR model from the view of financial economics, and apply the non-stationary $SSAR(1)$ model with time trend to the analysis of Nikkei 225 spot and futures indices. In Section 4, some concluding remarks on our econometric approach to the non-stationary and non-linear time series modelling will be given. Proofs of theorems will be gathered in the Appendix.

2. Stationary and Non-stationary SSAR models

2.1 The SSAR model

In this section we shall consider the multivariate simultaneous switching autoregressive (SSAR) model with moving-average (MA) disturbances. In the following representation the order of the autoregressive part is one without loss of generality. This is because we can consider the p-th order multivariate SSAR model similarly, which can also be re-written in a first order multivariate autoregressive form by using the standard Markovian representation well known in statistical time series analysis.

Let $\boldsymbol{y}_t$ be an $m \times 1$ vector of time series variables. The model we consider in this section is represented by

$$(2.1) \qquad \boldsymbol{y}_t = \begin{cases} \boldsymbol{\mu}_1 + \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{D}_1\boldsymbol{u}_t & \text{if} \quad \boldsymbol{e}'_m\boldsymbol{y}_t \geq \boldsymbol{e}'_m\boldsymbol{y}_{t-1} \\ \boldsymbol{\mu}_2 + \boldsymbol{B}\boldsymbol{y}_{t-1} + \boldsymbol{D}_2\boldsymbol{u}_t & \text{if} \quad \boldsymbol{e}'_m\boldsymbol{y}_t < \boldsymbol{e}'_m\boldsymbol{y}_{t-1} \end{cases},$$

where $\boldsymbol{e}'_m = (0, \cdots, 0, 1)$ and $\boldsymbol{\mu}'_i$ $(i = 1, 2)$ are $1 \times m$ vectors of constants, $\boldsymbol{A}$ and $\boldsymbol{B}$ are $m \times m$ matrices, and $\boldsymbol{D}_i$ $(i = 1, 2)$ are $m \times n$ matrices.
The disturbance terms $\{\boldsymbol{u}_t\}$ are a sequence of $I(d)$ process in the sense that

$$(2.2) \qquad \Delta^d \boldsymbol{u}_t = \sum_{j=0}^{\infty} \boldsymbol{C}_j \boldsymbol{v}_{t-j},$$

where $\boldsymbol{C}_0 = \boldsymbol{I}_n$,

$$(2.3) \qquad \sum_{j=0}^{\infty} \|\boldsymbol{C}_j\| < +\infty ,$$

and $\{\boldsymbol{v}_t\}$ are a sequence of martingale differences with $E(\boldsymbol{v}_t|\mathcal{F}_{t-1}) = \boldsymbol{o}$ and

$$(2.4) \qquad E(\boldsymbol{v}_t\boldsymbol{v}'_t|\mathcal{F}_{t-1}) = \boldsymbol{\Omega}_t \ (a.s.) .$$

In the above notations the σ−field $\mathcal{F}_{t-1}$ is generated by $\{\boldsymbol{y}_s, \boldsymbol{u}_s; s \leq t-1\}$, Δ is the difference operator, $I(d)$ denotes the integrated linear stochastic process

$(d \geq 0)$, and $\boldsymbol{\Omega}_t$ represents the conditional covariance matrix. Although the order of the moving-average (MA) terms in (2.3) can be $+\infty$ in the general case, we shall only deal in this paper with the finite MA case when $\boldsymbol{C}_j = \boldsymbol{O}$ $(j > q)$. The distinction between m and n can be useful when we deal with the higher order SSAR models.

The most important feature of this representation is that the time series variables may take quite different values in two different phases or regimes. This type of statistical time series models is called the threshold time series model in the recent time series literature. However, since the vector time series and two phases at time t are determined simultaneously, we shall refer to this type of time series models as simultaneous switching autoregressive (SSAR) time series models. It will appear later in this paper that this simultaneity has not only important economic interpretations, but also casts new light on the non-linear time series modelling.

We now consider the basic question whether the stochastic process defined by (2.1), (2.2), and (2.3) is meaningful in a proper statistical sense. The general answer to this question is negative and we need some additional conditions on the unknown parameters in the SSAR model. This issue has been called the coherency problem. We say the non-linear time series model (2.1) is coherent if and only if the correspondence between $\{\boldsymbol{y}_t\}$ and $\{\boldsymbol{u}_t\}$ is one-to-one given the initial condition $\mathcal{F}_0$. (See Gourieroux et. al. (1980) and Section 4 of Kunitomo and Sato (1996) for the detail.) The conditions of $\boldsymbol{e}'_m \boldsymbol{y}_t \geq \boldsymbol{e}'_m \boldsymbol{y}_{t-1}$ and $\boldsymbol{e}'_m \boldsymbol{y}_t < \boldsymbol{e}'_m \boldsymbol{y}_{t-1}$ can be rewritten as

$$\boldsymbol{e}'_m \boldsymbol{D}_1 \boldsymbol{u}_t \geq \boldsymbol{e}'_m (\boldsymbol{I}_m - \boldsymbol{A}) \boldsymbol{y}_{t-1} - \boldsymbol{e}'_m \boldsymbol{\mu}_1 \quad , \tag{2.5}$$

and

$$\boldsymbol{e}'_m \boldsymbol{D}_2 \boldsymbol{u}_t < \boldsymbol{e}'_m (\boldsymbol{I}_m - \boldsymbol{B}) \boldsymbol{y}_{t-1} - \boldsymbol{e}'_m \boldsymbol{\mu}_2 \quad , \tag{2.6}$$

respectively. When $m = n$, a set of conditions on the coherency for the SSAR model can be summarized by a $1 \times m$ vector $\boldsymbol{d}'$ and a $1 \times (1+m)$ vector $\boldsymbol{r}'$:

$$\frac{1}{\sigma_1} \boldsymbol{e}'_m \boldsymbol{D}_1 = \frac{1}{\sigma_2} \boldsymbol{e}'_m \boldsymbol{D}_2 = \boldsymbol{d}' \ , \tag{2.7}$$

and

$$\begin{aligned} \frac{1}{\sigma_1} [-\boldsymbol{e}'_m \boldsymbol{\mu}_1, \ \boldsymbol{e}'_m (\boldsymbol{I}_m - \boldsymbol{A})] &= \frac{1}{\sigma_2} [-\boldsymbol{e}'_m \boldsymbol{\mu}_2, \ \boldsymbol{e}'_m (\boldsymbol{I}_m - \boldsymbol{B})] \\ &= \boldsymbol{r}' \ , \end{aligned} \tag{2.8}$$

where σ_i $(i = 1, 2)$ are unknown scale parameters and $\boldsymbol{d}'\boldsymbol{d} = 1$ for normalization. We then have the following proposition, a proof for which is given in the Appendix.

Theorem 2.1 : *Suppose (i) $m = n$, (ii) $\sigma_i > 0$ $(i = 1, 2)$, $|\boldsymbol{D}_1 \boldsymbol{D}_2| > 0$, and (iii) the conditions (2.7) and (2.8) hold. Then the correspondence between two stochastic processes $\{\boldsymbol{u}_t\}$ and $\{\boldsymbol{y}_t\}$ defined in $\boldsymbol{R}^m$ is one-to-one given the initial condition $\mathcal{F}_0$.*

This proposition means that the SSAR model consisting of (2.1), (2.2), and (2.3) is coherent as an econometric model under the assumptions in Theorem 2.1. Hence the number of structural parameters in the SSAR model is less than the number of parameters appearing in (2.1), which can be regarded as a reduced form representation. When $m = n = 1$, we do not need (2.7) because it is automatically satisfied. In this case we use the notation $\sigma_i = \boldsymbol{D}_i > 0$ $(i = 1, 2)$ without loss of generality.

We define the indicator functions by

$$I_t^{(1)} = 1_{\{\boldsymbol{e}_m'\boldsymbol{y}_t \geq \boldsymbol{e}_m'\boldsymbol{y}_{t-1}\}} \tag{2.9}$$

and

$$I_t^{(2)} = 1_{\{\boldsymbol{e}_m'\boldsymbol{y}_t < \boldsymbol{e}_m'\boldsymbol{y}_{t-1}\}} , \tag{2.10}$$

where $1_{\{\omega\}} = 1$ if the event ω occurs and $1_{\{\omega\}} = 0$ otherwise. By the use of this notation, it is often more convenient to rewrite (2.1) in the following form:

$$\boldsymbol{y}_t = \boldsymbol{\mu}(t) + \boldsymbol{A}(t)\boldsymbol{y}_{t-1} + \boldsymbol{D}(t)\boldsymbol{u}_t , \tag{2.11}$$

where

$$\boldsymbol{\mu}(t) = \sum_{i=1}^{2} I_t^{(i)} \boldsymbol{\mu}_i , \tag{2.12}$$

$$\boldsymbol{A}(t) = \boldsymbol{A}\, I_t^{(1)} + \boldsymbol{B}\, I_t^{(2)} , \tag{2.13}$$

and

$$\boldsymbol{D}(t) = \sum_{i=1}^{2} I_t^{(i)} \boldsymbol{D}_i . \tag{2.14}$$

There are several special cases of (2.11), which are interesting from the viewpoint of possible econometric applications. Here we shall make mention of only three examples in the class of the SSAR models we introduced.

Example 1 : Consider the SSAR model when $d = q = 0$. This is the case which we have investigated in some detail (Kunitomo and Sato (1996)). We assumed that the disturbance terms $\{\boldsymbol{u}_t\}$ are a sequence of martingale differences with conditional homoskedasticity, i.e.

$$E(\boldsymbol{u}_t|\mathcal{F}_{t-1}) = \mathbf{o} , \tag{2.15}$$

and

$$E(\boldsymbol{u}_t\boldsymbol{u}_t'|\mathcal{F}_{t-1}) = \boldsymbol{I}_n , \tag{2.16}$$

where $\mathcal{F}_{t-1}$ is the σ-field generated by the random variables $\{\boldsymbol{y}_s, \boldsymbol{u}_s; s \leq t-1\}$. We investigated the conditions for the ergodicity and basic properties of the stationary distributions and their moments (Kunitomo and Sato (1996)). In particular, the necessary and sufficient conditions for the ergodicity when $m = n = 1$ are $A < 1, B < 1$, and $AB < 1$. It should be noted that the conditions $|A| < 1$ and $|B| < 1$ are sufficient, but not necessary for the geometric ergodicity of the stationary SSAR model. This illustrates one of interesting differences between the linear time series models and the non-linear time series models. There are interesting economic interpretations for these differences. For instance, we originally introduced the stationary SSAR model from the reduced form of a disequilibrium econometric model (Kunitomo and Sato (1996)). It seems that the conditions for ergodicity in the disequilibrium econometric model are much weaker than those for the corresponding equilibrium econometric model.

Example 2 : We can illustrate some possible applications by using the multivariate SSAR models. For this purpose, we take $d = 0, m = 2$, $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{o}$, and $\boldsymbol{e}_1^{'}\boldsymbol{A} = \boldsymbol{e}_1^{'}\boldsymbol{B} = (1,0)$, for simplicity. Then by using the coherency condition (2.8) we have the representation

$$\Delta \boldsymbol{y}_t = \boldsymbol{\alpha}(t)\boldsymbol{r}^{'}\boldsymbol{y}_{t-1} + \boldsymbol{D}(t)\boldsymbol{u}_t \; , \tag{2.17}$$

where $\boldsymbol{\alpha}(t)^{'} = (0, -\sum_{i=1}^{2} I_t^{(i)}\sigma_2^{(i)})$, and a 1×2 vector $\boldsymbol{r}^{'}$ and $\sigma_2^{(j)} (j = 1,2)$ are unknown (constant) parameters. We further take $\boldsymbol{D}(t) = (d_{ij}(t))$, $d_{11}(t) = \sigma_1$, $d_{22}(t) = \sigma_2^{(1)} I_t^{(1)} + \sigma_2^{(2)} I_t^{(2)}$, and $d_{12}(t) = d_{21}(t) = 0$. Then the vector $\boldsymbol{r}$ could be called a co-integrated vector in a non-linear sense because the stochastic process defined by

$$x_t = \boldsymbol{r}^{'}\boldsymbol{y}_t$$

is ergodic and stationary if and only if [1]:

$$a < 1, b < 1, ab < 1 \; , \tag{2.18}$$

where $a = 1 - \boldsymbol{e}_2^{'}\boldsymbol{r}\sigma_2^{(1)}$ and $b = 1 - \boldsymbol{e}_2^{'}\boldsymbol{r}\sigma_2^{(2)}$. We then have the 2-dimensional SSAR model in which the first variable of $\boldsymbol{y}_t$ follows a linear $I(1)$ process, the second variable follows a non-linear $I(1)$ process, and two variables are co-integrated in a non-linear sense. This situation may be interesting for some applications in financial time series (see the discussion in Section 3.3). When $\boldsymbol{A} = \boldsymbol{B}$ and $\boldsymbol{D}_1 = \boldsymbol{D}_2$, the form of (2.17) has been called the error-correction representation of a non-stationary linear time series model (see Engle and Granger (1987), for instance).

