97-F-25

Estimation of Asymmetrical Volatility for Asset Prices: The Simultaneous Switching ARIMA Approach

Naoto Kunitomo
Faculty of Economics, University of Tokyo
and
Seisho Sato
Institute of Statistical Mathematics

August 1997

Estimation of Asymmetrical Volatility for Asset Prices : The Simultaneous Switching ARIMA Approach *

Naoto Kunitomo†
and
Seisho Sato ‡

July 1997

Abstract

The asymmetrical movement between the downward and upward phases of the sample paths of many financial time series has been commonly noted by economists. Since this feature cannot be described by the Autoregressive Integrated Moving-average (ARIMA) model and the Autoregressive Conditional Heteroskedastic (ARCH) model, we introduce a class of the Simultaneous Switching Autoregressive Integrated Moving-Average (SSARIMA) model with ARCH disturbances. The asymmetrical volatility function of financial time series with daily effects can easily be estimated by this modelling. We also report a simple empirical result on stock price daily indices of the Nikkei-225 and SP-500.

Key Words

Asymmetrical Volatility, Stock Prices, Simultaneous Switching ARIMA Model, Conditional Heteroskedasticity, Daily Effect.

---

*This paper is a revised version of Discussion Paper No.96-E-24, Institute for Monetary and Economic Studies, Bank of Japan. The research was conducted while the first author was a visiting research fellow at the Institute of Monetary and Economic Studies, Bank of Japan. We thank Dr. Kunio Okina for some suggestions on interpreting our empirical findings and the participants of the seminar at Osaka University for helpful comments.

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

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

1. Introduction

In the past decade, several non-linear time series models have been proposed by econometricians. In particular, considerable attention has been paid to the autoregressive conditional heteroskedasticity (ARCH) model, which was originally proposed by Engle (1982). A number of extensions of the standard ARCH model have been proposed and some of them have been used in empirical studies. In addition, some non-parametric and semi-parametric estimation methods for the conditional heteroskedasticities in asset prices and returns have been proposed. The related issues of modelling the conditional heteroskedasticities in asset prices and their volatilities have been discussed by Bollerslev (1986) and Chapter 21 of Hamilton, for instance.

In this paper we shall propose to use a new class of non-linear time series models called the simultaneous switching autoregressive integrated moving-average (SSARIMA) models for analyzing stock prices and possibly other asset prices. In particular, we shall use the SSARIMA model with the autoregressive conditional heteroskedastic (ARCH) disturbances to estimate the volatility functions of stock prices. The main reason for using this class of non-linear time series models is because we are convinced that the class of autoregressive integrated moving-average (ARIMA) time series model and the standard ARCH models cannot describe one important aspect in some financial time series, that is, the asymmetrical movement in the upward phase (or regime) and in the downward phase (or regime). It has often been observed and argued that major financial series including stock prices display some kind of asymmetrical movements in the upward and downward phases. This feature of time series can be regarded as one form of time irreversibility discussed in the statistical time series analysis : see Chapter 4 of Tong (1990).

Earlier, we introduced a 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

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

where A, B, σ_i $(\sigma_i > 0, i = 1, 2)$ are scalar unknown coefficients, and $\{v_t\}$ are a sequence of i.i.d. random variables with $E(v_t) = 0$ and $E(v_t^2) = 1$. By imposing the coherency condition given by

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

this time series model has the Markovian representation

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

where γ is an unknown parameter and $I(\cdot)$ is the indicator function. When $\sigma_1 = \sigma_2 = \sigma$, then the Simultaneous Switching Autoregressive (SSAR) model becomes the standard $AR(1)$ model by re-parametrizing $A = B = 1 - \sigma\gamma$. As we have shown (Kunitomo and Sato (1996)), even this simplest univariate SSAR model, called SSAR(1), gives us some explanations and descriptions of a very important aspect of the asymmetrical movement of time series in two different phases. Although this characteristic of economic time series has been observed by a number of economists, as yet there has not been a simple but general time series model incorporating this feature directly as far as we know in the econometric literature. It is important to note that the SSAR(1) model defined by (1.1)-(1.3) is different from the threshold ARMA models in non-linear time series analysis. The simplest first order threshold autoregressive model is given by

$$(1.4) \qquad y_t = \begin{cases} Ay_{t-1} + \sigma_1 v_t & \text{if} \quad y_{t-d} \geq 0 \\ \\ By_{t-1} + \sigma_2 v_t & \text{if} \quad y_{t-d} < 0 \end{cases} ,$$

where d is a positive integer parameter. The most important issue in connection with the SSAR models in (1.1)-(1.3) is in the fact that the phase at t is determined when y_{t-d} is observed in the threshold autoregressive (TAR) models. Hence once we have an observation at $t - d$ we can perfectly predict the phase at t in the future. See Tong (1990) for the details on the threshold autoregressive models and other related non-linear time series models.

In a subsequent work (Kunitomo and Sato (1997)), we have extended the basic SSAR model that we proposed in Kunitomo and Sato (1996) into important directions for econometric applications. We have introduced a class of non-stationary SSAR models, which is useful for applications to major financial time series. Also we allow that the disturbance terms in the SSAR model can be autocorrelated and have a finite moving-average structure in addition to the non-linear autoregressive part in the time series models. It is important to note that it is not possible to describe the kind of asymmetrical patterns in the upward and downward phases by using the standard linear non-stationary time series models including the ARIMA time series model and the standard ARCH model proposed by Engle (1982). This issue has been previously pointed out by Nelson (1991) and Harvey and Shephard (1993) in the context of the volatility function of stock prices. Although there have been some proposed estimation methods for volatility functions by Nelson (1991) and Harvey and Shephard (1993), the time series models often become very complicated once the asymmetrical forms of the volatility functions are introduced. Since the stationary SSARMA and non-stationary SSARIMA models are natural extensions of the corresponding ARMA and ARIMA models in some sense, they can easily be extended to handle the asymmetrical conditional heteroskedasticities.

The main purpose of this paper is to give a new estimation method of the volatility functions of asset prices, which can be asymmetrical in two phases, by the use of the SSARIMA modelling with ARCH disturbances. Unlike other parametric, non-parametric, and semi-parametric methods already available, our formulation is a very simple parametric approach and it provides an easy way to handle daily effects in the volatility functions of asset prices. Since these effects have been observed by some financial economists and practitioners in financial markets, our method for estimating the volatility functions may have very real applications. Furthermore, our estimation method for the simultaneous switching autoregressive integrated (SSARI) models can be justified by its asymptotic properties and hence the model selection procedure within a class of SSARIMA models based on the information criteria can be developed rather straightforwardly.

In Section 2, we shall introduce the univariate SSARIMA model with ARCH disturbances. We also shall investigate some properties of the SSARIMA(p, q) model with a time trend and ARCH(r) disturbances in some detail and give the method of estimating the asymmetrical conditional heteroskedasticities in asset prices. We shall also discuss the asymptotic properties of the estimation method and develop the model selection procedure. Then in Section 3, we shall apply the SSARI model with a time trend for the analysis of the Nikkei 225 spot index at Tokyo and the SP 500 spot index at New York. In Section 4, some concluding remarks on our econometric approach and empirical findings will be given. The proofs of some theoretical results obtained in this paper will be gathered in the Appendix.

2. The SSARIMA Model

There has been growing interest in the last decade among econometricians to investigate financial time series data by using statistical time series analysis. There are several interesting features often observed in financial time series data. First, many financial time series such as stock prices, bond prices, interest rates, and foreign exchange rates are often too volatile to use the stationary time series models used in standard statistical time series analysis. There are cogent arguments in financial economics that there are equivalent martingale measures for asset prices : see Duffie (1992). Then there are some reasons why it is difficult to describe the observed asset prices by stationary time series models. Second, the distributions of financial prices, yields, and returns are often not well approximated by the Gaussian distribution. It has often been found that the kurtosis calculated from the daily returns for stock prices is much larger than 3, for instance. However, there is little consensus on the class of distributions to be used for describing financial time series among econometricians. Third, the estimated historical volatility functions for many financial time series are often not constant over time. This leads to the argument for that the conditional variances of time series are not constant over time. Fourth, some financial time series exhibit asymmetrical movements in the upward phase and the downward phase. In particular, a number of economists have observed this type of asymmetrical time series movement in stock prices.

The standard linear time series models such as the autoregressive integrated moving average (ARIMA) process can go towards explaining the first and second features, but not to the third and fourth features. The standard autoregressive conditional heteroskedasticity (ARCH) process, which was originally proposed by Engle (1982) and has been sometimes used in recent econometric applications, is consistent with the second and third features, but not with the fourth one. There have thus been several attempts to extend the standard ARCH model. (See Hamilton (1994), for instance.)

2.1 A Simple Model of Stock Prices

In this section we first extend the simple econometric model of stock prices that we used in Kunitomo and Sato (1997). We began by modifying the well-known economic model in 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. 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 anticipated price V_t^*, which is given by

$$(2.1) \qquad P_t - P_{t-1} = \begin{cases} g_1(V_t - V_t^*) & \text{if} \quad V_t - V_t^* \geq 0 \\ g_2(V_t - V_t^*) & \text{if} \quad V_t - V_t^* < 0 \end{cases} .$$

The anticipated price V_t^* at t calculated by the past realizations of prices is given by

$$(2.2) \qquad V_t^* = \sum_{i=0}^{p-1} \beta_i P_{t-i-1}$$

and the adjustment coefficients in the stock price equation satisfy the condition $g_i \geq 0$ $(i = 1, 2)$.