Example 3 : When $d = 1$, the stochastic process defined by (2.1), (2.2), and (2.3) is non-stationary. In subsequent analysis in this paper we shall mainly forcus on the non-stationary and univariate case, that is, the SSAR model when $m = n = d = 1$. Thus we are extending the stationary SSAR model discussed in Kunitomo and Sato (1994a,b) to a class of the non-stationary SSAR time series models. Since the integrated autoregressive moving-average (ARIMA) process has been a useful class of non-stationary time series models, we can call the stochastic process under consideration a simultaneous switching autoregressive integrated moving-average (SSARIMA) process. In Section 3.2, we shall argue that there are compelling reasons why the non-stationary SSAR model we in-

[1]Since the proof in the present case is similar to that of Theorem 2.3 in Section 5, we omit the detail.

troduce as Example 3 is useful for some applications to analyzing financial time series.

2.2 Characterization of a non-stationary SSAR model

When $\{\boldsymbol{u}_t\}$ in (2.2) is an $I(1)$ process, the stochastic process $\{\boldsymbol{y}_t\}$ is a non-ergodic process. Hence there are basic questions on the properties of the stochastic process defined by $\{\boldsymbol{y}_t\}$ when $d = 1$. By using the representation of (2.11), the time series model for $\{\Delta \boldsymbol{y}_t\}$ can be written as

(2.19)

$$\begin{aligned} \Delta \boldsymbol{y}_t &= \boldsymbol{D}(t)\Delta[\boldsymbol{D}(t)^{-1}\boldsymbol{\mu}(t)] \\ &+ \boldsymbol{D}(t)\boldsymbol{D}(t-1)^{-1}\Delta \boldsymbol{y}_{t-1} \\ &- \boldsymbol{D}(t)[\boldsymbol{D}(t)^{-1}(\boldsymbol{I}_m - \boldsymbol{A}(t))\boldsymbol{y}_{t-1} - \boldsymbol{D}(t-1)^{-1}(\boldsymbol{I}_m - \boldsymbol{A}(t-1))\boldsymbol{y}_{t-2}] \\ &+ \boldsymbol{D}(t)\Delta \boldsymbol{u}_t \,. \end{aligned}$$

Further when $m = 1$ we can simplify some coefficients by the coherency conditions (2.8). In this case we have the relations $\mu(t) = -r_0 D(t)$, and $1 - A(t) = r_1 D(t)$, where $\boldsymbol{r}' = (r_0, r_1)$. Hence we have the following characterization result on $\{\Delta y_t\}$.

Theorem 2.2 : *Suppose $d = m = 1$. Define the non-linear transformation of $\{\Delta y_t\}$ by*

(2.20) $$T(\Delta y_t) = D(t)^{-1}\Delta y_t \,.$$

Then the transformed stochastic process $\{T(\Delta y_t)\}$ satisfies

(2.21) $$T(\Delta y_t) = A(t-1)T(\Delta y_{t-1}) + \Delta u_t \,.$$

The time series model defined by (2.21) has been called the first order threshold autoregressive (TAR) model with MA disturbances in the non-linear time series analysis. From this result we know that $\{\Delta y_t\}$ is slightly different from the TAR(1) model with MA disturbances, which has been known to be useful for application in statistical time series analysis.

From the above discussions, we can deduce some properties of the differenced time series $\{\Delta y_t\}$. We can thus investigate the univariate non-stationary SSAR model when $d = m = 1$ in further detail. In the empirical application we shall report in Section 3, we also include the time trend variable in the univariate SSAR model. Thus the non-linear and non-stationary SSAR model to be considered is given by

(2.22) $$y_t = \begin{cases} A_0 + A_1 t + A_2 y_{t-1} + \sigma_1 u_t & (\text{if} \quad y_t \geq y_{t-1}) \\ B_0 + B_1 t + B_2 y_{t-1} + \sigma_2 u_t & (\text{if} \quad y_t < y_{t-1}) \end{cases} ,$$

where we take $\sigma_i > 0$ $(i = 1, 2)$. Then by the same argument used to obtain (2.8), we can derive the coherency conditions for this model. The resulting conditions can be summarized by

$$(2.23) \quad -\frac{A_0}{\sigma_1} = -\frac{B_0}{\sigma_2} = r_0 \,, \quad -\frac{A_1}{\sigma_1} = -\frac{B_1}{\sigma_2} = r_1 \,, \quad \frac{1 - A_2}{\sigma_1} = \frac{1 - B_2}{\sigma_2} = r_2 \,.$$

Since $\{y_t\}$ is a non-ergodic process, we need to investigate the stochastic process defined by (2.22). For this purpose it is convenient to use the indicator functions $I_t^{(1)} = I(\Delta y_t \geq 0)$ and $I_t^{(2)} = I(\Delta y_t < 0)$. Also we use the notation of $D(t) = \sigma_1 I_t^{(1)} + \sigma_2 I_t^{(2)}$ and re-write the disturbance terms $\{u_t\}$ as

$$(2.24) \qquad u_t = \frac{1}{D(t)} \Delta y_t + r_0 + r_1 t + r_2 y_{t-1} \,.$$

Then, given the information available at $t - 1$, there are four phases for Δy_t at t to be considered depending on $I_t^{(i)}$ and $I_{t-1}^{(i)}$ $(i = 1, 2)$. By taking the difference operation in (2.24) and re-arranging each term, we have the representation

$$(2.25) \qquad \Delta y_t = D(t) \left\{ -r_1 + \left(-r_2 + \frac{1}{D(t-1)} \right) \Delta y_{t-1} + \Delta u_t \right\} \,.$$

Hence the stochastic process $\{\Delta y_t\}$ has the representation

$$(2.26) \Delta y_t = \begin{cases} A_1 + A_2 \Delta y_{t-1} + \sigma_1 \Delta u_t & (\text{if} \quad \Delta y_{t-1} \geq 0, \ \Delta y_t \geq 0) \\ A_1 + \left(\frac{\sigma_1}{\sigma_2}\right) B_2 \Delta y_{t-1} + \sigma_1 \Delta u_t & (\text{if} \quad \Delta y_{t-1} < 0, \ \Delta y_t \geq 0) \\ B_1 + \left(\frac{\sigma_2}{\sigma_1}\right) A_2 \Delta y_{t-1} + \sigma_2 \Delta u_t & (\text{if} \quad \Delta y_{t-1} \geq 0, \ \Delta y_t < 0) \\ B_1 + B_2 \Delta y_{t-1} + \sigma_2 \Delta u_t & (\text{if} \quad \Delta y_{t-1} < 0, \ \Delta y_t < 0) \end{cases} .$$

By this form of representation, we notice that the differenced process $\{\Delta y_t\}$ from the SSIAR model has not only the simultaneous switching characteristic, but also a characteristic of the threshold type time series model. When d=1 and q is finite, $\{y_t\}$ has Markovian representation. More generally, the stochastic process $\{\Delta^d y_t\}$ has similar characteristics when the underlying process $\{u_t\}$ is an $I(d)$. For the stochastic process $\{\Delta y_t\}$ defined by (2.26), we can establish the necessary and sufficient conditions for its ergodicity. A proof is set out in the Appendix.

Theorem 2.3 : *Suppose (i) the order of MA terms q on $\{\Delta u_t\}$ is a finite number, (ii) the coherency condition (2.23) holds, (iii) the density function $g(v)$ of $\{v_t\}$ is everywhere positive in $\boldsymbol{R}^1$, and (iv) $\sup_{t \geq 1} E[|v_t|] < +\infty$. Then the Markov chain defined by (2.26) for $\{\Delta y_t\}$ is ergodic if and only if*

$$(2.27) \qquad A_2 < 1, \ B_2 < 1, \ A_2 B_2 < 1 \,.$$

For the precise definition and discussions on the ergodicity for Markov chains on a general state space, see Tweedie (1975), Liu and Susko (1992), or Meyn and Tweedie (1993). It is interesting to see that the conditions given by (2.27) are identical to the ergodicity conditions for the stationary SSAR(1) model derived in our previous work (Kunitomo and Sato (1996)). However, for current purposes, no additional conditions as we used then on $\{v_t\}$ are necessary.

The non-stationary SSAR given by (2.11) is a complicated stochastic process. In order to get some idea of its statistical properties, we did a set of simulation for the simplest case. When $m = d = 1$, the simplest SSAR model can be re-written as

$$y_t - \mu = \begin{cases} A(y_{t-1} - \mu) + \sigma_1 u_t & \text{if} \quad y_t \geq y_{t-1} \\ \\ B(y_{t-1} - \mu) + \sigma_2 u_t & \text{if} \quad y_t < y_{t-1} \end{cases} , \tag{2.28}$$

where we re-define μ as a location parameter and σ_i $(i = 1, 2)$ as scale parameters. In our notation (2.28) corresponds to (2.22) with $A = A_2, B = B_2$, and $A_1 = B_1 = 0$. The disturbance terms $\{u_t\}$ follow the random walk process satisfying

$$u_t = u_{t-1} + v_t \ . \tag{2.29}$$

The innovation terms $\{v_t\}$ in (2.29) are independently and identically distributed random variables and follow $N(0, 1)$. The condition on coherency in this case is given by

$$\frac{1 - A}{\sigma_1} = \frac{1 - B}{\sigma_2} = r \ . \tag{2.30}$$

For the sake of simplicity, we set $\mu = 0$ in our simulations. Although there are four unknown parameters A, B and σ_i $(i = 1, 2)$ in (2.28), there are only three free parameters A, B and r.

We took several sets of values on these parameters and did a set of simulations in a systematic way. From these we present just three simulated time series data cases in Figure 2.1, which were generated from the same set of random numbers used for $\{v_t\}$ in (2.29). The middle one among three cases shows the sample path of the simulated time series when $A = B = 0.5$, which means that the non-stationary SSAR(1) model is actually the standard ARIMA(1,1,0) model. When $A \neq B$, we notice some asymmetrical patterns in the sample paths of the simulated time series. For economic time series, the case when $A = 0.8$ and $B = 0.2$ may be the most interesting one. Even though we use a very simple non-stationary SSAR model, we can get very interesting asymmetrical patterns on the sample paths of $\{y_t\}$ along the simulated random walk of $\{u_t\}$. This aspect can not be easily realized by the linear non-stationary time series models such as the ARIMA model. Since the dominant factor in the present case is the random walk part, however, it is generally more difficult to distinguish the asymmetrical case from the symmetric case than in the stationary SSAR models. We have investigated the sample paths of the time series generated by the stationary SSAR models in a previous study (Kunitomo and Sato (1996)).

2.3 Maximum Likelihood Estimation

The SSAR model is quite complex as a statistical model in its several aspects when $m = n$ and $d \geq 1$. The first aspect is its similarity to the threshold autoregressive model in that the present state variables depend on the past realized values of time series. Another aspect is that there is a simultaneity between the present phase and the present value of the time series variables. The last aspect is that the SSAR model when $d \geq 1$ is a non-linear and non-stationary stochastic process. As we discussed for the simple stationary SSAR model (Sato and Kunitomo (1996)), the standard least squares estimation method for data set in each phase separately gives inconsistent estimates for its unknown parameters. In addition, our simulations showed that the bias of the least squares estimator is numerically quite significant in most cases. This aspect is quite different from the estimation problem for the standard TAR time series models. The main reason for this is because there is an important simultaneity involved in the SSAR models. Instead of the least squares method, we are proposing to use the maximum likelihood method for the non-stationary SSAR model in this paper.

We shall consider the case when $m = n$ and $d = 1$. We set the initial conditions such that $\boldsymbol{v}_0 = \boldsymbol{v}_{-1} = \cdots = \boldsymbol{v}_{-q} = \boldsymbol{o}$ and $\Delta \boldsymbol{y}_1$ is fixed for the simplicity. Then the Jacobian of the transformation from $\{\Delta \boldsymbol{u}_t, 2 \leq t \leq T\}$ to $\{\Delta \boldsymbol{y}_t, 2 \leq t \leq T\}$ is given by

$$|J(\Delta \boldsymbol{u}_t \rightarrow \Delta \boldsymbol{y}_t)|_+ = \prod_{t=2}^{T} |\boldsymbol{D}(t)|^{-1} . \tag{2.31}$$

The Jacobian of the transformation from $\{\boldsymbol{v}_t, 2 \leq t \leq T\}$ to $\{\Delta \boldsymbol{u}_t, 2 \leq t \leq T\}$ is one provided that (2.2) is an invertible MA process.

Under the assumption that the disturbance terms $\{\boldsymbol{v}_t\}$ are independently and normally distributed random variables, the conditional log-likelihood function when $d = 1$ for $\{\Delta \boldsymbol{y}_t,\ 2 \leq t \leq T\}$ given the initial conditions can be written as

$$\begin{aligned} \log L_T(\boldsymbol{\theta}) &= -\frac{(T-1)m}{2} \log 2\pi \\ &- \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{2} I_t^{(i)} \log |\boldsymbol{D}_i \boldsymbol{\Omega}(\boldsymbol{\theta}) \boldsymbol{D}_i'| \\ &- \frac{1}{2}\sum_{t=2}^{T} \boldsymbol{v}_t'(\boldsymbol{\theta})\ \boldsymbol{\Omega}(\boldsymbol{\theta})^{-1}\ \boldsymbol{v}_t(\boldsymbol{\theta}) , \end{aligned} \tag{2.32}$$

where $\{\boldsymbol{v}_t(\boldsymbol{\theta})\}$ are $\{\boldsymbol{v}_t\}$ rewritten from (2.1) and (2.2) as functions of $\{\Delta \boldsymbol{y}_t\}$, $\boldsymbol{\Omega}(\boldsymbol{\theta})$ is the covariance matrix of $\boldsymbol{v}_t$ whose diagonal elements are ones, and $\boldsymbol{\theta}$ is a vector of structural parameters as appeared in the original SSAR model. When $m = n = 1$, we use the notation $\sigma_i = D_i$ $(i = 1, 2)$ and the parameter vector is given by $\boldsymbol{\theta}' = (r_1, r_2, \sigma_1, \sigma_2, c_1, \cdots, c_q)$.