Earlier, we dealt with the case when $V_t^* = P_{t-1}$ (Kunitomo and Sato (1997)), which is included in (2.2) as a special case. The anticipated price V_t^* in (2.2) includes the optimal forecasts of price levels given the past information if P_t could have followed some ARIMA models. Because there are new shocks or news available at t in markets, V_t could be different from V_t^*. In addition, 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. For instance, when $V_t \geq V_t^*$ the current price has been under-evaluated and there is economic pressure mainly from the demand side to force the price up. On the other hand, when $V_t < V_t^*$ the current price has been over-evaluated and there is economic pressure mainly from the supply side to push the price down. Since there are two main forces during the actual price determination process in financial markets, the two coefficients g_i $(i = 1, 2)$ could be different. However, instead of discussing the rigorous economic justifications, here we simply point out that this formulation includes many situations as special cases, which are theoretically or practically interesting in financial economics. When $g_1 = g_2$, (2.1) is reduced to the standard linear adjustment model. Further, when $g_1 = g_2 = 1$ and $V_t^* = P_{t-1}$, then $V_t = P_t$

and the intrinsic value of a security is always equal to its observed price. Hence, by using the formulation we have adopted in (2.1) it is possible to examine from the observed time series data if these conditions are reasonable descriptions of reality.

In the 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,

$$V_t = V_{t-1} + \sigma e_t + \mu \ , \tag{2.3}$$

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

$$I_t^{(1)} = I\left(P_t \geq P_{t-1}\right)$$

and

$$I_t^{(2)} = I\left(P_t < P_{t-1}\right),$$

where $I(\omega) = 1$ if the event ω occurs and $I(\omega) = 0$ otherwise. Then we rearrange (2.1) and (2.2). By combining (2.3) with (2.1) and (2.2), we can get the representation of ΔP_t as

$$\Delta P_t = g(t)[\frac{1}{g(t-1)} - \beta_0]\Delta P_{t-1} - g(t)\sum_{i=1}^{p-1}\beta_i \Delta P_{t-i} + g(t)[\mu + \sigma e_t] \ , \tag{2.4}$$

where $g(t) = g_1 I_t^{(1)} + g_2 I_t^{(2)}$ and Δ is the difference operator such as $\Delta P_t = P_t - P_{t-1}$. In this representation, $I_t^{(1)} = 1$ if and only if $V_t - V_t^* \geq 0$. But then (2.1) implies that $I_t^{(1)} = 1$ if and only if $\Delta P_t \geq 0$. When $p = 1$ and $\beta_0 = 1$, (2.4) is identical to our SSARI model in Kunitomo and Sato (1997).

2.2 The SSARIMA(p, q)–ARCH(r) model

In this section we shall generalize (2.4) in the above section and introduce a new class of the simultaneous switching autoregressive integrated moving-average (SSARIMA) model with autoregressive conditional heteroskedasticity (ARCH). For the specific application to financial time series in this paper we first consider the univariate SSARMA model given by

$$y_t = \begin{cases} a_1^* + a_2^* t + \sum_{i=1}^{p} a_i y_{t-i} + \sigma_1 u_t & (\text{if} \quad y_t \geq y_{t-1}) \\ \\ b_1^* + b_2^* t + \sum_{i=1}^{p} b_i y_{t-i} + \sigma_2 u_t & (\text{if} \quad y_t < y_{t-1}) \end{cases} , \tag{2.5}$$

where we take $\sigma_i > 0$ $(i = 1, 2)$. The disturbance terms $\{u_t\}$ are a sequence of $I(1)$ process satisfying

$$\Delta u_t = \sum_{j=0}^{q} c_j v_{t-j}, \tag{2.6}$$

where $c_0 = 1$ for the normalization factor. The random variables $\{v_t\}$ are martingale differences with $E(v_t|\mathcal{F}_{t-1}) = 0$ and

$$E(v_t^2|\mathcal{F}_{t-1}) = 1 + \sum_{i=1}^{r} \alpha_i v_{t-i}^2 \;\; a.s. \,, \tag{2.7}$$

where the unknown coefficients $\{\alpha_i, i = 1, \cdots, r\}$ satisfy some restrictions on the positivity of conditional variances $\alpha_i \geq 0$ $(i = 1, \cdots, r)$. In the above notation the $\sigma-$field $\mathcal{F}_{t-1}$ is generated by a set of random variables $\{y_s, v_s \,; s \leq t-1\}$. Although the conditional heteroskedasticity of $\{v_t\}$ has a simple form as in the original ARCH model and is symmetrical, the resulting conditional heteroskedasticities for $\{y_t\}$ and $\{\Delta y_t\}$ can be asymmetrical.

The univariate non-linear time series model we introduced in (2.5) has the first order multivariate autoregressive form with moving-average disturbances by using the standard state space representation in time series analysis. Let us define $p \times 1$ vectors $\boldsymbol{y}_t$ and $\boldsymbol{u}_t$ by

$$\boldsymbol{y}_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix} , \boldsymbol{u}_t = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \tag{2.8}$$

$p \times 2$ matrices $\boldsymbol{\mu}_i$ $(i = 1, 2)$ by

$$\boldsymbol{\mu}_1 = \begin{pmatrix} a_1^* & a_2^* \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} , \quad \boldsymbol{\mu}_2 = \begin{pmatrix} b_1^* & b_2^* \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} , \tag{2.9}$$

and $p \times p$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ by

$$\boldsymbol{A} = \begin{pmatrix} a_1 & \cdots & \cdots & a_p \\ 1 & & \boldsymbol{0} & \\ & \ddots & & \\ \boldsymbol{0} & & 1 & 0 \end{pmatrix} , \quad \boldsymbol{B} = \begin{pmatrix} b_1 & \cdots & \cdots & b_p \\ 1 & & \boldsymbol{0} & \\ & \ddots & & \\ \boldsymbol{0} & & 1 & 0 \end{pmatrix} . \tag{2.10}$$

Then the non-linear time series model we have introduced can be represented as

$$\boldsymbol{y}_t = \begin{cases} \boldsymbol{\mu}_1 \boldsymbol{z}_t^* + \boldsymbol{A}\boldsymbol{y}_{t-1} + \sigma_1 \boldsymbol{u}_t & \text{if} \quad \boldsymbol{e}_1' \boldsymbol{y}_t \geq \boldsymbol{e}_1' \boldsymbol{y}_{t-1} \\ \boldsymbol{\mu}_2 \boldsymbol{z}_t^* + \boldsymbol{B}\boldsymbol{y}_{t-1} + \sigma_2 \boldsymbol{u}_t & \text{if} \quad \boldsymbol{e}_1' \boldsymbol{y}_t < \boldsymbol{e}_1' \boldsymbol{y}_{t-1} \end{cases} , \tag{2.11}$$

where $\boldsymbol{e}_1' = (1, 0, \cdots, 0)$ and $\boldsymbol{z}_t^* = (1, t)'$ is the vector of strictly exogenous variables.

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 model could be classified as the threshold 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 this type of time series models to a class of simultaneous switching autoregressive (SSAR) time series model. The univariate time series model consisting of (2.5) and (2.6) can be called the simultaneous switching autoregressive integrated moving-average (SSARIMA) model. We denote this class of models as SSARIMA$(p, 1, q)$ or simply SSARIMA(p, q), which is a direct extension of the standard ARIMA$(p, 1, q)$ model in the statistical time series analysis. In the SSARIMA model we need some restrictions on unknown parameters in order to make the stochastic process defined by (2.11) meaningful in a proper statistical sense. This issue has been called the coherency problem and extensively discussed in Kunitomo and Sato (1996, 1997) [1]. We say the non-linear time series model (2.11) is coherent if and only if the correspondence between $\{\boldsymbol{y}_t\}$ and $\{\boldsymbol{u}_t\}$ is one-to-one given the initial condition $\mathcal{F}_0$.

The conditions that $\boldsymbol{e}_1'\boldsymbol{y}_t \geq \boldsymbol{e}_1'\boldsymbol{y}_{t-1}$ and $\boldsymbol{e}_1'\boldsymbol{y}_t < \boldsymbol{e}_1'\boldsymbol{y}_{t-1}$ can be rewritten as

$$\sigma_1 u_t \geq \boldsymbol{e}_1'(\boldsymbol{I}_p - \boldsymbol{A})\boldsymbol{y}_{t-1} - \boldsymbol{e}_1'\boldsymbol{\mu}_1 \tag{2.12}$$

and

$$\sigma_2 u_t < \boldsymbol{e}_1'(\boldsymbol{I}_p - \boldsymbol{B})\boldsymbol{y}_{t-1} - \boldsymbol{e}_1'\boldsymbol{\mu}_2 \ , \tag{2.13}$$

respectively. Then the set of conditions on coherency in the present case can be summarized by a 1×2 vector $\boldsymbol{\gamma}^{*'} = (\gamma_1^*, \gamma_2^*)$ and a $1 \times p$ vector $\boldsymbol{\gamma}' = (\gamma_1, \cdots, \gamma_p)$:

$$\begin{aligned} \frac{1}{\sigma_1}[-\boldsymbol{e}_1'\boldsymbol{\mu}_1,\ \boldsymbol{e}_1'(\boldsymbol{I}_p - \boldsymbol{A})] &= \frac{1}{\sigma_2}[-\boldsymbol{e}_1'\boldsymbol{\mu}_2,\ \boldsymbol{e}_1'(\boldsymbol{I}_p - \boldsymbol{B})] \\ &= (\boldsymbol{\gamma}^{*'}, \boldsymbol{\gamma}') \ . \end{aligned} \tag{2.14}$$

Because the univariate SSARIMA(p, q) has some specific structure in the class of the general multivariate SSARIMA model, it has a simple representation. For instance, we do not need any additional conditions to (2.14) for the normalization. By using Theorem 2.1 of Kunitomo and Sato (1996), we have the one-to-one correspondence between the stochastic processes $\{y_t\}$ and $\{v_t\}$ under the condition given by (2.14).

In order to obtain a useful representation of the process $\{\Delta y_t\}$, we use the indicator functions

$$I_t^{(1)} = I(\boldsymbol{e}_1'\boldsymbol{y}_t \geq \boldsymbol{e}_1'\boldsymbol{y}_{t-1})$$

and

$$I_t^{(2)} = I(\boldsymbol{e}_1'\boldsymbol{y}_t < \boldsymbol{e}_1'\boldsymbol{y}_{t-1}) \ .$$

Then we can obtain the representation

[1] The terminology "coherency" has been adopted from Gourieroux, Laffont, and Monfort (1980) in econometrics.

$$\boldsymbol{y}_t = \boldsymbol{\mu}(t)\boldsymbol{z}_t^* + \boldsymbol{A}(t)\boldsymbol{y}_{t-1} + \sigma(t)\boldsymbol{u}_t \ , \tag{2.15}$$

where

$$\boldsymbol{\mu}(t) = \sum_{i=1}^{2} \boldsymbol{\mu}_i I_t^{(i)} \ , \tag{2.16}$$

$$\boldsymbol{A}(t) = \boldsymbol{A} I_t^{(1)} + \boldsymbol{B} I_t^{(2)} \ , \tag{2.17}$$

and

$$\sigma(t) = \sum_{i=1}^{2} \sigma_i I_t^{(i)} \ . \tag{2.18}$$

When the first component of $\{\boldsymbol{u}_t\}$ in (2.15) is an $I(1)$ process, the stochastic process $\{y_t\}$ is a non-ergodic process. Hence it is of interest to investigate the ergodicity conditions for $\{\Delta y_t\}$. In the present univariate case we can simplify some coefficients by the coherency conditions (2.14), that is, we have the relations

$$\boldsymbol{\mu}(t) = -(\gamma_1^*, \gamma_2^*)\sigma(t) \tag{2.19}$$

and

$$\boldsymbol{e}_1' - \boldsymbol{e}_1'\boldsymbol{A}(t) = (\gamma_1, \cdots, \gamma_p)\sigma(t) \ . \tag{2.20}$$

By using these relations, we re-write the disturbance terms $\{u_t\}$ as

$$u_t = \frac{1}{\sigma(t)}\Delta y_t + \gamma_1^* + \gamma_2^* t + \sum_{i=1}^{p} \gamma_i y_{t-i} \ . \tag{2.21}$$

Because of (2.6) and (2.21), given the information available at $t-1$, we have to consider four phases for Δy_t at t depending on $I_t^{(i)}$ and $I_{t-1}^{(i)}$ $(i = 1, 2)$. By taking the difference operation on (2.21) and using (2.14) and (2.15), we have the representation as

$$\Delta y_t = \sigma(t)\left\{A_0 + \sum_{i=1}^{p} A_i(t-i)\sigma(t-i)^{-1}\Delta y_{t-i} + \Delta u_t\right\} \ , \tag{2.22}$$

where $A_0 = -\gamma_2^*$,

$$A_1(t-1) = 1 - \gamma_1\sigma(t-1) \ ,$$

and

$$A_i(t-i) = -\gamma_i\sigma(t-i) \ \ (i = 2, \cdots, p) \ .$$

It is immediately clear that the time series model defined by (2.4) is a special case of (2.22). Moreover, we have the following characterization result on $\{\Delta y_t\}$, which is a simple extension of Theorem 2.1 of Kunitomo and Sato (1997).

Theorem 2.1 : *Define the non-linear transformation of* $\{\Delta y_t\}$ *by*

$$T(\Delta y_t) = \sigma(t)^{-1}\Delta y_t \ . \tag{2.23}$$

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

$$T(\Delta y_t) = A_0 + \sum_{i=1}^{p} A_i(t-i)T(\Delta y_{t-1}) + \Delta u_t \ , \tag{2.24}$$

where A_0 and $A_i(t-i)$ $(i = 1, \cdots, p)$ are defined by (2.22).

The time series model defined by (2.24) can be called the threshold autoregressive moving-average (TARMA) model with time-varying coefficients in the non-linear time series analysis. From this result it can be deduced that the stochastic process for $\{\Delta y_t\}$ is slightly different from the TARMA(p,q) model. Also the stochastic process $\{\Delta y_t\}$ has the representation

$$\Delta y_t = \begin{cases} a_2^* + \sum_{i=1}^{p} a_i \Delta y_{t-i} + \sigma_1 \Delta u_t & (\text{if} \quad \Delta y_{t-1} \geq 0, \ \Delta y_t \geq 0) \\ a_2^* + \left(\frac{\sigma_1}{\sigma_2}\right) b_1 \Delta y_{t-1} + \sum_{i=2}^{p} a_i \Delta y_{t-i} + \sigma_1 \Delta u_t & (\text{if} \quad \Delta y_{t-1} < 0, \ \Delta y_t \geq 0) \\ b_2^* + \left(\frac{\sigma_2}{\sigma_1}\right) a_1 \Delta y_{t-1} + \sum_{i=2}^{p} b_i \Delta y_{t-i} + \sigma_2 \Delta u_t & (\text{if} \quad \Delta y_{t-1} \geq 0, \ \Delta y_t < 0) \\ b_2^* + \sum_{i=1}^{p} b_i \Delta y_{t-i} + \sigma_2 \Delta u_t & (\text{if} \quad \Delta y_{t-1} < 0, \ \Delta y_t < 0) \end{cases} . \tag{2.25}$$

By this form of representation, we notice that the differenced process $\{\Delta y_t\}$ from the SSARIMA model has not only the simultaneous switching characteristic, but also a characteristic of the threshold type time series model. For the stochastic process $\{\Delta y_t\}$ defined by (2.25), we can present a set of sufficient conditions for its ergodicity. A proof is provided in the Appendix.

Theorem 2.2 : *Suppose (i) p is a finite number and $q = 0$, (ii) the coherency condition (2.14) holds, (iii) the density function $g(v)$ of $\{v_t\}$ is continuous and everywhere positive with respect to the Lebesgue measure, and (iv) $\sup_{t \geq 1} E[|v_t|] < +\infty$. Then the Markov chain defined by (2.25) for $\{\Delta y_t\}$ is ergodic if we have the conditions (v)*

$$a_1 + \sum_{j=2}^{p} |a_j| < 1 \ , \tag{2.26}$$

$$b_1 + \sum_{j=2}^{p} |b_j| < 1 \ , \tag{2.27}$$

$$[a_1 - \sum_{j=2}^{p} |a_j|][b_1 - \sum_{j=2}^{p} |b_j|] < 1 \ , \tag{2.28}$$

$$\sum_{j=2}^{p} |a_j| < 1 \,, \tag{2.29}$$

and

$$\sum_{j=2}^{p} |b_j| < 1 \,. \tag{2.30}$$

For the precise definition and the related discussions on the ergodicity for Markov chains on a general state space, see Tweedie (1973), Liu and Susko (1992), or Meyn and Tweedie (1993). When $p = 1$, we have shown (Kunitomo and Sato (1997)) that the necessary and sufficient conditions for the ergodicity of $\{\Delta y_t\}$ are given by

$$a_1 < 1 \,, b_1 < 1 \,, a_1 b_1 < 1 \,. \tag{2.31}$$

By a simple calculation, the sufficient conditions in the above theorem reduce to these ones when $p = 1$. When $p \geq 2$ and $q = 0$, the naive sufficient conditions for the stability of the stochastic process $\{\Delta y_t\}$ are given by

$$\sum_{i=1}^{p} |a_i| < 1, \ \sum_{i=1}^{p} |b_i| < 1 \,. \tag{2.32}$$

But they exclude many interesting non-linear phenomena such as the case when there are some over-reactions and subsequent gradual adjustments to its mean level. It seems that the conditions we have for the ergodicity in the avove theorem cover some of these important cases, but they are too strong as the necessary conditions when $p \geq 2$. From our limited number of simulations, we have found that our sufficient conditions exclude some important cases even when $p = 2$.

When q (> 0) is a finite number, we were not be able to prove the corresponding results as Theorem 2.2 (and Theorem 2.5 below) at present. However, we can prove some results even in this general case. The first result on the SSARIMA(p, q) model characterizes one of the main differences between the SSARIMA models in this paper and the threshold ARIMA (TARIMA) model in the non-linear time series analysis.

Theorem 2.3 : *Let $\{v_t\}$ in the SSARIMA(p, q) model be independently and identically distributed random variables with the density function $g(v)$, which is everywhere positive in $\mathbf{R}$. Then given*

$$(\Delta y_{t-1}, \cdots, \Delta y_{t-p}, v_{t-1}, \cdots, v_{t-q}) = (z_1, \cdots, z_p, z_{p+1}, \cdots, z_{p+q}),$$

the conditional probability

$$\Pr\{\Delta y_t \leq y | \Delta y_{t-1} = z_1, \cdots, \Delta y_{t-p} = z_p, v_{t-1} = z_{p+1}, \cdots, v_{t-q} = z_{p+q}\} \tag{2.33}$$

is a continuous function of $\boldsymbol{z}' = (z_1, \cdots, z_{p+q})$.

Contrary to the SSARIMA models discussed in this paper, the TARMA and TARIMA models do not share this continuity property. (See Tong (1990), for instance.)