The maximum likelihood (ML) estimator $\hat{\theta}_{ML}$ can be defined by the maximum of $\log L_T(\boldsymbol{\theta})$ with respect to the unknown parameters in $\boldsymbol{\theta}$, where the parameter space $\boldsymbol{\Theta}$ is restricted by the coherency conditions given by (2.7) and (2.8). By using $\hat{\theta}_{ML}$ and the initial condition, we can also estimate other parameters such as r_0. The asymptotic properties of the ML estimator in the non-stationary SSAR model when $m = n = d = 1$ can be established, that is, the ML estimator is consistent and asymptotically normal. A proof of this is provided in the Appendix.

Theorem 2.4 : *For the non-stationary SSAR model given by (2.22), suppose (i) the sufficient conditions for the coherency in (2.23) and the ergodicity in Theorem 2.3 hold, (ii) the disturbances terms $\{v_t\}$ are independently distributed as $N(0,1)$, (iii) the MA order q is a finite number and (2.2) is invertible, and (iv) $\sigma_i > 0$ $(i = 1, 2)$. Also suppose (v) the true parameter vector $\boldsymbol{\theta}_0$ is an interior point of a compact set $\boldsymbol{\Theta}_0$ in the parameter space $\boldsymbol{\Theta}$. Then the ML estimators $\hat{\boldsymbol{\theta}}_{ML}$ of unknown parameters in $\boldsymbol{\theta}$ are consistent and asymptotically normally distributed as*

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}\right) \xrightarrow{d} N\left[0,\, I_{\boldsymbol{\theta}}^{-1}\right] \quad , \tag{2.33}$$

provided

$$I_{\boldsymbol{\theta}} = \operatorname*{plim}_{T\to\infty} \frac{1}{T}\left[-\frac{\partial^2 \log L_T(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right] \tag{2.34}$$

is a positive definite matrix.

We do not have a simpler form of the information matrix [2] at present. Then we need to use

$$\frac{1}{T}\left[-\frac{\partial^2 \log L_T(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}}$$

as its consistent estimator for statistical inferences.

We have also investigated the finite sample properties of the ML estimator in a systematic way. Because their mathematical expressions are intractable, we have utilized simulation procedures. We generated the simulated time series $\{\Delta y_t\}$ and $\{y_t\}$ for the non-stationary SSAR model when $d = m = 1$ and $q = 0$, i.e. the SSIAR model without time trend. We used the standard normal random numbers for the disturbance terms $\{v_t\}$. Then we obtained tables of the sample mean of the ML estimator from 5,000 replications. Among many tables, we show the numerical results only for the case when $T = 100$ and $T = 500$ in Table 1. The numerical values of the means in our tables should be accurate to two digits at least. From these tables, we find that the bias of the ML estimator is negligible when the sample size is about 100 and the estimates based on the ML estimation are reliable. These findings are very similar to those from our previous investigations on the ML estimation method on a stationary SSAR model (Sato

[2]We conjecture that the asssumption of its positive definiteness can be removed because we can show it in a simple case.

and Kunitomo (1996)). Thus there is strong support for the use of the ML estimation method for the stationary and the non-stationary SSAR models.

2.4 Asymmetry with Conditional Heteroskedasticity

In recent econometric analysis of financial data, considerable attention has been paid to the autoregressive conditional heteroskedasticity (ARCH) model, which was originally proposed by Engle (1982). The critical argument against using the standard ARCH model seems to lie in the fact that the original ARCH model cannot represent the asymmetrical nature of volatility functions of some financial time series. Several generalizations on the original ARCH model and other approaches have been proposed. For instance, see Bollerslev (1986), Nelson (1991), and Harvey and Shepard (1993) on the related problems and technical difficulties in the previous studies.

In this respect, the SSAR modelling provides a natural way to handle the asymmetry with conditional heteroskedasticities. In order to illustrate this point, we take the simple case in (2.1)-(2.4) when $m = d = 1$ and $q = 0$. Let Ω_t $(= E(v_t^2|\mathcal{F}_{t-1}))$ be the conditional heteroskedasticity function of disturbances $\{v_t\}$ for the stochastic process $\{\Delta y_t\}$ represented by

$$\Delta y_t = D(t)\left\{-r_1 + \left(-r_2 + \frac{1}{D(t-1)}\right)\Delta y_{t-1} + v_t\right\}, \tag{2.35}$$

where $E(v_t) = 0$. If we further assume that

$$\Omega_t = 1 + \sum_{i=1}^{p} \alpha_i v_{t-i}^2 \tag{2.36}$$

and the unknown coefficients $\{\alpha_j, j = 1, \cdots, p\}$ satisfy some restrictions on the stationarity and the positivity of conditional variances of $\{v_t\}$, we implicitly include an asymmetrical conditional variance function for $\{\Delta y_t\}$. In this way the ARCH type models for the symmetric conditional heteroskedasticity can be easily incorporated into the SSAR modelling and the resulting volatility function for the original process can be asymmetrical in two phases. If we want to have an efficient estimation procedure of the ARCH effects, the likelihood function in (2.32) when $m = n = d = 1$ should be modified by simply substituting $\sigma_i^2\Omega_t(\boldsymbol{\theta})$ $(i = 1, 2)$ and $v_t(\boldsymbol{\theta})^2\Omega_t(\boldsymbol{\theta})^{-1}$ for $D_i\Omega(\boldsymbol{\theta})D_i'$ and $\boldsymbol{v}_t(\boldsymbol{\theta})'\Omega(\boldsymbol{\theta})^{-1}\boldsymbol{v}_t(\boldsymbol{\theta})$, respectively. The structural parameter vector in this case is given by $\boldsymbol{\theta}' = (r_1, r_2, \sigma_1, \sigma_2, c_1, \cdots, c_q, \alpha_1, \cdots, \alpha_p)$.

3. An Application to Financial Data

3.1 Financial Time Series

The main reason to introduce the non-stationary SSAR model is its applicability to economic time series data. Especially, there has been growing interest among econometricians and statisticians in the last decade to investigate financial time series data by using statistical time series analysis. There have been several interesting features in financial time series data. First, many financial time series such as stock prices, bond prices, interest rates, foreign exchange rates, and their derivatives are often too volatile to use the stationary time series models of statistical time series analysis. The results of the prediction based on the stationary linear time series models have therefore been unsatisfactory. Furthoremore, there is considerable reason in financial economics to believe that there is a martingale measure for the stock prices : see Harrison and Kreps (1979). Second, the distributions of prices and yields are often not well approximated by the Gaussian distribution. It has often been found by econometricians that the kurtosis calculated from the stock returns is much larger than 3. Third, the estimated volatility functions for many financial time series are not constant over time. This leads to the idea that the conditional variances of time series are not constant over time. Fourth, some financial time series including stock prices exhibit asymmetrical movements between in the up-ward phase and in the down-ward phase. These features are not consistent with standard linear time series models such as the autoregressive integrated moving average (ARIMA) process and the standard autoregressive conditional heteroskedasticity (ARCH) process, which have sometimes been used in recent econometric applications.

We should stress that the non-stationary SSAR model introduced in Section 2.2 has statistical properties that are consistent with all of the above observations on many financial time series. Thus we hope that the non-stationary and non-linear time series model we introduced in Section 2 will potentially be useful for applications in many financial data.

Before presenting our empirical application, however, we have examined the asymmetrical property by using a set of Japanese stock indices data as the preliminary data analysis. For this purpose we have used time series data set of the Nikkei 225 spot index from January 1985 to May 1986 and the Nikkei 225 futures index from January 1990 to August 1991 as we shall explain in Section 3.3. We first fit the non-linear regression model [3]

$$\Delta y_t = \beta_0 + \beta_1^+ \Delta y_{t-1}^+ + \beta_1^- \Delta y_{t-1}^- + v_t \tag{3.1}$$

where β_0, β_1^+, and β_1^- are unknown regression coefficients and $\{v_t\}$ are the disturbance terms with $E(v_t|\mathcal{F}_{t-1}) = 0$ and $E(v_t^2|\mathcal{F}_{t-1}) = \sigma^2$. The signed lagged explanatory variables are defined by

$$\Delta y_{t-1}^+ = \begin{cases} \Delta y_{t-1} & \text{if } \Delta y_{t-1} \geq 0 \\ 0 & \text{if } \Delta y_{t-1} < 0 \end{cases} ,$$

and

$$\Delta y_{t-1}^- = \begin{cases} 0 & \text{if } \Delta y_{t-1} \geq 0 \\ \Delta y_{t-1} & \text{if } \Delta y_{t-1} < 0 \end{cases} .$$

[3] We have estimated other non-linear regression models including the case when $(\Delta y_{t-1}^+)^2$ and $(\Delta y_{t-1}^-)^2$ are in the explanatory variables, for example, by using the full data sets explained in Section 3.3. Since the estimated results are similar, however, we have omitted the details.

In particular, if $\beta_1^+ = \beta_1^- = \beta_1$, the asymmetrical terms in the model disappear and we have an AR(1) model for Δy_t.

Using the time series data of the Nikkei 225 indices, we have estimated (3.1) by the standard least squares method and the results of our estimation are summarized in Table 2. From Table 2, we find that the estimated unknown coefficients β_1^+ and β_1^- are significantly different in the spot index equation. The F-value for the hypothesis

$$H_0 : \beta_1^+ = \beta_1^-$$

is 8.306. Furthermore, the estimated coefficient of β_1^+ is significantly different from zero while the estimated coefficient of β_1^- is *not* significantly different from zero. In order to see the degree of model fitting, we have calculated the value of AIC (Akaike's information criterion) in each case. The value of AIC for the regression is -4090.4 while the value of AIC for the $AR(1)$ fitting is -4084.1. We have found that the estimated models with Δy_{t-1}^+ and Δy_{t-1}^- are better than those with only Δy_{t-1} in many sample periods for the spot index by the minimum AIC. Therefore we tentatively conclude that by using the model given by (3.1), we have picked up an asymmetrical property in the spot index data.

On the other hand, the estimated unknown coefficients β_1^+ and β_1^- are *not* significantly different in the futures index equation. The F-value for the hypothesis $H_0 : \beta_1^+ = \beta_1^-$ is 0.19. The value of AIC for the regression is -3237.97 while the value of AIC for the $AR(1)$ fitting is -3239.78. (The t-value for the estimated coefficient is 2.57.) We also have found that the estimated models with Δy_{t-1}^+ and Δy_{t-1}^- are *not* better than those with only Δy_{t-1} in many sample periods for the futures index. Because these estimated results from the threshold type regressions are typical, we tentatively conclude that there are some differences in the non-linear aspect discussed between the spot index and futures index, which may be of interest to financial economists.

3.2 A Simple Model of Stock Prices

In this section we first discuss a simple econometric model of stock prices, which leads mathematically to the non-stationary SSAR model. The main reason for the following discussion is not to develop the financial economics, but to illustrate why the SSAR model is useful and applicable to many financial time series. For this purpose, we slightly modify the well-known economic model in the micro-market structure literature of financial economics developed by Amihud and Mendelson (1987).

Let the intrinsic value of a security at time t and its observed price be V_t and P_t respectively. We distinguish the intrinsic value of a security and its observed price. There are some economic reasons why they can be different : see Amihud and Mendelson (1987) and its references to the the recent literature on micro-market structures in financial economics. Since the two values V_t and P_t can be different, we can introduce a partial-adjustment model when the intrisic value V_t at t deviates from the observed past price P_{t-1} at $t-1$ as follows

$$P_t - P_{t-1} = \begin{cases} g_1(V_t - P_{t-1}) & \text{if} \quad V_t - P_{t-1} \geq 0 \\ g_2(V_t - P_{t-1}) & \text{if} \quad V_t - P_{t-1} < 0 \end{cases} , \tag{3.2}$$

where V_t and P_t are in logarithms and the adjustment coefficients g_i satisfy $g_i \geq 0$ $(i = 1, 2)$.

We note that we have modified the adjustment process used in Amihud and Mendelson (1987) in two ways. First, we have omitted the contemporary noise factor in the right hand side. We did this because of the resulting simplicity. Second, we have allowed the adjustment coefficients g_i $(i = 1, 2)$ to take different values. There could be intuitive economic reasons why they can be different [4]. Because there are new shocks or news available at t in financial markets, V_t could be different from P_{t-1}. When $V_t \geq P_{t-1}$, the intrisic value at t is above the past realized price and there is economic pressure mainly from the demand side to make the price go up. When $V_t < P_{t-1}$, on the other hand, there is economic pressure mainly from the supply side to make the price go down. Since there are two main forces during the actual price determination process in financial markets, the two coefficents g_i $(i = 1, 2)$ could be different. Instead of discussing their details, however, we simply point out that this formulation covers many cases which are theoretically or practically interesting in financial economics. When $g_1 = g_2$, (3.2) is reduced to the standard linear adjustment model. Further, when $g_1 = g_2 = 1$, $V_t = P_t$ and the intrinsic value of a security is always equal to its observed price. Hence, by using the formulation we have adopted in (3.2) it is possible to examine from the observed time series data if these conditions are reasonable descriptions of reality.