We also have obtained a result on the existence of stationary distribution and moments of $\{\Delta y_t\}$ for the SSARIMA(p,q) models. Kunitomo and Sato (1997) have proved the same result when $p=1$ as their Lemma A.2. Since it is straightforward to extend the proof for an arbitrary p, we omit the details.

Theorem 2.4 : *In the SSARIMA(p,q) model given by (2.25), assume (i) the coherency conditions (2.14), (ii) the ergodicity conditions (2.26)-(2.30), and (iii)* $\sup_{t\geq 1} E[|v_t|^k] < +\infty$ *for some* $k \geq 1$. *Then there exists a stationary solution* $\{\Delta y_t\}$ *satisfying (2.25) and*

$$\sup_{t\geq 1} E[|\Delta y_t|^k] < +\infty \,. \tag{2.34}$$

2.3 Estimation and Model Selection

In order to estimate the SSARIMA models, we have proposed and investigated the maximum likelihood (ML) method (Kunitomo and Sato (1997)). We are also proposing to use the ML method for estimating the SSARIMA models with ARCH disturbances. However, because of the ARCH effects for the disturbance terms, we need to modify the likelihood function for the standard SSARIMA model slightly.

We set the initial conditions such that $v_0 = v_{-1} = \cdots = v_{-\max\{q,r\}} = 0$ and $\Delta y_t, 1 \leq t \leq p$ are fixed for the simplicity. Then there is an important aspect in the present model that the Jacobian of the transformation from $\{\Delta u_t, p+1 \leq t \leq T\}$ to $\{\Delta y_t, p+1 \leq t \leq T\}$ is given by

$$\prod_{t=p+1}^{T} \sigma(t)^{-1}.$$

The Jacobian of the transformation from $\{v_t, p+1 \leq t \leq T\}$ to $\{\Delta u_t, p+1 \leq t \leq T\}$ is 1 provided that (2.6) is an invertible MA process.

Under the assumption that the disturbance terms $\{v_t\}$ given $\mathcal{F}_{t-1}$ are conditionally normal random variables, the normalized log-likelihood function for $\{\Delta y_t,\ p+1 \leq t \leq T\}$ given the initial conditions can be written as

$$\begin{aligned} \log L_T(\boldsymbol{\theta}) &= -\frac{(T-1)m}{2T}\log 2\pi \\ &- \frac{1}{2T}\sum_{t=p+1}^{T}\sum_{i=1}^{2} I_t^{(i)} \log[\sigma_i^2 \omega_t(\boldsymbol{\theta})] - \frac{1}{2T}\sum_{t=p+1}^{T} \omega_t(\boldsymbol{\theta})^{-1} v_t(\boldsymbol{\theta})^2 \,, \end{aligned} \tag{2.35}$$

where $\{v_t(\boldsymbol{\theta})\}$ are $\{v_t\}$ rewritten from (2.5) and (2.6) as functions of $\{\Delta y_t\}$, $\boldsymbol{\theta}$ is a vector of structural parameters appearing in the original SSARIMA model, and

$$(2.36) \qquad \omega_t(\boldsymbol{\theta}) = 1 + \sum_{i=1}^{r} \alpha_i v_{t-i}^2(\boldsymbol{\theta}) \ .$$

The unknown coefficients $\{\alpha_i, \ i = 1, \cdots, r\}$ satisfy

$$(2.37) \qquad \alpha_i \geq 0 \ , \ \sum_{i=1}^{r} \alpha_i < 1 \ .$$

In this representation of the SSARIMA model with ARCH disturbances, the vector of unknown coefficients consists of

$$\boldsymbol{\theta}' = (\gamma_2^*, \gamma_1, \cdots, \gamma_p, c_1, \cdots, c_q, \sigma_1, \sigma_2, \alpha_1, \cdots, \alpha_r) \ .$$

We note the important property that $\{v_t(\boldsymbol{\theta})\}$ do not include the parameters $\{\alpha_i\}$ because they are constructed from (2.5) and (2.6). The maximum likelihood (ML) estimator can be defined as 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 (2.14). For an efficient ML computation, we can divide the full samples into samples corresponding to four different phases using (2.25) because $\{\Delta u_t\}$ is a function of $\{\Delta y_{t-i}, i = 0, \cdots, p\}$ and $\{\sigma(t-i), i = 0, 1\}$ from (2.21). Then we try to maximize the likelihood function with respect to the unknown parameters by an iterative optimization technique.

The asymptotic properties of the ML estimator in the SSARIMA(1,0) has been established by Kunitomo and Sato (1997). By modifying those results slightly, we can establish that the ML estimator is consistent and asymptotically normal in the general SSARIMA(p,0)-ARCH(r) model. A sketchy proof of this result is given in the Appendix.

Theorem 2.5 : *Let $\{\Delta y_t\}$ be the stationary process followed by the SSARIMA(p,0) model (2.25) with ARCH(r) disturbances. Assume (i) the sufficient conditions for the coherency in (2.14), (ii) the sufficient conditions for the ergodicity in (2.26)-(2.30), (iii) the disturbance terms $\{v_t\}$ are stationary and conditionally normally distributed given $\mathcal{F}_{t-1}$ with $E[v_t^{4+\delta}] < +\infty$ for some $\delta > 0$, and (iv) $\alpha_i \geq 0$ $(i = 1, \cdots, r), 0 \leq \sum_{i=1}^{r} \alpha_i < 1$, and $\sigma_i > 0$ $(i = 1, 2)$. Also suppose (v) the true parameter vector $\boldsymbol{\theta}_0$ is an interior point of the parameter space $\boldsymbol{\Theta}$. Then the ML estimators $\hat{\boldsymbol{\theta}}_{ML}$ of the unknown parameters in $\boldsymbol{\theta}$ are consistent and asymptotically normally distributed as*

$$(2.38) \qquad \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left[0, \ I(\boldsymbol{\theta}_0)^{-1}\right] \ ,$$

where $\boldsymbol{\theta}_0$ is the true parameters vector and

$$(2.39) \qquad I(\boldsymbol{\theta}_0) = \lim_{T\to\infty} \left[-\frac{\partial^2 \log L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right] \ .$$

The conditions in the above theorem are the same as in Theorem 2.4 of Kunitomo and Sato (1997) except (ii) and (iii). In the present case, we need some sufficient conditions for the non-negativity of conditional variance functions and the existence of unconditional moments, which imply some restrictions on the parameter space. At present we also have the asymptotic results only when $q = 0$. These conditions could be certainly relaxed. The asymptotic information matrix can be consistently estimated by

$$\left[-\frac{\partial^2 \log L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] |_{\hat{\boldsymbol{\theta}}_{ML}} . \tag{2.40}$$

Although we have obtained the consistency and the asymptotic normality of the ML estimator under a set of sufficient conditions, presently we do not know much of the behavior of the ML estimator when the sample size is not very large and the conditional distribution of disturbances are not normal. In this respect, we have investgated the finite sample properties of the ML estimator in a systematic way for the SSARIMA(1, 0) model when $T = 100$ and $T = 500$ (Kunitomo and Sato (1997)). We have produced some evidence on the use of the ML estimation method for the SSARIMA(1, 0) model even when the sample size is not very large and the conditional distribution of disturbances are different from normal to a certain extent. From those results we expect that similar finite sample properties could hold in the present situation.

It is important to note that the class of the SSARIMA(p, q)-ARCH(r) models we have introduced in this paper includes many statistical time series models as special cases. For instance, it includes the ARIMA$(p, 1, q)$ models with ARCH(r) disturbances. In order to select an appropriate model within the class of the SSARIMA(p, q)-ARCH(r) models, we propose to use the minimum Akaike's Information Criterion (AIC): see Akaike (1973). Although there have been many proposed criteria for model selection procedures, it is widely known that the minimum AIC is useful and successful in many applications. The value of AIC for the SSARIMA(p, q)-ARCH(r) model in the present case can be defined by

$$\begin{aligned} AIC(p,q,r) &= \frac{1}{T}\sum_{i=1}^{2} \sum_{T=p+1}^{T} I_t^{(i)} \log[\hat{\sigma}_i{}^2 \omega_t(\hat{\boldsymbol{\theta}})] \\ &+ \frac{1}{T} \sum_{t=p+1}^{T} \omega_t(\hat{\boldsymbol{\theta}})^{-1} v_t(\hat{\boldsymbol{\theta}})^2 + 2(p+q+r) , \end{aligned} \tag{2.41}$$

where $\hat{\sigma}_i$ $(i = 1, 2)$ denote the ML estimator of σ_i $(i = 1, 2)$ and $\hat{\boldsymbol{\theta}}$ denotes the vector of the ML estimator for $\boldsymbol{\theta}$.

3. An Application to Financial Data

In this section we shall report an empirical result using some time series data on stock price indices. In our data analysis we have used the time series data of the

Nikkei-225 indices which are the most popular stock price index traded in Japan and the SP-500 from the U.S. The first data sets are the closing daily data of the Nekkei-225 Spot index from January 1985 to December 1994 and the second data sets are the closing daily data of the SP-500 from January 1985 to December 1994. All data were transformed into their logarithms before the estimation of the non-stationary SSARMA-ARCH model. It may be of some interests for economists to compare the time series movements of these spot prices, which are reprentative indices in two major stock markets in the world.

It has often been argued by financial economists and practitioners in major financial markets that there have been significant daily effects in the volatility functions of financial prices. The typical argument for the existence of daily effects is from the observation that the information flows are different from day to day and there are some trading day effects and holiday effects. By taking account of this observation, in actual estimation we have the volatility function

$$\omega_t(\boldsymbol{\theta}) = 1 + \sum_{i=1}^{r} \alpha_i v_{t-i}^2 + \sum_{i=1}^{r^*} \beta_i D_{it} \ , \tag{3.1}$$

where D_{it} are the dummy variables defined by

$$D_{it} = \begin{cases} 1 & \text{if} \quad t \in \boldsymbol{J}_i \\ 0 & \text{if} \quad t \notin \boldsymbol{J}_i \end{cases} \ , \tag{3.2}$$

and the index sets $\boldsymbol{J}_i$ $(i = 1, \cdots, r^*)$ denote the calendar days [2].

We have estimated the SSARIMA$(p, 0)-$ARCH(r) model with the dummy variables of the form (3.1). The estimation of structural parameters in the SSARIMA$(p, 0)-$ARCH(r) model has been conducted by the ML method under the assumption of the conditionally normal disturbances. Since we cannot obtain an explicit formula for the ML estimators of unknown parameters, we have used a numerical non-linear optimization technique with the coherency restrictions on parameters. Because we do not have a rigorous justification for using the SSARIMA(p, q) models when $q > 0$ and we could not find any significant autocorrelations in residuals in most cases, we took $q = 0$ in actual estimation. We have divided the full sample period for Nikkei-225 into 4 sub-periods, each consisting of about 600 data points. This has been done because many different phenomena occurred in the full sample period and the assumption of constant coefficients in the underlying SSARIMA model could be unreasonable from a practical point of view [3]. The resulting estimation results are summarized by Table 1 for the Nikkei-225 data and Table 2 for the SP-500 data [4].

[2]We took $r^* = 5$ in our actual investigation and $\boldsymbol{J}_2$ corresponds to the Tuesday dummy, for instance. We could have incorporated other dummies such as holoday variables in a similar fashion.

[3]For instance, many changes in regulations, taxes, and market structures as well as financial innovations have occurred at financial markets in Japan during the full sample period.

[4]We have estimated both SSARIMA$(p, 0)-$ARCH(r) models with and without daily effects. In Table 1 and Table 2, SSARIMA(i) and ARIMA(i) $(i = 0, \cdots, 3)$ stand for SSARIMA$(i, 0)$

There are several interesting empirical observations from Table 1. First, the spot stock price index sometimes shows sharp asymmetrical movements either in the upward or downward phase. This phenomenon was 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 downward coefficients are often smaller than the corresponding upward coefficients. This agrees with the fact that the estimated volatility in the downward phase is often larger than the volatility in the upward phase. Second, we have found significant daily effects on the volatility functions in most cases. In our estimation of daily effects we have normalized the volatility coefficients such that the Monday volatility always takes the value of 1. Except for the first data period [5], we have found negative daily effects from the Monday volatility, which are statistically significant. This agrees with the notion that the Monday volatility effect is significantly larger than other days because of holiday effects, but also it is less than the holiday length from Friday to Monday times the normal level of volatility. In Japanese financial markets there was some trading on Saturday mornings during 1980s. It would not be possible to find any significant differences in the daily volatility differences because the trading on Saturday was terminated at the end of the 1980s. Also we found that the inclusion of daily effects in the form of (3.1) does not have a significant effect on the estimation of asymmetry and volatility level. In most cases the estimated coefficients of a_i, b_i, and α_i are similar for the model (2.25) with or without daily effects. Third, after around 1990 [6], there have not been many occasions as was the case previously when the asymmetrical movements of the price indices are evident and the estimated coefficients are not very large. The differences of AIC in ARIMA$(p,1,0)$ and SSARIMA$(p,0)$ are not very large in most cases in this period. Fourth, as far as the SP-500 data is concerned, we find some evidence of the asymmetry discussed during the periods before around 1990. This is quite evident for the data including that for 1987. In addition, we often find some asymmetry whenever we use a large quantity of data over more than 5 years before 1990. However, in the data of 1990s it is difficult to find any asymmetry and the estimated coefficients are generally not very large. We only show the estimation results by using the data from January 1990 to June 1992 in Table 2, which seem to be typical of recent SP-500 data. The model chosen by AIC is ARIMA$(0,1,0)$ with ARCH(2) and daily effects. It is often possible to find significant ARCH effects as well as daily effects on volatility in the 1990s by comparing the values of AIC.

---

and ARIMA$(i,1,0)$, respectively.

[5] The estimated Monday and Tuesday volatilities in this period are different from our intuition. However, Dr. Kunio Okina has suggested as possible explanation that there was a substantial amount of informational flow from the U.S. financial markets to the Tokyo stock market during the first period. Because there is a time lag in two countries, the Tuesday volatility in Tokyo was quite large.

[6] There can be economic as well as institutional reasons for this phenomenon. For instance, we have compared the movements of the Nikkei-225 spot index and futures index (Kunitomo and Sato (1997)). The trading of the Nikkei-225 futures in Japan became active around at the end of 1989.

These empirical problems and findings may have some implications for financial economists. Needless to say, these observations from our empirical results on stock indices are preliminary. 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 sometimes ignored in the past econometric studies on the volatilities of major financial time series. Since the asymmetrical pattern in the movements of time series between the upward phase and the downward phase often observed by economists can not be represented properly by the standard ARMA, ARIMA, and ARCH processes, we have proposed to use the class of simultaneous switching autoregressive integrated moving-average (SSARIMA) models with autoregressive conditional heteroskedastic (ARCH) disturbances, which was recently proposed by Kunitomo and Sato (1997). In this paper we have investigated some properties of the SSARIMA$(p, 0)$-ARCH(r) model and the asymptotic properties of the maximum likelihood estimation method for estimating its unknown parameters. Unlike other methods already available for estimating the volatilities of financial time series, our method is a very simple but general one for investigating whether they are asymmetrical or not from a medium size set of time series data. The model selection procedure in the class of the SSARIMA(p, q)-ARCH(r) can be straightforwardly developed.

We have also tried to show that there are some reasons why the SSARIMA model with ARCH disturbances is a useful tool to analyze many financial time series in financial markets. The point is that if we permit the intrisic value of security to be different from the observed price and have a non-linear adjustment process, the result is the SSARIMA model. Then the estimated coefficients in the upward phase and in the downward phase can be different and the resulting volatility of financial time series can be asymmetrical around the normal volatility level.

In this paper we have used the SSARIMA-ARCH modelling to examine the movements of the Nikkei-225 stock index at Tokyo and the SP-500 index at New York from 1985 to 1994. We have found some evidence of the asymmetrical movements of stock indices before around 1990 in particular, which has been consistent with the view of some financial economists and practitioners in the Japanese financial markets. By using our empirical example, at least we have shown that our modelling approach is useful for analysing some financial prices.

Finally, we should mention that our approach can be easily extended to a more general form of conditional heteroskedasticity models such as the generalized ARCH model by Bollerslev (1986). In addition, our formulation does not need a large quantity of data for estimation as some non-parametric and semi-parametric time series methods do. We think there are some advantageous aspects of our modelling compared over other methods.

5. Mathematical Appendix

In this appendix, we gather some mathematical details which we have omitted in the previous sections. However, since most of the theoretical results reported in this paper are rather straightforward extensions of the corresponding ones in Kunitomo and Sato (1997), we use some notations and related results in that paper. We use the notation $x_t = \Delta y_t$ in the Appendix.

Proof of Theorem 2.2 : Define a $p \times 1$ vector $\boldsymbol{X}_t$ by

$$\text{(A.1)} \qquad \boldsymbol{X}_t = \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-p+1} \end{pmatrix} .$$

We then consider the Markovian representation for $\{\boldsymbol{X}_t\}$. The condition $x_t \geq 0$ is equivalent to $v_t \geq \gamma_2^* + \boldsymbol{a}_{t-1}^{'} \boldsymbol{X}_{t-1}$, where

$$\text{(A.2)} \qquad \boldsymbol{a}_{t-1}^{'} = (\gamma_1 - \frac{1}{\sigma(t-1)}, \gamma_2, \cdots, \gamma_p) .$$

By using (2.21), we have the representation

$$\text{(A.3)} \qquad \boldsymbol{X}_t = \boldsymbol{H}(\boldsymbol{X}_{t-1}, v_t) ,$$

where

$$\text{(A.4)} \qquad \boldsymbol{H}(\boldsymbol{X}_{t-1}, v_t) = \begin{pmatrix} -\sigma(t)\gamma_2^* - \sigma(t)\boldsymbol{a}_{t-1}^{'}\boldsymbol{X}_{t-1} + \sigma(t)v_t \\ x_{t-1} \\ \vdots \\ x_{t-p+1} \\ v_t \end{pmatrix} .$$