In recent financial economics, there has been a convention that the logarithm of the intrisic security value $\{V_t\}$ follows an integrated process $I(1)$ with a drift,

$$V_t = V_{t-1} + \sigma e_t + \mu \ , \tag{3.3}$$

where μ represents the expected daily return and $\{e_t\}$ are a sequence of random variables generated by the linear stationary stochastic process possessing a MA representation.

By combining (3.2) and (3.3), we obtain the representation of ΔP_t as

$$\Delta P_t = g(t)[\frac{1}{g(t-1)} - 1]\Delta P_{t-1} + g(t)[\mu + \sigma e_t] \ , \tag{3.4}$$

where $g(t) = g_1 I_t^{(1)} + g_2 I_t^{(2)}$. In this representation, $I_t^{(1)} = 1$ if and only if $V_t - P_{t-1} \geq 0$. But then (3.2) implies that $I_t^{(1)} = 1$ if and only if $\Delta P_t \geq 0$. Hence (3.4) is a special case of the non-stationary SSAR model we have discussed in Section 2.1 when $m = n = d = 1$.

[4] There are also institutional factors such as the tax system, short-sale restrictions, trading and commission rules, and other regulations in major financial markets. Usually these factors determine the actual transaction costs.

By using Theorem 2.3 in Section 2.2, we have calculated the ergodic region for the process $\{\Delta P_t\}$ with respect to the adjustment coefficients g_i $(i = 1, 2)$: see Figure 3.1. We note that the ergodic region when $g_1 \neq g_2$ is quite large in comparison to when $g_1 = g_2$. This figure may be useful when we interpret the empirical results reported in the next sub-section.

3.3 An Empirical Analysis of Spot and Futures Indices

In this section we shall report an empirical result using the time series data in the Japanese financial markets. In our data analysis we have used time series data set of the Nikkei 225 indices which are the most popular stock price indices traded in Japan. They are the daily closing data of Nikkei Spot and Futures indices from January 1985 to December 1994. Trade on Nikkei index Futures started at the end of 1980s at the Osaka Stock Exchange, so we have used the data for Nikkei Futures from January 1990 to December 1994 [5]. All data were transformed into their logarithms before the estimation of the non-stationary SSAR model. It may be of some interest in financial economists to compare the time series movements of the spot price index and the corresponding futures price index.

Using these data, we have estimated the first order non-stationary univariate SSAR model with time trend given by (3.4), which could be written as SSIAR(1). The estimation of structural parameters in the SSIAR(1) model was conducted by the ML method under the assumption of the normal disturbances. Since we cannot obtain an explicit formula for the ML estimators of unknown parameters, we have used a numerical nonlinear optimization technique with the coherency restrictions on parameters given by (2.23). In the actual estimation we took $q = 0$ because we could not find any significant MA terms in most case. The resulting estimation results are summarized by Table 2. We should note that the estimated values of g_1 and g_2 correspond to $1 - A_2$ and $1 - B_2$, respectively. LK in figures stands for the maximized log-likelihood functions. For the purpose of comparison, we also have estimated the standard $ARIMA(1,1,0)$ process from our time series data set. In order to make a comparison, we have calculated the likelihood ratio statistic $LR(A = B)$ for testing the null hypothesis

$$H_0 \ : \ A_2 = B_2 \ .$$

Under the assumption of the Gaussian disturbances, the likelihood ratio statistic $LR(A_2 = B_2)$ is asymptotically [6] distributed as $\chi^2(1)$. Thus this test statistic

[5] To be more precise, we have constructed the discounted futures series from the original futures data. We have done this by using the monthly average of inter-bank interest rates. The main reason for this adjustment procedure was because the published futures series are discontinuous before the delivery dates roughly once in every three months for the Nikkei 225 futures series. See the report of Bank of Japan (1993) for the details of problems on the stock futures market and data in Japan. Nevertheless, we have found that the following empirical results have not been much affected by the above procedure.

[6] The asymptotic distribution can be derived by using the standard argument in time series analysis and Theorem 2.4.

gives us useful information on the asymmetrical movements of stock prices. The estimated results have been summarized in Table 3 and Figure 3.2.

There are several interesting empirical observations. First, the spot stock price index sometimes shows sharp asymmetrical movements either it is in the upward phase or in the downward phase. This phenomenon has been evident in 1985 and 1987. Actually we already knew that there was a sharp decline in October of 1987. During these sharp downward phases, the estimated values of the adjustment coefficient $d_2 = 1 - B_2$ are often greater than 1, which indicates that there were some over-reacting movements in the stock markets. On the other hand, the estimated coefficient $d_1 = 1 - A_2$ is less than one in the up-ward phases and we could not find any over-reacting movements in the same periods. Second, all of the χ^2–statistics for the futures stock index are not very large from 1990 to 1994. This suggests that the futures stock index does not show any significant asymmetry in two phases in comparison to the spot price index. There could be some economic interpretations for this observation. Third, after starting of the active trade of stock index futures in the Japanese financial markets, there have not been many occasions as was the case previously when the asymmetrical movements of the price indices are evident. In particular, we could not find many over-reacting movements in the downward phases by looking at the estimated adjustment coefficients in this period.

Following the estimation of asymmetrical movements of time series by the SSAR modelling, we can calculate the estimated residuals. In order to see the significance of the conditional heteroskedasticity in our data set, we have estimated the first order ARCH model ($ARCH(1)$) and calculated the likelihood ratio statistic for the hypothesis $H_0 : \alpha_1 = 0$ from the residuals. Among 7 sub-period the ARCH effects are significant in 4 cases with 1% significance level by using the χ^2 distribution. We have often observed strong evidence for the existence of significant conditional heteroskedasticity. It seems that the asymmetrical volatility function based on the SSAR models often fits the data of the Nikkei Index better than the standard ARCH model fitting. In this case the resulting estimated volatility functions for the Nikkei Index are asymmetrical in two phases. However, we should emphasize that the standard ARCH modelling can be easily implemented in the SSAR modelling.

These empirical problems and findings may have some implications for financial economists. Needless to say, these observations from our empirical results on the Japanese financial markets are preliminary and further considerations are needed. But clearly it has not been easy to detect these features of the financial time series data by using the existing methods and the linear time series modelling in particular.

4. Conclusions

In this paper we have focused on one important aspect in many financial economic time series, which has been often ignored in the past econometric studies. We have argued that the asymmetrical pattern in the movements of time series between the upward and downward phases often observed by economists can not be represented properly by the stationary and non-stationary linear time series models including the standard ARMA, ARIMA, and ARCH processes, which have been used in many empirical studies in the past.

We have therefore introduced the class of simultaneous switching autoregressive (SSAR) models, which is one type of non-linear switching time series models. It has the distinctive properties of simultaneity and time irreversibility. Since we have already investigated the stationary SSAR model (Kunitomo and Sato (1996)), we have focussed on the non-stationary SSAR model and investigated some of its properties in the univariate case. In this paper we have proposed the maximum likelihood estimation method for estimating the unknown parameters in the SSAR model. We hope that the results reported in this paper may shed new light on the time series properties often observed by many economists and statisticians.

We have also tried to show that there are some natural reasons why the non-stationary SSAR model introduced in Section 2 is a useful tool to analyze many financial time series in financial markets. We have illustrated this issue by suggesting a very simple model for stock price movements in Section 3.2. The point is that if we permit the intrisic value of security to be different from the observed price and have an adjustment process, the result is a new non-linear time series model. Then the estimated coefficients in the upward and downward phases can be different, and we can get some interesting information from the estimated adjustment coefficients and the resulting χ^2 statistics. We have illustrated this advantageous aspect of our modelling approach by examining the movements of the Nikkei stock index and the Nikkei futures stock index from 1985 to 1994 in the Japanese financial markets. Of course, there can be many possibilities to describe financial time series by non-stationary and non-linear time series modelling. At least we can conclude that the non-linear and non-stationary models we introduced in this paper give a class of interesting econometric and statistical models, which are useful for possible applications.

However, there are several important issues that remain to be solved. In this paper we have only investigated some special cases of the non-stationary SSAR model. In particular, there are some interesting situations when we have multivariate non-linear time series as illustrated in Example 2 in Section 2.3. Also we have shown that the conditional heteroskedasticities such as the ARCH model can easily be incorporated into the SSAR modelling and the resulting volatility function can be asymmetrical in two phases. Since there can be many non-linear time series models as we indicated in the Introduction and Section 2.4, however, the comparison or the discrimination of the SSAR models from other linear and non-linear statistical models will be necessary. Further studies will be necessary on these problems.

5. Mathematical Appendix

In this appendix, we gather some mathematical details which we have omitted in the previous sections.

Proof of Theorem 2.1 : Let $\boldsymbol{Y}_{1|0}^{(i)}$ $(i = 1, 2)$ be the partition of the sample space for $\boldsymbol{y}_1$ in $\boldsymbol{R}^m$ given $\boldsymbol{y}_0$, which is defined by the indicator functions $I_1^{(i)}$ $(i = 1, 2)$. We can then successively define $\boldsymbol{Y}_{t|0}^{(j)}$ $(j = 1, \cdots, 2^t)$ as the partition of the sample space for $\{\boldsymbol{y}_s, 1 \le s \le t\}$ in $\boldsymbol{R}^{mt}$ by $I_s^{(i)}$ $(s = 1, \cdots, t; i = 1, 2)$. By this sequence of partitions of the sample space, we have $\cap_i \boldsymbol{Y}_{t|0}^{(i)} = \phi$ and $\cup_i \boldsymbol{Y}_{t|0}^{(i)} = \boldsymbol{R}^{mt}$ for any $t > 0$.

Next, we use the indicator functions

(A.1) $$J_t^{(1)} = 1_{\{\boldsymbol{e}_m' \boldsymbol{D}_1 \boldsymbol{u}_t \ge \boldsymbol{e}_m'(\boldsymbol{I}_m - \boldsymbol{A})\boldsymbol{y}_{t-1} - \boldsymbol{e}_m' \boldsymbol{\mu}_1\}}$$

and

(A.2) $$J_t^{(2)} = 1_{\{\boldsymbol{e}_m' \boldsymbol{D}_1 \boldsymbol{u}_t < \boldsymbol{e}_m'(\boldsymbol{I}_m - \boldsymbol{A})\boldsymbol{y}_{t-1} - \boldsymbol{e}_m' \boldsymbol{\mu}_1\}} .$$

Let $\boldsymbol{U}_{1|0}^{(i)}$ $(i = 1, 2)$ be the partition of the sample space for $\boldsymbol{u}_1$ in $\boldsymbol{R}^m$ given $\boldsymbol{y}_0$, which is defined by the indicator functions $J_1^{(i)}$ $(i = 1, 2)$. We can also successively define $\boldsymbol{U}_{t|0}^{(i)}$ $(i = 1, \cdots, 2^t)$ as the partition of the sample space for $\{\boldsymbol{u}_s, 1 \le s \le t\}$ in $\boldsymbol{R}^{mt}$ by $J_s^{(i)}$ $(s = 1, \cdots, t; i = 1, 2)$ for any $t > 0$. Then under the assumptions of Theorem 2.1 it is straightforward to show that $\cap_i \boldsymbol{U}_{t|0}^{(i)} = \phi$, $\cup_i \boldsymbol{U}_{t|0}^{(i)} = \boldsymbol{R}^{mt}$, and the correspondence between $\boldsymbol{Y}_{t|0}^{(i)}$ $(i = 1, \cdots, 2^t)$ and $\boldsymbol{U}_{t|0}^{(i)}$ $(i = 1, \cdots, 2^t)$ is one-to-one. *Q.E.D.*

Proof of Theorem 2.3 : We shall use a method similar to the one used by Liu and Susko (1992) for the TAR(1) model with MA disturbances, which is based on a fixed-point theorem. However, we note that substantial changes in their method are necessary and we can establish stronger results than theirs because of the different features of the non-stationary SSAR model with MA disturbances.

(i) Sufficiency : Let $x_t = \Delta y_t$ and define $(1 + q) \times 1$ vector $\boldsymbol{X}_t$ by

(A.3) $$\boldsymbol{X}_t = \begin{pmatrix} x_t \\ v_t \\ v_{t-1} \\ \vdots \\ v_{t-q+1} \end{pmatrix} .$$

Then we consider the Markovian representation for $\{\boldsymbol{X}_t\}$. For the sake of simplicity, we set $r_1 = 0$. The condition $x_t \ge 0$ is equivalent to $v_t \ge \boldsymbol{a}_{t-1}' \boldsymbol{X}_{t-1}$, where

(A.4) $$\boldsymbol{a}_{t-1}' = (r_2 - \frac{1}{D(t-1)}, -c_1, -c_2, \cdots, -c_q)$$

and $\{c_j, 1 \le j \le q\}$ are the MA coefficients of $\{\Delta \boldsymbol{u}_t\}$. From (2.25) we have the representation

(A.5) $$\boldsymbol{X}_t = \boldsymbol{H}(\boldsymbol{X}_{t-1}, v_t) ,$$

where

$$\text{(A.6)} \qquad \boldsymbol{H}(\boldsymbol{X}_{t-1}, v_t) = \begin{pmatrix} -D(t)\boldsymbol{a}'_{t-1}\boldsymbol{X}_{t-1} + D(t)v_t \\ v_t \\ v_{t-1} \\ \vdots \\ v_{t-q+1} \end{pmatrix} .$$

We use the criterion function

$$\text{(A.7)} \qquad G(\boldsymbol{\xi}) = \sum_{i=1}^{q} h(\xi_i) ,$$

where $h(\xi_i) = |\xi_i|$ and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_q)$.
If we use a sequence of compact sets $K_n = [-n, n] \times \cdots \times [-n, n], n = 1, 2, \cdots$, then $\inf_{\boldsymbol{\xi} \in K_n^c} G(\boldsymbol{\xi}) \to +\infty$ as $n \to +\infty$. For $t \geq q$,

$$\text{(A.8)} \qquad \begin{aligned} E[G(\boldsymbol{X}_t)|\boldsymbol{X}_0] &= E[h(x_t) + \sum_{j=0}^{q-1} h(v_{t-j})|\boldsymbol{X}_0] \\ &\leq k_1 + E[E[h(x_t)|\boldsymbol{X}_{t-1}]|\cdots|\boldsymbol{X}_0] \end{aligned}$$

because $E[|v_t|] < +\infty$, where k_1 is a positive constant.

Let

$$\text{(A.9)} \qquad \begin{aligned} Q_{t|t-1} &= E[h(x_t)|\boldsymbol{X}_{t-1}] \\ &= E\left\{h[-D(t)\boldsymbol{a}'_{t-1}\boldsymbol{X}_{t-1} + D(t)v_t|\boldsymbol{X}_{t-1}]\right\} . \end{aligned}$$

We first consider the case when $x_{t-1} = x > 0$. In this case from (2.26) we have two phases at t given $x > 0$ and

$$\text{(A.10)} \quad \begin{aligned} Q_{t|t-1} &= \sigma_1 \int_{z \geq (r_2 - \frac{1}{\sigma_1})x - \boldsymbol{c}'\boldsymbol{z}_{t-1}} [-(r_2 - \frac{1}{\sigma_1})x + \boldsymbol{c}'\boldsymbol{z}_{t-1} + z]g(z)dz \\ &- \sigma_2 \int_{z < (r_2 - \frac{1}{\sigma_1})x - \boldsymbol{c}'\boldsymbol{z}_{t-1}} [-(r_2 - \frac{1}{\sigma_1})x + \boldsymbol{c}'\boldsymbol{z}_{t-1} + z]g(z)dz, \end{aligned}$$

where $\boldsymbol{c}' = (c_1, \cdots, c_q)$ and $\boldsymbol{z}_{t-1} = (v_{t-1}, \cdots, v_{t-q})$. Then by using (2.23), we have

$$\text{(A.11)} \quad \begin{aligned} Q_{t|t-1} \leq k_2(1 + \sum_{i=1}^{q} |v_{t-i}|) &+ A_2 x \int_{z \geq (r_2 - \frac{1}{\sigma_1})x - \boldsymbol{c}'\boldsymbol{z}_{t-1}} g(z)dz \\ &- \left(\frac{\sigma_2}{\sigma_1}\right) A_2 x \int_{z < (r_2 - \frac{1}{\sigma_1})x - \boldsymbol{c}'\boldsymbol{z}_{t-1}} g(z)dz, \end{aligned}$$

where k_2 is a positive constant. We note that we have the inequality

$$\text{(A.12)} \qquad -\frac{\sigma_2}{\sigma_1} A_2 < 1 .$$

This is because $A_2B_2 < 1$ and the coherency condition (2.23) implies

$$0 < r_2 < \frac{1}{\sigma_1} + \frac{1}{\sigma_2} \ . \tag{A.13}$$

Then we take $\delta_1 = \max\ \{A_2, -(\sigma_2/\sigma_1)A_2\}$. When $x_{t-1} = x < 0$, we can use the similar arguments and take $\delta_2 = \max\ \{B_2, -(\sigma_1/\sigma_2)B_2\}$. Then by taking $\delta = \max\{\delta_1, \delta_2\}$, we have the relation

$$E[G(\boldsymbol{X}_t)|\boldsymbol{X}_{t-1}] \le k_3 \left(1 + \sum_{i=1}^{q} |v_{t-i}|\right) + \delta G(\boldsymbol{X}_{t-1}) \ , \tag{A.14}$$

where $0 < \delta < 1$ and k_3 is a positive constant. By repeating this procedure, we have

$$E[G(\boldsymbol{X}_t)|\boldsymbol{X}_0] \le k_3 \sum_{k=0}^{t-1} \delta^k \left(1 + \sum_{j=1}^{q} \eta_{t-k-j}\right) + \delta^t G(\boldsymbol{X}_0) \ , \tag{A.15}$$

where $\eta_k = E[|v_k|] I(k > 0) + |v_k| I(k \le 0)$. Hence we have established the boundedness condition on the criterion function

$$\sup_{t \ge 1} E[G(\boldsymbol{X}_t)|\boldsymbol{X}_0] < +\infty \ . \tag{A.16}$$

Next, by using the following Lemma A.1 and a similar argument to Lemma 2.1 in Liu and Susko (1992), the Markov chain defined by (2.26) in the SSIAR(1) model satisfies the additional key condition in Liu and Susko (1992) (their Assumption 2.1). Thus there exists a finite positive invariant measure for the Markov chain $\{\boldsymbol{X}_t\}$ by Theorem 1.1 of Liu and Susko (1992). Also because we have assumed that $g(v)$ is everywhere positive in $\boldsymbol{R}$ and (2.26), we can show that the Markov chain is irreducible. The proof of Lemma A.1 is the result of a straightforward calculation using (2.26) and is therefore omitted.

Lemma A.1 : *Let $\{v_t\}$ in the SSIAR model given by (2.22) be independently and identically distributed random variables with the density function $g(v)$, which is everywhere positive in $\boldsymbol{R}$. Then given $(\Delta y_{t-1}, v_{t-1}, \cdots, v_{t-q}) = (z_0, z_1, \cdots, z_q)$, the conditional probability*

$$\Pr\{\Delta y_t \le x | \Delta y_{t-1} = z_0, v_{t-1} = z_1, \cdots, v_{t-q} = z_q\} \tag{A.17}$$

is a continuous function of $\boldsymbol{z}' = (z_0, z_1, \cdots, z_q)$.

(ii) Necessity : Without loss of generality we take $q = 0$. The essential part of the proof is similar to that for the TAR(1) model given by Chan et. al. (1985). However, there is one aspect in which we have to modify their proof for our model.

We have to consider the situation when the values of parameters are on their boundaries. For an illustration, we consider the case when $A_2 = 1$, $A_1 < 0$ and $B_2 < 1$. By using the coherency condition in this case these conditions imply $1 - \sigma_1 r_2 = 1$, $-\sigma_1 r_1 < 0$, and $1 - \sigma_2 r_2 < 1$. Then we have $\sigma_1 r_2 = 0$, $\sigma_1 r_1 > 0$, and $\sigma_2 r_2 > 0$. However, they are contradictory when $\sigma_1 > 0$ and $\sigma_2 > 0$. Other boundary cases can be treated similarly. *Q.E.D.*

Since we shall use some probability convergence arguments in the proof of Theorem 2.4, we need some results on the existence of moments for $\{\Delta y_t\}$. For this purpose, we prepare Lemma A.2, which is also of independent interest.

Lemma A.2 : *In the SSIAR model given by (2.22), assume (i) the coherency conditions (2.23), (ii) the ergodicity conditions (2.27), and (iii)* $\sup_{t\geq 1} E[|v_t|^k] < +\infty$ *for some* $k \geq 1$. *Then*

$$\sup_{t\geq 1} E[|\Delta y_t|^k] < +\infty \, . \tag{A.18}$$

Proof of Lemma A.2 : The method of proof is similar to that for the sufficiency part of Theorem 2.3 and we shall first show (A.18) for $k = 2$. Let $x_t = \Delta y_t$. Then from the assumptions we made in Lemma A.2, $E[|x_t|] < +\infty$ by (A.16) in Theorem 2.3. We take the criterion function $g(\boldsymbol{X}_t) = |x_t|^2$. We consider the case when $x_{t-1} = x > 0$ and evaluate $E[x_t^2|\mathcal{F}_{t-1}]$. We use the truncation argument by the event $x_{t-1} > M$ for some $M > 0$. Then

$$\begin{aligned}
&E[x_t^2|x_{t-1}, v_{t-1}, \cdots, v_{t-q}] \\
&= \sigma_1^2 \int_{z\geq(r_2-\frac{1}{\sigma_1})x-\boldsymbol{c}'\boldsymbol{z}_{t-1}} \left[-(r_2-\frac{1}{\sigma_1})x + \boldsymbol{c}'\boldsymbol{z}_{t-1} + z\right]^2 g(z)dz \\
&\quad +\sigma_2^2 \int_{z<(r_2-\frac{1}{\sigma_1})x-\boldsymbol{c}'\boldsymbol{z}_{t-1}} \left[-(r_2-\frac{1}{\sigma_1})x + \boldsymbol{c}'\boldsymbol{z}_{t-1} + z\right]^2 g(z)dz \\
&\leq k_4(M)[(1+x_{t-1})(1+\sum_{i=1}^{q}|v_{t-i}|) + \sum_{i=1}^{q} v_{t-i}^2] \\
&\quad +A_2^2 x_{t-1}^2 \int_{z\geq(r_2-\frac{1}{\sigma_1})x_{t-1}-\boldsymbol{c}'\boldsymbol{z}_{t-1},x_{t-1}>M} g(z)dz \\
&\quad +\left(-\frac{\sigma_2}{\sigma_1}A_2\right)^2 x_{t-1}^2 \int_{z<(r_2-\frac{1}{\sigma_1})x_{t-1}-\boldsymbol{c}'\boldsymbol{z}_{t-1},x_{t-1}>M} g(z)dz,
\end{aligned} \tag{A.19}$$

where $k_4(M)$ is a positive constant depending on $M > 0$. There are three cases to be considered for the coefficients of x_{t-1}^2. When $0 < r_2 < 1/\sigma_1$, we have $-\sigma_2/\sigma_1 < -(\sigma_2/\sigma_1)A_2 < 0$ and $1 > A_2 > 0$. In this situation the second integral in (A.19) can be small if we take a sufficiently large M. When $r_2 = 1/\sigma_1$, we have $-(\sigma_2/\sigma_1)A_2 = A_2 = 0$. When $1/\sigma_1 < r_2 < (1/\sigma_1) + (1/\sigma_2)$, we have $0 < -(\sigma_2/\sigma_1)A_2 < 1$ and $0 > A_2 > -(\sigma_1/\sigma_2)$. In this situation the first integral in (A.19) can be small if we take a sufficiently large M. Hence by taking a sufficiently large $M > 0$, there exists δ_1 such that $0 < \delta_1 < 1$ and

$$\text{(A.20)} \qquad E[x_t^2|x_{t-1}, v_{t-1}, \cdots, v_{t-q}]$$
$$\leq k_4(M)[(1+x_{t-1})(1+\sum_{i=1}^{q}|v_{t-i}|) + \sum_{i=1}^{q} v_{t-i}^2] + \delta_1 x_{t-1}^2 \,.$$

Also when $x_{t-1} = x < 0$, we use the truncation argument for the corresponding integrals to (A.19) by the event $x_{t-1} < -M < 0$ and we can find δ_2 $(0 < \delta_2 < 1)$ for the above inequalities. Then we can take $\delta = \max\{\delta_1, \delta_2\}$ and $0 < \delta < 1$. By repeating this procedure on conditional expectations, we have

$$\text{(A.21)} \qquad E[x_t^2|\mathcal{F}_0]$$
$$\leq k_5(M)\sum_{k=0}^{t-1}\delta^k E[(1+x_{t-1-k})(1+\sum_{i=1}^{q}|v_{t-i-k}|) + \sum_{i=1}^{q} v_{t-i-k}^2|\mathcal{F}_0] + \delta^t x_0^2 \,,$$

where $k_5(M)$ is a positive constant depending on $M > 0$. Because we can show that $E[|x_{t-1}v_{t-1}|]$ is bounded by using the assumptions we have made in Lemma A.3, we obtain that $\sup_{t\geq 1} E[x_t^2] < +\infty$.

Next, for an arbitrary $k > 2$, we take the criterion function $G(\boldsymbol{X}_t) = |x_t|^k$ and use the induction with respect to k. The remaining arguments are similar to those for $k = 2$. *Q.E.D.*

Proof of Theorem 2.4 : The method of our proof is similar to the one used in Sato and Kunitomo (1994). However, substantial modifications are necessary because the SSIAR model with the MA disturbances is different from the stationary SSAR model in several important aspects.