We use the criterion function

$$\text{(A.5)} \qquad G(\boldsymbol{x}) = 1 + \max_{1 \leq j \leq p} |x_j| \rho_j ,$$

for $\boldsymbol{x} = (x_i)$ and some $\rho_1 > \rho_2 > \cdots > \rho_p > 0$. First, we consider the case when $x_{t-1} \geq 0$. For $t \geq 1$ and $\boldsymbol{x} = (x_j)$,

$$\begin{aligned} \text{(A.6)} \qquad & E[G(\boldsymbol{X}_t)|\boldsymbol{X}_{t-1} = \boldsymbol{x}] \\ = \ & 1 + E[\max\{\max_{1 \leq j \leq p} \rho_j |x_{t+1-j}||\boldsymbol{X}_{t-1} = \boldsymbol{x}] \\ \leq \ & c_1 + E[\max\{(a_1\rho_1 x_{t-1} + \sum_{j=2}^{p} |a_j|\rho_1|x_{t-j}|)I_t^{(1)} \\ & -(\frac{\sigma_2}{\sigma_1}a_1\rho_1 x_{t-1} + \sum_{j=2}^{p} |b_j|\rho_1|x_{t-j}|)I_t^{(2)}, \max_{2 \leq j \leq p} \rho_j |x_{t+1-j}|\}] , \end{aligned}$$

where c_1 is a positive constant. We have the last inequality since $E[|v_t|] < +\infty$. Because we assumed the conditions (v), there exist $0 < \theta < 1$ and $\rho_1 > \rho_2 > \cdots > \rho_p > 0$ such that $\theta > \rho_{j+1}/\rho_j$ $(j = 1, \cdots, p-1)$ and

$$a_1 + \sum_{j=2}^{p} |a_j| \frac{\rho_1}{\rho_j} < \theta < 1 , \tag{A.7}$$

$$b_1 + \sum_{j=2}^{p} |b_j| \frac{\rho_1}{\rho_j} < \theta < 1 , \tag{A.8}$$

$$\sum_{j=2}^{p} |a_j| \frac{\rho_1}{\rho_j} < \theta < 1 \tag{A.9}$$

and

$$\sum_{j=2}^{p} |b_j| \frac{\rho_1}{\rho_j} < \theta < 1 . \tag{A.10}$$

We note that the condition (2.28) in (v) implies an inequality

$$\gamma_1 + \sum_{j=2}^{p} |\gamma_j| < \frac{1}{\sigma_1} + \frac{1}{\sigma_2} . \tag{A.11}$$

Hence we can take the same θ in (A.7)-(A.10) satisfying $0 < \theta < 1$,

$$\sigma_1(\gamma_1 + \sum_{j=2}^{p} |\gamma_j| \frac{\rho_1}{\rho_j} - \frac{1}{\sigma_2}) < \theta < 1 , \tag{A.12}$$

and

$$\sigma_2(\gamma_1 + \sum_{j=2}^{p} |\gamma_j| \frac{\rho_1}{\rho_j} - \frac{1}{\sigma_1}) < \theta < 1 . \tag{A.13}$$

Then by using (2.14), (A.7), and (A.13), we have

$$\begin{aligned} & E[G(\boldsymbol{X}_t)|\boldsymbol{X}_{t-1} = \boldsymbol{x}] \\ \leq\ & c_2 + \max\{[(\max(a_1, 0)\rho_1 x_{t-1} + \sum_{j=2}^{p} |a_j| \frac{\rho_1}{\rho_j} \rho_j |x_{t-j}|) I_t^{(1)} \\ & +(\max(-\frac{\sigma_2}{\sigma_1} a_1, 0)\rho_1 x_{t-1} + \sum_{j=2}^{p} |b_j| \frac{\rho_1}{\rho_j} \rho_j |x_{t-j}|) I_t^{(2)}], \theta \max_{2 \leq j \leq p} \rho_j |x_j|\} \\ <\ & c_3 + \theta[1 + \max_{1 \leq j \leq p} \rho_j |x_j|] \\ =\ & c_3 + \theta G(\boldsymbol{x}) , \end{aligned} \tag{A.14}$$

where c_2 and c_3 are some positive constants.

Next, we consider the case when $x_{t-1} < 0$. By the similar arguments based on the conditions in (v), we have a similar inequality for $t \geq 1$ as (A.14). Then for any $\boldsymbol{X}_{t-1} = \boldsymbol{x}$, we have

$$(A.15) \qquad E[G(\boldsymbol{X}_t)|\boldsymbol{X}_{t-1} = \boldsymbol{x}] < c_4 + \theta G(\boldsymbol{x}) \, ,$$

where c_4 is a positive constant.

The rest of our proof is similar to the one for Lemma 3.1 of Chan and Tong (1985). Since the Markov chain for $\boldsymbol{X}_t$ is aperiodic and ϕ−irreducible due to Condition (iii), we can apply Theorem 4 in Tweedie (1983) and establish that $\{\boldsymbol{X}_t\}$ is geometrically ergodic. *Q.E.D.*

A Sketch of Proof of Theorem 2.5 : The method of our proof is similar to the one used in Kunitomo and Sato (1997). Some lengthy modifications are necessary because there is an ARCH effect in the disturbance terms and p is an arbitrary integer. However, basically we can develop the method similar to the one used by Weiss (1986), which in turn utilized the results of Basawa, Feigin, and Heyde (1976). We give only a sketch of the proof.

(i) ***Step*** 1 : Let the stochastic process $\{\Delta u_t(\boldsymbol{\theta})\}$ be defined by

$$(A.16) \; \Delta u_t(\boldsymbol{\theta}) = D(t)^{-1}\Delta y_t + \gamma_2^* + [\gamma_1 - D(t-1)^{-1}]\Delta y_{t-1} + \sum_{i=2}^{p} \gamma_i \Delta y_{t-i} \, ; ,$$

which is identical to $\{\Delta u_t\}$ in (2.25). We denote the vector of true parameter values of $\boldsymbol{\theta}' = (\gamma_2^*, \gamma_1, \cdots, \gamma_p, \sigma_1, \sigma_2)$ as $\boldsymbol{\theta}_0' = (\gamma_2^{*(0)}, \gamma_1^{(0)}, \cdots, \gamma_p^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)})$. By substituting Δy_t evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ into (A.16), we have

$$\begin{aligned}
(A.17) \qquad \Delta u_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0) &= \sigma(t)^{-1}\sigma_t^{(0)}\Delta \boldsymbol{u}_t(\boldsymbol{\theta}_0) \\
&+ [\gamma_2^* - \sigma(t)^{-1}\sigma_t^{(0)}\gamma_2^{*(0)}] \\
&+ \sum_{i=1}^{p}[\gamma_i - \sigma(t)^{-1}\sigma_t^{(0)}\gamma_i^{(0)}]\Delta y_{t-i}(\boldsymbol{\theta}_0) \\
&+ [\sigma(t)^{-1}\sigma(t)^{(0)}\sigma_{t-1}^{(0)-1} - \sigma(t-1)^{-1}]\Delta y_{t-1}(\boldsymbol{\theta}_0),
\end{aligned}$$

where $\sigma_t^{(0)} = \sigma(t)$ evaluated at $\sigma_i = \sigma_i^{(0)}$ $(i = 1, 2)$ and $\Delta y_{t-1}(\boldsymbol{\theta}_0) = \Delta y_{t-1}$ evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Thus we have $\sigma(t)^{-1}\sigma_t^{(0)} = (\sigma_1^{(0)}/\sigma_1)I_t^{(1)} + (\sigma_2^{(0)}/\sigma_2)I_t^{(2)}$, for instance. By the Ergodic Theorem for the Markov chain (see Chapter 17 of Meyn and Tweedie (1993), for instance),

$$\begin{aligned}
(A.18) \qquad Q(\boldsymbol{\theta}, \boldsymbol{\theta}_0) &= -\frac{1}{2} \lim_{T\to\infty} \frac{1}{T} \sum_{t=2}^{T} \sum_{i=1}^{2} I_t^{(i)} \log(\sigma_i^2 \omega_t(\boldsymbol{\alpha})) \\
&\quad -\frac{1}{2} \lim_{T\to\infty} \frac{1}{T} \sum_{t=2}^{T} \frac{\Delta u_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0)^2}{\omega_t(\boldsymbol{\alpha})}
\end{aligned}$$

$$= -\frac{1}{2}E[\sum_{i=1}^{2} I_t^{(i)}\log(\sigma_i^2\omega_t(\boldsymbol{\alpha}))] - \frac{1}{2}E[\frac{\Delta u_t(\boldsymbol{\theta},\boldsymbol{\theta}_0)^2}{\omega_t(\boldsymbol{\alpha})}] ,$$

where $\boldsymbol{\alpha} = (\alpha_i)$. Because (A.17) is a linear function of $\gamma_2^*, \gamma_1, \cdots, \gamma_p$, and $1/\sigma_1$ ($= \eta_1$), $1/\sigma_2$ ($= \eta_2$), 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). Also by using the conditional normality assumption on $\{v_t\}$, we have the key relation

(A.19)
$$\int_{c_t}^{+\infty} (v^2 - \omega_t(\boldsymbol{\alpha}^{(0)}) - c_t v)\phi(v)dv = 0 ,$$

where c_t is a constant given $\mathcal{F}_{t-1}$ and $\phi(\cdot)$ is the density function of the standard normal distribution. Then we can obtain that for $\boldsymbol{\theta} = (\theta_i)$,

(A.20)
$$\frac{\partial Q}{\partial \theta_i}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = 0 .$$

Next, from (A.16) we rewrite

(A.21)
$$I_t^{(i)}\Delta y_t - I_{t-1}^{(i)}\Delta y_{t-1} = \eta_i^{-1} I_t^{(i)}[v_t + c_t] - I_{t-1}^{(i)}\Delta y_{t-1} \ (i = 1, 2) ,$$

where we take

$$c_t = -\gamma_2^{*(0)} + \sigma_{t-1}^{(0)-1}\Delta y_{t-1} - \sum_{i=1}^{p}\gamma_i \Delta y_{t-i} .$$

Then for $i = 1, 2$ we have the equation

(A.22)
$$\begin{aligned}\frac{\partial Q}{\partial \eta_i}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= E\{(\omega_t(\boldsymbol{\alpha})^{-1} v_t[\eta_i^{-1} I_t^{(i)}(v_t + c_t) - I_{t-1}^{(i)}\Delta y_{t-1}]\} \\ &= 0 .\end{aligned}$$

By using (A.19) with c_t in (A.21), we have $E[I_t^{(i)}(\omega_t(\boldsymbol{\alpha})^{-1}\omega_t(\boldsymbol{\alpha}^{(0)}) - 1)] = 0$ ($i = 1, 2$) . Then we can use the inequality

(A.23)
$$\begin{aligned} 0 &= \log E[\omega_t(\boldsymbol{\alpha})^{-1}\omega_t(\boldsymbol{\alpha}^{(0)})] \\ &\geq E[\log(\omega_t(\boldsymbol{\alpha})^{-1}\omega_t(\boldsymbol{\alpha}^{(0)}))] .\end{aligned}$$

We notice that the first term of (A.18) does not depend on γ_2^*, γ_j ($j = 1, \cdots, p$) and the maximum of the function $Q(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ with respect to these parameters occur at some linear combinations of η_i ($i = 1, 2$). Also the first order conditions for η_i ($i = 1, 2$) are given by

(A.24)
$$\eta_i^{-1} E[I_t^{(i)}] - E[\omega_t(\boldsymbol{\alpha})^{-1} v_t \frac{\partial \Delta u_t(\boldsymbol{\theta},\boldsymbol{\theta}_0)}{\partial \eta_i}] = 0 ,$$

which implies

$$E[\omega_t(\boldsymbol{\alpha})^{-1} v_t \sum_{i=1}^{2} \eta_i \frac{\partial \Delta u_t(\boldsymbol{\theta},\boldsymbol{\theta}_0)}{\partial \eta_i}] = \sum_{i=1}^{2} E[I_t^{(i)}] = 1 .$$

Hence the maximization of the function $Q(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ with respect to γ_2^*, γ_i $(i = 1,\cdots,p)$, and η_i $(i=1,2)$ implies the condition

$$E[\omega_t(\boldsymbol{\alpha})^{-1}\Delta u_t(\boldsymbol{\theta},\boldsymbol{\theta}_0)^2] = E[\omega_t(\boldsymbol{\alpha})^{-1}v_t(\boldsymbol{\beta}^{(0)})^2] = E[\omega_t(\boldsymbol{\alpha})^{-1}\omega_t(\boldsymbol{\alpha}^{(0)})] = 1 \ ,$$

where $\boldsymbol{\theta}_0' = (\boldsymbol{\beta}^{(0)'}, \boldsymbol{\alpha}^{(0)'})$. Then we have

$$\text{(A.25)} \qquad \begin{aligned} & -\frac{1}{2}E[\sum_{i=1}^{2} I_t^{(i)} \log(\sigma_i^2\omega_t(\boldsymbol{\alpha}))] - \frac{1}{2}E[\frac{\Delta u_t(\boldsymbol{\theta},\boldsymbol{\theta}_0)^2}{\omega_t(\boldsymbol{\alpha})}] \\ \leq\ & -\frac{1}{2}E[\sum_{i=1}^{2} I_t^{(i)} \log(\sigma_i^{(0)2}\omega_t(\boldsymbol{\alpha}^{(0)}))] - \frac{1}{2}E[\frac{\Delta u_t(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)^2}{\omega_t(\boldsymbol{\alpha}^{(0)})}] \end{aligned}$$

where $\boldsymbol{\alpha}^{(0)} = (\alpha_i^{(0)})$ and the equality holds only if when $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{(0)}$. Thus the function $Q(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ is uniquely maximized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

(ii) ***Step*** 2 : As the results of straightforward calculations, we have a representation of the information matrix evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ as

$$\text{(A.26)} \ \mathcal{I}(\boldsymbol{\theta}_0) = E[\omega_t(\boldsymbol{\alpha}^{(0)})^{-1} \begin{pmatrix} 1 \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \\ I_t^{(1)}\Delta y_t - I_{t-1}^{(1)}\Delta y_{t-1} \\ I_t^{(2)}\Delta y_t - I_{t-1}^{(2)}\Delta y_{t-1} \\ \omega_t(\boldsymbol{\alpha}^{(0)})^{-1}v_t\frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_1} \\ \vdots \\ \omega_t(\boldsymbol{\alpha}^{(0)})^{-1}v_t\frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_p} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \\ I_t^{(1)}\Delta y_t - I_{t-1}^{(1)}\Delta y_{t-1} \\ I_t^{(2)}\Delta y_t - I_{t-1}^{(2)}\Delta y_{t-1} \\ \omega_t(\boldsymbol{\alpha}^{(0)})^{-1}v_t\frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_1} \\ \vdots \\ \omega_t(\boldsymbol{\alpha}^{(0)})^{-1}v_t\frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_r} \end{pmatrix}']$$

$$+ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & E[I_t^{(1)}]\eta_1^{(0)-2} & 0 & \vdots \\ 0 & 0 & 0 & E[I_t^{(2)}]\eta_2^{(0)-2} & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2}I(\boldsymbol{\alpha}^{(0)},\boldsymbol{\alpha}^{(0)}) \end{pmatrix},$$

where

$$\text{(A.27)} \qquad I(\boldsymbol{\alpha}^{(0)},\boldsymbol{\alpha}^{(0)}) = E[\omega_t(\boldsymbol{\alpha}^{(0)})^{-2} \begin{pmatrix} \frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_1} \\ \vdots \\ \frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_r} \end{pmatrix} \begin{pmatrix} \frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_1} \\ \vdots \\ \frac{\partial\omega_t(\boldsymbol{\alpha})}{\partial\alpha_r} \end{pmatrix}']$$

is a positive definite matrix and $\eta_i^{(0)} = 1/\sigma_i^{(0)}$ $(i=1,2)$. (If (A.27) is not positive definite, then we can lead an obvious contradition. We can use this type of arguments in the following derivation.) We notice that $E[I_t^{(i)}] > 0$ $(i=1,2)$, the matrix

$$E[\begin{pmatrix} 1 \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \end{pmatrix}\begin{pmatrix} 1 \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \end{pmatrix}'] \tag{A.28}$$

is positive definite, and the upper-right hand corner and the lower-left hand corner of $I(\boldsymbol{\theta}_0)$ are zero sub-matrices. By taking the conditional expectation with respect to $\mathcal{F}_{t-1}$ at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and using (A.19), we also have the relation

$$\begin{aligned} &E\{[\omega_t(\boldsymbol{\alpha}^{(0)})v_t(\boldsymbol{\theta}_0)\frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_j}][I_t^{(i)}\Delta y_t - I_{t-1}^{(i)}\Delta y_{t-1}]\} \\ = \ &\eta_i^{(0)-1}E\{I_t^{(i)}\frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_j}\} \end{aligned} \tag{A.29}$$

for $i = 1, 2$ and $j = 1, \cdots, r$. Then we can utilize that

$$E[\omega_t(\boldsymbol{\alpha}^{(0)})^{-2}\begin{pmatrix} \omega_t(\boldsymbol{\alpha}^{(0)})\eta_1^{(0)-1}I_t^{(1)} \\ \omega_t(\boldsymbol{\alpha}^{(0)})\eta_2^{(0)-1}I_t^{(2)} \\ \frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_1} \\ \vdots \\ \frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_r} \end{pmatrix}\begin{pmatrix} \omega_t(\boldsymbol{\alpha}^{(0)})\eta_1^{(0)-1}I_t^{(1)} \\ \omega_t(\boldsymbol{\alpha}^{(0)})\eta_2^{(0)-1}I_t^{(2)} \\ \frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_1} \\ \vdots \\ \frac{\partial \omega_t(\boldsymbol{\alpha})}{\partial \alpha_r} \end{pmatrix}'] \tag{A.30}$$

is a positive definite matrix.
Next, we shall use the standard formula that for any partitioned $(m+n) \times (m+n)$ non-negative definite matrix

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix},$$

$$|\boldsymbol{A}| = |\boldsymbol{A}_{22}||\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}|,$$

where $\boldsymbol{A}_{22}$ is a positive definite matrix. Then by setting $\boldsymbol{A}_{22} = (1/2)I(\boldsymbol{\alpha}^{(0)}, \boldsymbol{\alpha}^{(0)})$ and after some calculations, the informatiom matrix defined by

$$I(\boldsymbol{\theta}_0) = -\frac{\partial^2 Q}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \tag{A.31}$$

is positive definite for $\boldsymbol{\theta} = (\theta_i)$.

(iii) ***Step*** 3 : The rest of the proof is to check Conditions (A2) given by Basawa, Feigin, and Heyde (1976). The arguments in our proof are standard because we are dealing with a stationary ergodic Markovian process. By using the Ergodic theorem for the Markov chain and the consistency of $\hat{\boldsymbol{\theta}}_{ML}$, we have the convergence of (2.39) and its consistent estimation. Also since we have assumed the conditional normality for disturbances and the existence of unconditional moments of order $4+\delta$, we can use Theorem 2.4 for $k=4$. Thus the expectations of third order derivatives of the log-likelihood function

$$|\frac{\partial^3 \log L_T(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j\partial\theta_k}|$$

are bounded for any i, j, and k for $\boldsymbol{\theta}=(\theta_i)\in\boldsymbol{\Theta}$. Then we can check the Condition (A2) given by Basawa, Feigin, and Heyde (1976), which implies that there exists a consistent root of the equation

$$\frac{\partial \log L_T(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=0\ .$$

(iv) ***Step*** 4 : In order to prove the asymptotic normality of the ML estimator in the SSIAR model with ARCH disturbances under the assumptions we have made, an important step is the martingale property of the partial derivatives of the log-likelihood function. Then we can use the central limit theorem for martingales. (See Basawa, Feigin, and Heyde (1976) or Hall and Heyde (1980), for instance.)