(i) Consistency : Let the stochastic process $\{\Delta u_t(\boldsymbol{\theta})\}$ be defined by

$$\text{(A.22)} \qquad \Delta u_t(\boldsymbol{\theta}) = D(t)^{-1}\Delta y_t + r_1 + [r_2 - D(t-1)^{-1}]\Delta y_{t-1} \,,$$

which is identical to $\{\Delta u_t\}$ in (2.25). We denote the vector of true parameter values of $\boldsymbol{\theta}' = (r_1, r_2, \sigma_1, \sigma_2, c_1, \cdots, c_q)$ as $\boldsymbol{\theta}_0' = (r_1^{(0)}, r_2^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)}, c_1^{(0)}, \cdots, c_q^{(0)})$. By substituting Δy_t evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ into (A.22), we have

$$\text{(A.23)} \qquad \begin{aligned} \Delta u_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0) &= D(t)^{-1}D_t^{(0)}\Delta \boldsymbol{u}_t(\boldsymbol{\theta}_0) \\ &+ [r_1 - D(t)^{-1}D_t^{(0)}r_1^{(0)}] \\ &+ [r_2 - D(t)^{-1}D_t^{(0)}r_2^{(0)}]\Delta y_{t-1}(\boldsymbol{\theta}_0) \\ &+ [D(t)^{-1}D_t^{(0)}D_{t-1}^{(0)-1} - D(t-1)^{-1}]\Delta y_{t-1}(\boldsymbol{\theta}_0), \end{aligned}$$

where $D_t^{(0)} = D(t)$ evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\Delta y_{t-1}(\boldsymbol{\theta}_0) = \Delta y_{t-1}$ evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

Let also $\Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = (\Delta u_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0), 2 \leq t \leq T)$ be a $(T-1) \times 1$ vector. Then we can write

$$\text{(A.24)} \qquad \frac{1}{T} \sum_{t=2}^{T} v_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0)^2 = \frac{1}{T} \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0)' \boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta}) \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0),$$

where $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ is the $(T-1) \times (T-1)$ covariance matrix of $\{\Delta u_t(\boldsymbol{\theta})\}$ and $v_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ is defined by the right hand side. Under the assumption on the initial conditions we have the decomposition $\boldsymbol{\Sigma}_T(\boldsymbol{\theta}) = \boldsymbol{K}_T \boldsymbol{K}_T'$, where $\boldsymbol{K}_T$ is a lower triangular matrix with 1 in its diagonal elements by normalization and $|\boldsymbol{\Sigma}_T(\boldsymbol{\theta})| = 1$.

Let

$$\text{(A.25)} \quad Q(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = \operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{t=2}^{T} \sum_{i=1}^{2} I_t^{(i)} \log \sigma_i - \frac{1}{2} \operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{t=2}^{T} v_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0)^2 .$$

Because the probability limits in (A.25) exist under the assumptions we have made, we shall consider the criterion function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$. First, we notice that $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ is a quadratic function of (r_1, r_2), and a concave function of (σ_1, σ_2) by the following Lemma A.3. (The proof of Lemma A.3 is a result of direct calculations.) Then for $\boldsymbol{r}' = (r_1, r_2)$,

$$\begin{aligned}
\text{(A.26)} \qquad \frac{\partial Q}{\partial \boldsymbol{r}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= (-1) \operatorname*{plim}_{T \to \infty} \frac{1}{T} \frac{\partial \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0)'}{\partial \boldsymbol{r}} \boldsymbol{\Sigma}_T(\boldsymbol{\theta})^{-1} \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \\
&= (-1) \operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{s,t=2}^{T} \left(\begin{array}{c} 1 \\ \Delta y_{t-1}(\boldsymbol{\theta}_0) \end{array} \right) \sigma_T^{t,s} \Delta u_s(\boldsymbol{\theta}_0) \\
&= (-1) \operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{s=2}^{T} \sum_{t=2}^{s} \beta_{s-t} \left(\begin{array}{c} 1 \\ \Delta y_{t-1}(\boldsymbol{\theta}_0) \end{array} \right) v_s(\boldsymbol{\theta}_0),
\end{aligned}$$

where $v_s(\boldsymbol{\theta}_0) = v_s(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, $\sigma_T^{t,s}$ is the (t, s) component of $\boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta})$, and $\beta_{s-t} = \boldsymbol{e}_t'(T-1) \boldsymbol{K}_T'^{-1} \boldsymbol{e}_s(T-1)$ $(s \geq t)$ for a $(T-1) \times 1$ vector $\boldsymbol{e}_t(T-1)$ with 1 in its t-th component and zeros in other components. By using the invertibility condition for the MA process, we have $|\beta_k| = O(\rho^k)$ $(0 < \rho < 1)$. Since $v_s(\boldsymbol{\theta}_0)$ $(s \geq t)$ are uncorrelated with $\mathcal{F}_{t-1}$, we can use the standard truncation arguments for the last summation in (A.26). Then we have

$$\text{(A.27)} \qquad \frac{\partial Q}{\partial \boldsymbol{r}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0 .$$

Let $\eta_i = 1/\sigma_i$ $(i = 1, 2)$. Then

$$\begin{aligned}
\text{(A.28)} \qquad & \frac{\partial Q}{\partial \eta_1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\
&= \left(\operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{t=2}^{T} I_t^{(1)}\right) \frac{1}{\eta_1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} - \operatorname*{plim}_{T \to \infty} \frac{1}{T} \frac{\partial \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0)'}{\partial \eta_1} \boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta}) \Delta \boldsymbol{u}(\boldsymbol{\theta}, \boldsymbol{\theta}_0)
\end{aligned}$$

$$= \sigma_1^{(0)}(\underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{t=2}^{T} I_t^{(1)}) - \sigma_1^{(0)} \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{s,t=2}^{T} I_t^{(1)} w_t(\boldsymbol{\theta}_0)\sigma_T^{t,s}\Delta u_s(\boldsymbol{\theta}_0)$$

$$+ \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{s,t=2}^{T} I_{t-1}^{(1)}\Delta y_{t-1}(\boldsymbol{\theta}_0)\sigma_T^{t,s}\Delta u_s(\boldsymbol{\theta}_0) ,$$

where the stochastic process $w_t(\boldsymbol{\theta}_0)$ is defined by

$$\text{(A.29)} \quad w_t(\boldsymbol{\theta}_0) = \Delta u_t(\boldsymbol{\theta}_0) - [r_1^{(0)} + r_2^{(0)}\Delta y_{t-1}(\boldsymbol{\theta}_0)] + D_{t-1}^{(0)-1}\Delta y_{t-1}(\boldsymbol{\theta}_0) .$$

The third term of (A.29) is zero by similar arguments to (A.27). Because $\Delta u_t(\boldsymbol{\theta}_0) = \sum_{j=0}^{q} c_j^{(0)} v_{t-j}(\boldsymbol{\theta}_0)$ ($c_j^{(0)}$ are the true MA parameter values), the probability limit of the first two terms is given by

$$\text{(A.30)} \quad Q^*(\boldsymbol{\theta}_0)$$

$$= \sigma_1^{(0)}(\underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{s=2}^{T} I_s^{(1)}) - \sigma_1^{(0)} \underset{T\to\infty}{\text{plim}} \frac{1}{T}\sum_{s=2}^{T}\sum_{t=2}^{s} I_s^{(1)}\beta_{s-t} w_t(\boldsymbol{\theta}_0)v_s(\boldsymbol{\theta}_0)$$

$$= (-\sigma_1^{(0)})E[I_s^{(1)}\{v_s(\boldsymbol{\theta}_0)^2 + \xi_{s-1}(\boldsymbol{\theta}_0)v_s(\boldsymbol{\theta}_0) - 1\}] ,$$

where

$$\text{(A.31)} \quad \xi_{s-1}(\boldsymbol{\theta}_0) = \sum_{j=1}^{q} c_j^{(0)} v_{s-j}(\boldsymbol{\theta}_0) - r_1^{(0)} + [-r_2^{(0)} + D_{s-1}^{(0)-1}]\Delta y_{s-1}(\boldsymbol{\theta}_0) .$$

By using the normality assumption on $\{v_t\}$, we have the relation

$$\text{(A.32)} \qquad \int_c^{+\infty} (v^2 - 1 - cv)\phi(v)dv = 0 ,$$

where c is a constant and $\phi(\cdot)$ is the density of the standard normal distribution. We can also use the arguments for η_2 as η_1. Hence we conclude that

$$\text{(A.33)} \qquad \frac{\partial Q}{\partial \eta_i}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0 \; (i = 1, 2) .$$

Lemma A.3 : *Let $g(\boldsymbol{x})$ be defined by*

$$\text{(A.34)} \qquad g(\boldsymbol{x}) = \sum_{i=1}^{p} a_i \log x_i - \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x},$$

where $\boldsymbol{x} = (x_i)$ is a $p \times 1$ vector, $x_i > 0, a_i > 0$ $(i = 1, \cdots, p)$, and $\boldsymbol{A}$ is a non-negative definite matrix. Then

$$\text{(A.35)} \qquad g(c\boldsymbol{x} + (1-c)\boldsymbol{y}) > cg(\boldsymbol{x}) + (1-c)g(\boldsymbol{y})$$

for any $0 < c < 1$, where $\boldsymbol{y} = (y_i)$ is a $p \times 1$ vector with $y_i > 0$ $(i = 1, \cdots, p)$.

Next, we have to deal with the MA parameters in the SSIAR model. For this purpose, we set $r_i = r_i^{(0)}$ and $\sigma_i = \sigma_i^{(0)}$ $(i = 1, 2)$. Then by using (A.23), we have $\Delta u_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = \Delta u_t(\boldsymbol{\theta}_0)$. By applying Proposition 10.8.3 of Brockwell and Davis (1991), we have

$$\text{(A.36)} \qquad \operatorname*{plim}_{T\to\infty} \frac{1}{T} \Delta \boldsymbol{u}(\boldsymbol{\theta}_0)' \boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta}) \Delta \boldsymbol{u}(\boldsymbol{\theta}_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{c(\lambda; \boldsymbol{\theta})}{c(\lambda; \boldsymbol{\theta}_0)} d\lambda \, ,$$

where $\Delta \boldsymbol{u}(\boldsymbol{\theta}_0) = (\Delta u_t(\boldsymbol{\theta}_0))$ is a $(T-1)\times 1$ vector and $c(\lambda; \boldsymbol{\theta})$ is the spectral density function of $\{\Delta u_t(\boldsymbol{\theta})\}$. Then by using Proposition 10.8.1 of Brockwell and Davis (1991), (A.36) is uniquely maximized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Hence the function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ is uniquely maximized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. In addition, the probability convergence in (A.25) and (A.36) is uniform in $\boldsymbol{\Theta}_0 \subset \boldsymbol{\Theta}$ with respect to $\boldsymbol{\theta}$. Then by applying Theorem 4.1.1 of Amemiya (1985), we have the consistency of $\hat{\boldsymbol{\theta}}_{ML}$.

(ii) Asymptotic Normality : In order to prove the asymptotic normality of the ML estimator in the SSIAR model with the MA disturbances under the assumptions we have made, the most important step is the martingale property of the partial derivatives of the log-likelihood function summarized in the following Lemma A.4. The second step is to use the central limit theorem for martingales. (See Hall and Heyde (1980), or Anderson and Kunitomo (1992), for instance.) The rest of the proof is similar to the arguments used in Sato and Kunitomo (1994) by making use of Lemma A.2. For instance, it is straightforward to show

$$\text{(A.37)} \qquad E\left[-\frac{\partial^2 \log L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right] = E\left[\frac{\partial \log L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \log L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] .$$

Hence we omit the details. *Q.E.D.*

Lemma A.4 : *Let $\boldsymbol{\theta}$ be a vector of unknown parameters in the SSIAR model given by (2.25) except r_0 . Then we have*

$$\text{(A.38)} \qquad E\left[\frac{\partial \log L_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} | \mathcal{F}_{t-1}\right] = \frac{\partial \log L_{t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}},$$

where $\mathcal{F}_{t-1}$ is the $\sigma-$field generated by $\{\boldsymbol{y}_s, s \leq t-1; \boldsymbol{v}_s, s \leq t-1\}$.

A Sketch Proof of Lemma A.4 : Using the notations in (i), the conditional log-likelihood function for $\{\Delta y_s, 2 \leq s \leq t\}$ is proportional to

$$\text{(A.39)} \qquad \begin{aligned} L_t(\boldsymbol{\theta}) \ \propto \ & -\frac{1}{2} \log |\boldsymbol{\Sigma}_t(\boldsymbol{\theta})| \\ & - \frac{1}{2} \sum_{s=1}^{t} \sum_{i=1}^{2} I_s^{(i)} \log \sigma_i^2 - \frac{1}{2} \boldsymbol{h}_t(\boldsymbol{\theta})' \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta}) \boldsymbol{h}_t(\boldsymbol{\theta}), \end{aligned}$$

where $\boldsymbol{h}_t(\boldsymbol{\theta})$ is a $(t-1) \times 1$ vector with $\Delta u_s(\boldsymbol{\theta})$ $(2 \leq s \leq t)$ in its s-th component, and $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ is the covariance matrix of $\boldsymbol{h}_t(\boldsymbol{\theta})$. Under the assumption on the initial conditions, we have $|\boldsymbol{\Sigma}_t(\boldsymbol{\theta})| = 1$.

Take $\theta_1 = \sigma_1$ for example. In this case

$$\text{(A.40)} \qquad \frac{\partial \log L_t(\boldsymbol{\theta})}{\partial \theta_1} = -\frac{1}{\sigma_1}\sum_{s=1}^{t} I_s^{(1)} - \boldsymbol{h}_t(\boldsymbol{\theta})' \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{h}_t(\boldsymbol{\theta})}{\partial \sigma_1} ,$$

where

$$\text{(A.41)} \qquad \frac{\partial \Delta u_s(\boldsymbol{\theta})}{\partial \sigma_1} = -\frac{1}{{\sigma_1}^2} I_s^{(1)} \Delta y_s + \frac{1}{{\sigma_1}^2} I_{s-1}^{(1)} \Delta y_{s-1} .$$

By using (2.25), we have the relations

$$\text{(A.42)} \qquad (\frac{1}{\sigma_1}) I_s^{(1)} \Delta y_s = I_s^{(1)} (v_s(\boldsymbol{\theta}) + \xi_{s-1}(\boldsymbol{\theta})) ,$$

and

$$\text{(A.43)} \quad \boldsymbol{h}_t(\boldsymbol{\theta})' \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{h}_t(\boldsymbol{\theta})}{\partial \sigma_1}$$

$$= \sum_{s=2}^{t} v_s(\boldsymbol{\theta}) \sum_{s'=2}^{s} \boldsymbol{e}_s'(t) \boldsymbol{K}_t^{-1} \boldsymbol{e}_{s'}(t) [-\left(\frac{1}{\sigma_1}\right)^2 I_{s'}^{(1)} \Delta y_{s'} + \left(\frac{1}{\sigma_1}\right)^2 I_{s'-1}^{(1)} \Delta y_{s'-1}] .$$

Then by using the relation

$$\text{(A.44)} \qquad E\left\{ v_t(\boldsymbol{\theta}) \boldsymbol{e}_t'(t) \boldsymbol{K}_t'^{-1} \boldsymbol{e}_t(t) [(-\frac{1}{\sigma_1}) I_t^{(1)} (v_t(\boldsymbol{\theta}) + \xi_{t-1}(\boldsymbol{\theta})) + \left(\frac{1}{\sigma_1}\right)^2 I_{t-1}^{(1)} \Delta y_{t-1}] | \mathcal{F}_{t-1} \right\}$$

$$= E\left\{ v_t(\boldsymbol{\theta}) (-\frac{1}{\sigma_1}) I_t^{(1)} (v_t(\boldsymbol{\theta}) + \xi_{t-1}(\boldsymbol{\theta})) | \mathcal{F}_{t-1} \right\} ,$$

and (A.32), we have the martingale property

$$\text{(A.45)} \qquad E\left[\frac{\partial \log L_t(\boldsymbol{\theta})}{\partial \sigma_1} | \mathcal{F}_{t-1} \right]$$

$$= \frac{\partial \log L_{t-1}(\boldsymbol{\theta})}{\partial \sigma_1} + \frac{1}{\sigma_1} E\left[(v_t(\boldsymbol{\theta})^2 - 1) I_t^{(1)} + \xi_{t-1}(\boldsymbol{\theta}) v_t(\boldsymbol{\theta}) I_t^{(1)} | \mathcal{F}_{t-1} \right]$$

$$= \frac{\partial \log L_{t-1}(\boldsymbol{\theta})}{\partial \sigma_1} .$$

For other parameters in $\boldsymbol{\theta}$, it is straightforward to show (A.38) by similar arguments. *Q.E.D.*

References

[1] Amemiya, T. (1985), *Advanced Econometrics*, Basil-Blackwell.

[2] Amihud, Y. and Mendelson, H. (1987), "Trading Mechanisms and Stock Returns : An Empirical Investigation", *Journal of Finance*, Vol.XLII, **3**, 533-555.

[3] Anderson, T.W. and Kunitomo, N. (1992), "Asymptotic Distributions of Regression and Autoregression Coefficients with Martingale Differences", *Journal of Multivariate Analysis*, **40**, 221-243.

[4] Bank of Japan Monthly Bulletin (1993), "The Report on Functions of Financial Futures and Problems in Japan", January Issue (in Japanese).

[5] Bollerslev,T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, **31**, 307-327.

[6] Brockwell, P.J. and Davis, R.A. (1991), *Time Series: Theory and Method*, 2nd Edition, Springer-Verlag.

[7] Chan, K.S., Petruccelli, J. D., Tong, H. and Woolford, S. W. (1985), "A Multiple Threshold AR(1) model", *Journal of Applied Probability*, **22**, 267-279.

[8] Engle, R. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variances of U.K. Inflation," *Econometrica*, **50**, 987-1008.

[9] Engle, R. and Granger, C.W.J. (1987), "Cointegration and Error Correction: representation, estimation, and testing," *Econometrica*, **55**, 2, 251-276.

[10] Gourieroux, J., Laffont,J., and Monfort,A. (1980), "Coherency Conditions in Simultaneous Linear Equation Models with Endogenous Switching Regimes," *Econometrica*, **48**, 3, 675-695.

[11] Granger, C. W. and Andersen, A. P. (1978), *An Introduction to Bilinear Time Series Models*, Vanderhoek and Ruprecht, Gottingen.

[12] Hall, P. and Heyde, C. (1980), *Martingale Limit Theory and its Applications*, Academic Press.

[13] Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press.

[14] Harrison, J.M. and Kreps, D.M. (1979), "Martingales and Arbitrage in Multi-period Securities Markets," *Journal of Economic Theory*, **20**, 381-408.

[15] Harvey, A.C. and Shephard, N. (1993), "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns," Unpublished manuscript.

[16] Kunitomo, N. and Sato, S. (1994), "A Non-linearity in Economic Time Series and Disequilibrium Econometric Models", in Takeuchi, K. and Takemura, A. (ed.), *Theory and Application of Mathematical Statistics*, University of Tokyo Press, (in Japanese).

[17] Kunitomo, N. and Sato, S. (1996), "Asymmetry in Economic Time Series and Simultaneous Switching Autoregressive Model", *Structural Change and Economic Dynamics*, **7**, 1-34.

[18] Liu, J. and Susko, E. (1992), "On Strict Stationarity and Ergodicity of a Non-linear ARMA Model", *Journal of Applied Probability*, **29**, 363-373.

[19] Meyn, S.P. and Tweedie, R.L. (1993), *Markov Chain and Stochastic Stability*, Springer-Verlag.

[20] Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, **59**, 2, 347-370.

[21] Ozaki, T. and Oda, H. (1978), "Non-linear Time Series Model Identification by Akaike's Information Criterion", *Proceedings of IFAC Workshop on Information and Systems*, Campiegn, France.

[22] Petruccelli, J. D. and Woolford, S. W. (1984), "A Threshold AR(1) model", *Journal of Applied Probability*, **24**, 207-286.

[23] Sato, S. and Kunitomo, N. (1996), "Some Properties of the Maximum Likelihood Estimator in Simultaneous Switching Autoregressive Model", *Journal of Time Series Analysis*, **17**, 287-307.

[24] Tong, H. (1983), *Threshold Models in Non-linear Time Series Analysis*, Lecture Notes in Statistics No. 21, Springer, Heidelberg.

[25] Tong, H. (1990), *Non-linear Time Series*, Oxford University Press, Oxford.

[26] Tweedie, R.L. (1975), "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space," *Stochastic Processes and their Applications*, **3**, 385-403.

Table 1: The mean of the ML estimator of SSIAR model[7]

T = 100

	B = 0.8		B = 0.2		B = 0.0		B = -0.2		B = -1.5	
	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$
A =	0.7922	0.7921	0.7946	0.1844	0.7932	-0.0220	0.7957	-0.2181	0.7984	-1.5410
0.8	(0.059)	(0.058)	(0.044)	(0.169)	(0.043)	(0.203)	(0.038)	(0.230)	(0.027)	(0.475)
A =	0.1796	0.7912	0.1966	0.2011	0.1945	-0.0107	0.1915	-0.2006	0.2036	-1.5354
0.2	(0.173)	(0.044)	(0.117)	(0.115)	(0.110)	(0.144)	(0.100)	(0.155)	(0.061)	(0.336)
A =	-0.0215	0.7952	-0.0060	0.1943	-0.0062	-0.0045	-0.0008	-0.1966	0.0013	-1.5138
0.0	(0.203)	(0.041)	(0.133)	(0.108)	(0.130)	(0.129)	(0.117)	(0.142)	(0.070)	(0.300)
A =	-0.2159	0.7979	-0.2060	0.1958	-0.1957	-0.0028	-0.2032	-0.2049	-0.1933	-1.5185
-0.2	(0.230)	(0.038)	(0.158)	(0.099)	(0.150)	(0.120)	(0.140)	(0.139)	(0.082)	(0.291)
A =	-1.5277	0.7986	-1.5065	0.2031	-1.5052	0.0044	-1.4935	-0.2007	NA	NA
-1.5	(0.452)	(0.027)	(0.346)	(0.063)	(0.304)	(0.070)	(0.289)	(0.083)	(NA)	(NA)

[7] In Table 1, A and B correspond to A_2 and B_2 respectively with $A_1 = B_1 = 0$ in (2.22). The value in parentheses shows the root mean squared error. "NA" corresponds to the case when it is not ergodic. We did not have investigated the ML estimator in this case.

T = 500

	B = 0.8		B = 0.2		B = 0.0		B = -0.2		B = -1.5	
	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$	$\hat{A}$	$\hat{B}$
A =	0.7963	0.7964	0.7987	0.1929	0.7998	-0.0014	0.7990	-0.2033	0.7996	-1.4972
0.8	(0.029)	(0.028)	(0.019)	(0.072)	(0.018)	(0.087)	(0.017)	(0.101)	(0.012)	(0.210)
A =	0.1949	0.7982	0.1992	0.1997	0.2006	-0.0011	0.1997	-0.2024	0.2004	-1.5079
0.2	(0.075)	(0.020)	(0.052)	(0.052)	(0.048)	(0.062)	(0.043)	(0.066)	(0.027)	(0.139)
A =	-0.0042	0.7987	-0.0014	0.1998	-0.0069	-0.0014	0.0025	-0.1991	-0.0002	-1.4945
0.0	(0.084)	(0.018)	(0.061)	(0.046)	(0.061)	(0.056)	(0.053)	(0.066)	(0.031)	(0.133)
A =	-0.2046	0.7994	-0.1998	0.1986	-0.1987	0.0004	-0.2003	-0.1969	-0.2000	-1.5007
-0.2	(0.100)	(0.016)	(0.072)	(0.045)	(0.063)	(0.051)	(0.061)	(0.062)	(0.035)	(0.131)
A =	-1.5146	0.7997	-1.5081	0.1992	-1.5015	-0.0020	-1.5035	-0.1996	NA	NA
-1.5	(0.193)	(0.011)	(0.149)	(0.028)	(0.132)	(0.032)	(0.126)	(0.036)	(NA)	(NA)

Table 2: Estimated Results of Non-linear Regression

Spot Index Equation

	β_1^+	β_1^-
Estimate	0.3847	-0.0185
S.D.	0.0841	0.0864
t-value	4.572	-0.2137

Futures Index Equation

	β_1^+	β_1^-
Estimate	0.1626	0.0957
S.D.	0.0927	0.08955
t-value	1.7533	1.0683

Table 3: Estimated Results: Nikkei Index

Spot 1985 – 1989

	period	SSIAR(1)			ARIMA(1,1,0)		
		A_2	B_2	LK	$A_2(=B_2)$	LK	χ^2
1	1985.01.04-1985.09.10	0.207	-0.093	737.63	0.067	732.70	9.843 **
2	1985.09.11-1986.05.30	0.343	0.216	744.78	0.298	743.57	2.429
3	1986.05.31-1987.02.20	0.243	0.108	640.22	0.191	639.12	2.213
4	1987.02.23-1987.11.07	0.233	-0.951	555.16	-0.118	518.30	73.733 **
5	1987.11.09-1988.08.03	0.178	0.172	651.99	0.176	651.99	0.004
6	1988.08.04-1989.05.15	0.103	-0.093	732.00	0.027	730.34	3.326
7	1989.05.16-1989.12.29	-0.028	-0.074	605.20	-0.047	605.13	0.130

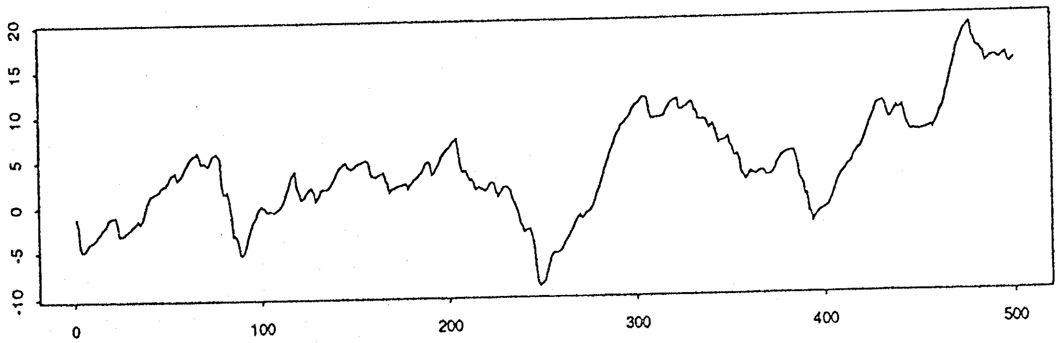
* 10% significance — $\chi^2(1)$
** 1% significance — $\chi^2(1)$

Spot 1990 – 1994

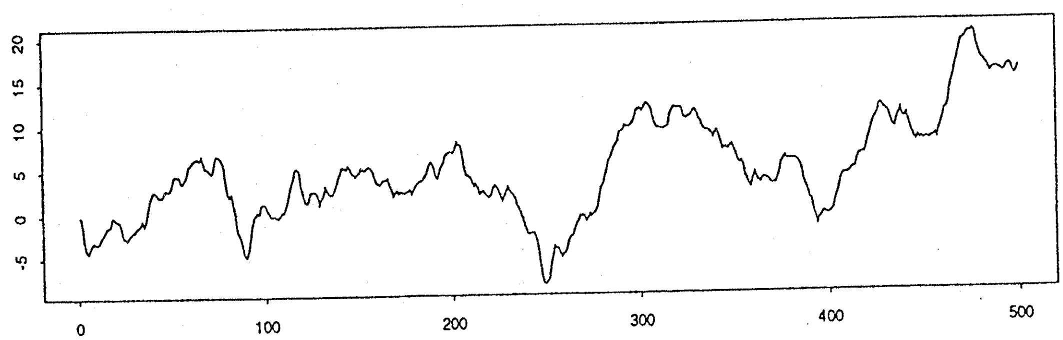
	period	SSIAR(1)			ARIMA(1,1,0)		
		A_2	B_2	LK	$A_2(=B_2)$	LK	χ^2
1	1990.01.04-1990.10.22	0.034	0.196	485.74	0.106	484.30	2.877
2	1990.10.23-1991.08.15	-0.026	0.072	565.59	0.020	565.15	0.878
3	1991.08.16-1992.06.11	-0.080	0.065	530.93	-0.005	530.00	1.847
4	1992.06.12-1993.03.31	-0.011	0.026	527.12	0.006	527.06	0.121
5	1993.04.01-1994.01.24	0.040	-0.069	570.08	-0.013	569.57	1.013
6	1994.01.25-1994.12.15	-0.151	0.066	693.63	-0.059	691.40	4.465 *

Futures 1990 – 1994

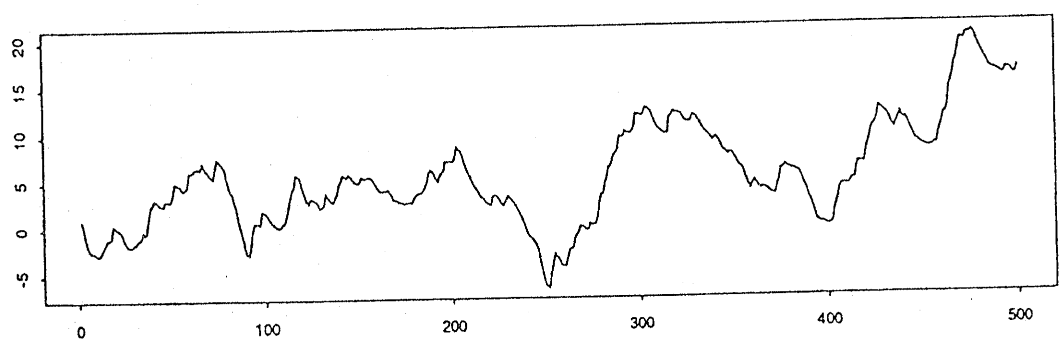
	period	SSIAR(1)			ARIMA(1,1,0)		
		A_2	B_2	LK	$A_2(=B_2)$	LK	χ^2
1	1990.01.04-1990.10.22	0.199	0.240	505.53	0.219	505.43	0.206
2	1990.10.23-1991.08.15	-0.020	-0.040	562.79	-0.030	562.77	0.035
3	1991.08.16-1992.06.11	0.076	0.172	551.36	0.126	550.83	1.045
4	1992.06.12-1993.03.31	0.059	0.006	538.57	0.033	538.45	0.236
5	1993.04.01-1994.01.24	-0.027	-0.084	565.06	-0.053	564.93	0.257
6	1994.01.25-1994.12.15	-0.103	-0.102	698.90	-0.103	698.90	0.000



A= 0.8 B= 0.2 r= 1



A= 0.5 B= 0.5 r= 1



A= 0.2 B= 0.8 r= 1

Figure 2.1: The sample paths of SSIAR(1)

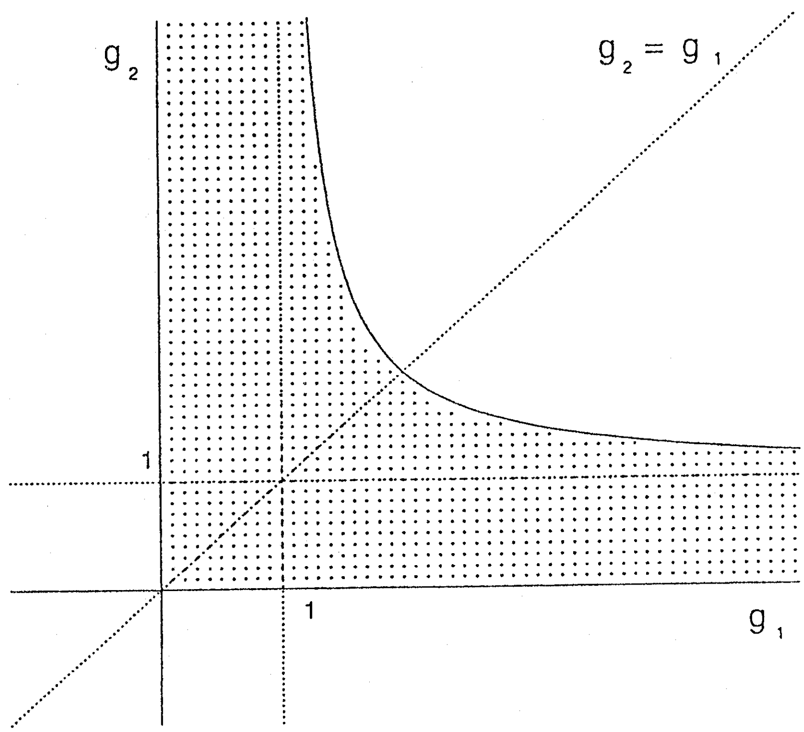
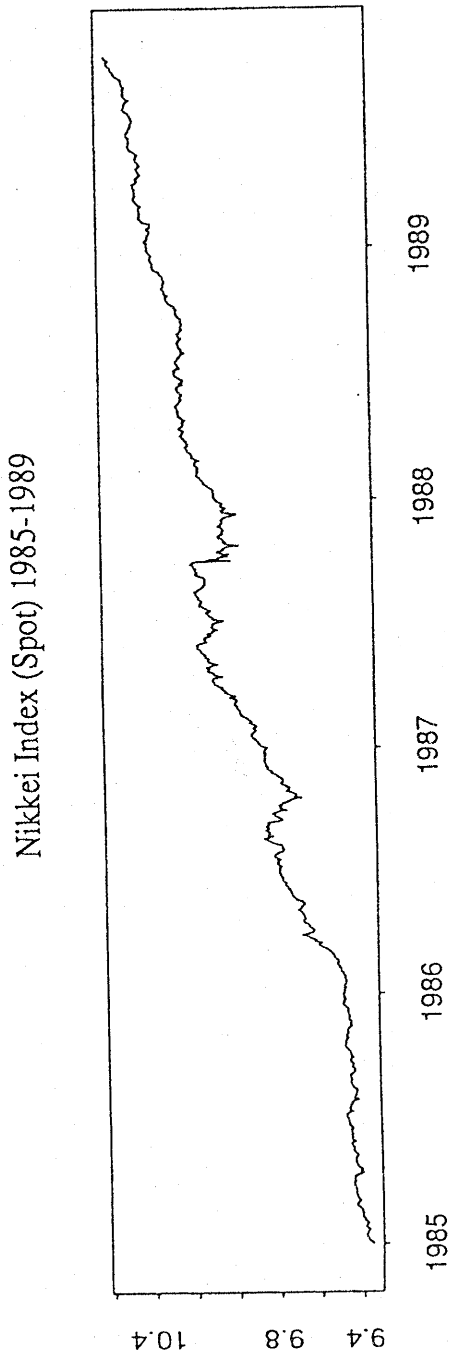


Figure 3.1: The region of ergodicity



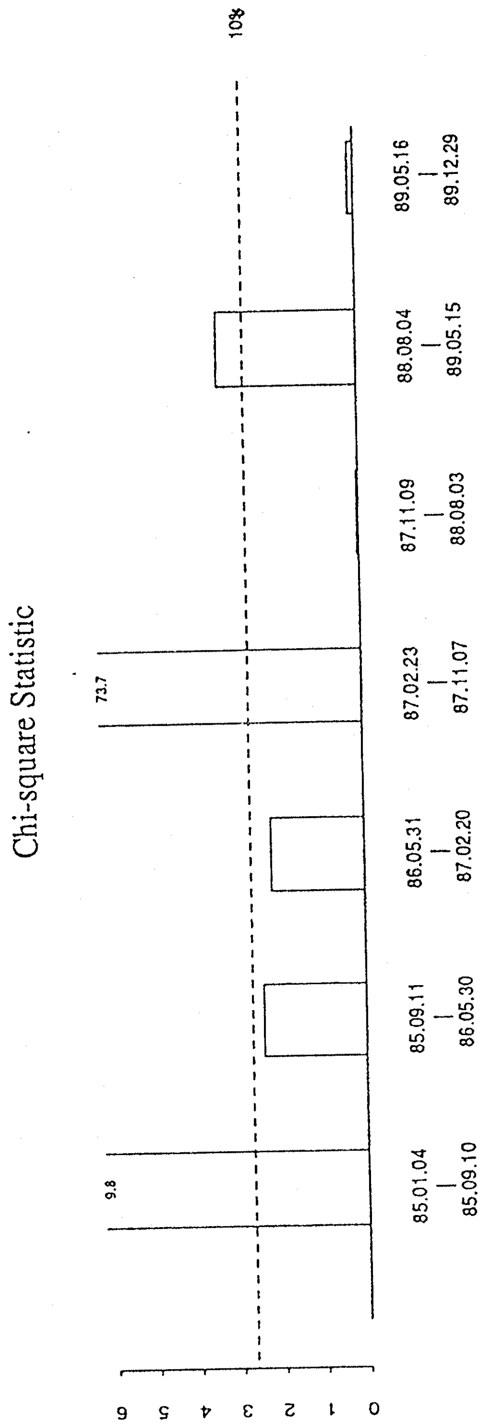


Figure 3.2: Result of Nikkei Index 225

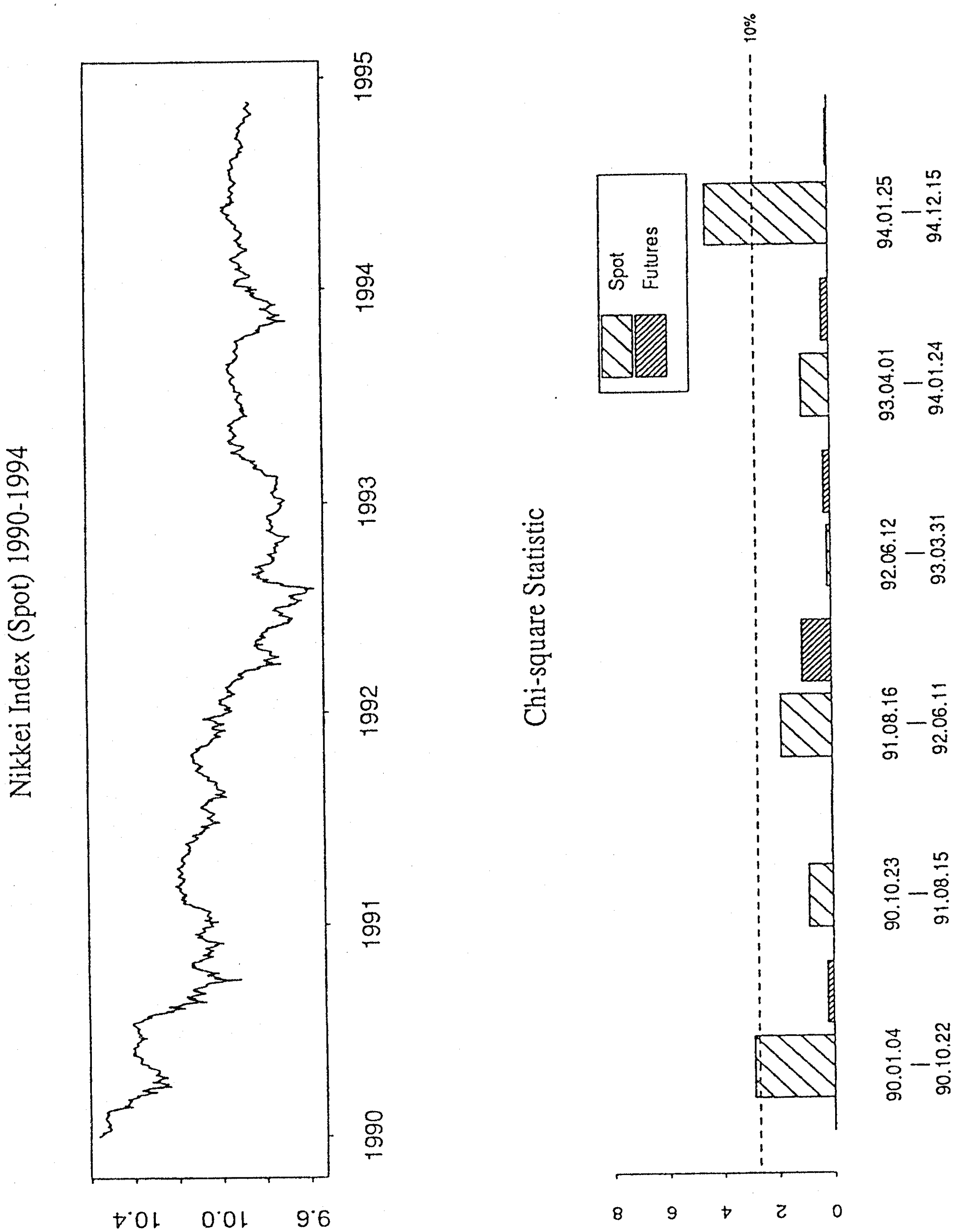


Figure 3.2: Result of Nikkei Index 225 (continued)