For the parameters in $\boldsymbol{\theta}$ except α_j $(j=1,\cdots,r)$, we can obtain the result by slightly modifying the derivation of Lemma A.5 of Kunitomo and Sato (1997). For the ARCH parameters α_j $(j=1,\cdots,r)$, we have

$$\text{(A.32)}\qquad \frac{\partial \log L_T(\boldsymbol{\theta})}{\partial\alpha_j}=\frac{1}{T}\sum_{t=1}^{T}(\frac{1}{2})\omega_t^{-1}\frac{\partial\omega_t}{\partial\alpha_j}[v_t^2(\boldsymbol{\theta})\omega_t^{-1}-1]\ ,$$

evaluated at $\boldsymbol{\theta}=\boldsymbol{\theta}_0$. Then it is straightforward to show that

$$\text{(A.33)}\qquad E\left[\frac{\partial \log L_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}|\mathcal{F}_{t-1}\right]=\frac{\partial \log L_{t-1}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\ ,$$

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

In order to use the central limit theorem for martingales, we can utilize the existence of unconditional moments $E[v_t^{4+\delta}]<\infty$ to show the conditional Lindeberg type condition. For instance, define a $1\times r$ vector $\boldsymbol{Z}_t=(z_{jt}; j=1,\cdots,r)$ by

$$\text{(A.34)}\qquad z_{jt}=\omega_t^{-1}v_{t-j}^2[v_t^2\omega_t^{-1}-1]\ .$$

Then we have

$$\text{(A.35)}\qquad \frac{1}{T}\sum_{t=1}^{T}E[||\boldsymbol{Z}_t||^2|\mathcal{F}_{t-1}]$$

converges to $E[||\boldsymbol{Z}_t||^2]$ in probability under the assumptions we have made. Also we have an inequality

$$(A.36)\quad \frac{1}{T}\sum_{t=1}^{T} E[||\boldsymbol{Z}_t||^2 I(||\boldsymbol{Z}_t|| > \sqrt{T}c)|\mathcal{F}_{t-1}] \leq \frac{1}{T}(\frac{1}{c\sqrt{T}})^{\delta'} \sum_{t=1}^{T} E[||\boldsymbol{Z}_t||^{2+\delta'}|\mathcal{F}_{t-1}]$$

for any positive constant c and some positive constant δ'. The last term converges to zero in probability because its expectation is bounded by

$$(A.37)\quad \frac{1}{c^{\delta'}}(\frac{1}{T})^{\delta'/2} \sup_{t\geq 1} E[||\boldsymbol{Z}_t||^{2+\delta'}]$$

and we can take $\delta = 2\delta'$.

Given that the maximum likelihood estimator of $\boldsymbol{\theta}$ is consistent, the rest of proof is to check Condition (A3) in Basawa, Feigin, and Heyde (1976) for its asymptotic normality. Because it is standard to do this procedure, we omit the details. *Q.E.D.*

References

[1] Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principle," in Petrov, B.N. and Csaki, F. eds., *Second International Symposium on Information Theory*, Akademia Kiado, Budapest, 267-281.

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

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

[4] Basawa,I.V., Feigin, P.D., and Heyde, C.C. (1976), "Asymptotic Properties of Maximum Likelihood Estimators for Stochastic Processes," *Sankhya*, **36**, 259-270.

[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. and Tong, H. (1985), "On the Use of the Deterministic Lyapunov Function for the Ergodicity of Stochastic Difference Equation," *Adv. Appl. Prob.*, Vol. 17, 666-678.

[8] Duffie, D. (1992), *Dynamic Asset Pricing Theory*, Princeton University Press.

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

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

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

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

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

[14] 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).

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

[16] Kunitomo, N. and Sato, S. (1997), "Stationary and Non-stationary Simultaneous Switching Autoregressive Models with an Application to Financial Time Series," Unpublished Manuscript.

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

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

[19] 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.

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

[21] Tweedie, R.L. (1973), "Criteria for rates of convergence of Markov chains with application to queueing theory," In *Papers in Probability, Statistics, and Analysis*, ed., J.F.C.Kingman and G.E.H.Reuter, Cambridge University Press.

[22] Weiss, A.A. (1986), "Asymptotic Theory for ARCH Models: Estimation and Testing," *Econometric Theory*, **2**, 107-131.

Table 1: Nikkei 225 Index (Jan. 1, 1985 – Feb. 20, 1987)

(i) Modelling with daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4174.02	-4239.82	-4297.53	-4298.33
SSARIMA(2)	-4178.76	-4239.16	-4297.63	-4297.67
SSARIMA(3)	-4176.76	-4237.32	-4295.67	-4295.83
ARIMA(0)	-4143.68	-4213.05	-4268.96	-4270.93
ARIMA(1)	-4166.34	-4234.70	-4286.93	-4287.64
ARIMA(2)	-4170.94	-4234.17	-4286.98	-4287.19
ARIMA(3)	-4169.05	-4232.41	-4285.00	-4285.52

Selected Model: SSARIMA(1,0) – ARCH(3) AIC: -4298.33

	1	2	3	σ
a (up)	0.2676	–	–	0.0030
b (down)	0.0752	–	–	0.0038

	α_1	α_2	α_3
ARCH coef.	0.19	0.37	0.07

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Daily effects	1	3.20	1.97	1.81	2.26	0.60

(ii) Modelling without daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4163.98	-4227.60	-4284.09	-4282.21
SSARIMA(2)	-4165.67	-4227.98	-4282.41	-4280.46
SSARIMA(3)	-4163.97	-4226.54	-4280.61	-4278.66
ARIMA(0)	-4135.46	-4207.97	-4255.02	-4253.02
ARIMA(1)	-4154.55	-4223.05	-4271.77	-4269.77
ARIMA(2)	-4155.80	-4224.08	-4270.08	-4268.08
ARIMA(3)	-4154.64	-4222.85	-4268.51	-4266.51

Selected Model: SSARIMA(1,0) – ARCH(2) AIC: -4284.09

	1	2	3	σ
a (up)	0.2719	–	–	0.0043
b (down)	0.0693	–	–	0.0056

	α_1	α_2	α_3
ARCH coef.	0.23	0.34	–

Table 1 (cont.): Nikkei 225 Index (Feb. 21, 1987 – Dec. 31, 1989)

(i) Modelling with daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4828.46	-5114.55	-5167.47	-5167.29
SSARIMA(2)	-4836.28	-5129.81	-5166.09	-5165.70
SSARIMA(3)	-4834.52	-5129.29	-5164.11	-5163.80
ARIMA(0)	-4741.86	-5096.03	-5151.75	-5154.42
ARIMA(1)	-4740.41	-5098.90	-5163.12	-5163.86
ARIMA(2)	-4762.00	-5102.56	-5161.77	-5162.18
ARIMA(3)	-4761.40	-5122.66	-5159.77	-5160.23

Selected Model: SSARIMA(1,0) – ARCH(2) AIC: -5167.47

	1	2	3	σ
a (up)	0.2083	–	–	0.0058
b (down)	0.0776	–	–	0.0068

	α_1	α_2	α_3
ARCH coef.	0.70	0.16	–

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Daily effects	1	0.68	0.63	0.56	0.58	0.87

(ii) Modelling without daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4782.42	-5111.23	-5169.37	-5168.64
SSARIMA(2)	-4787.27	-5122.35	-5167.88	-5167.01
SSARIMA(3)	-4785.35	-5121.75	-5165.89	-5165.07
ARIMA(0)	-4658.77	-5071.68	-5149.73	-5152.83
ARIMA(1)	-4658.20	-5086.33	-5161.78	-5162.11
ARIMA(2)	-4666.46	-5091.56	-5160.47	-5160.42
ARIMA(3)	-4664.74	-5093.18	-5158.50	-5158.43

Selected Model: SSARIMA(1,0) – ARCH(2) AIC: -5169.37

	1	2	3	σ
a (up)	0.2177	–	–	0.0048
b (down)	0.0624	–	–	0.0057

	α_1	α_2	α_3
ARCH coef.	0.70	0.17	–

Table 1 (cont.): Nikkei 225 Index (Jan. 1, 1990 – Jun 11, 1992)

(i) Modelling with daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-3117.69	-3174.44	-3203.24	-3210.30
SSARIMA(2)	-3135.06	-3193.48	-3222.36	-3225.73
SSARIMA(3)	-3133.11	-3191.66	-3220.53	-3223.77
ARIMA(0)	-3115.75	-3171.58	-3201.35	-3209.02
ARIMA(1)	-3115.24	-3174.18	-3201.64	-3208.62
ARIMA(2)	-3132.40	-3192.57	-3222.68	-3225.92
ARIMA(3)	-3130.46	-3190.85	-3220.73	-3223.93

Selected Model: ARIMA(2,0) – ARCH(3)   AIC: -3225.92

	1	2	3	σ
a (up)	0.0866	-0.2126	–	0.0153
b (down)	0.0866	-0.2126	–	0.0153

	α_1	α_2	α_3
ARCH coef.	0.21	0.28	0.09

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Daily effects	1	0.42	0.66	0.34	0.50	–

(ii) Modelling without daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-3112.69	-3166.62	-3191.60	-3200.92
SSARIMA(2)	-3129.99	-3185.10	-3212.33	-3217.27
SSARIMA(3)	-3128.13	-3183.23	-3210.41	-3215.28
ARIMA(0)	-3110.39	-3162.32	-3189.49	-3199.39
ARIMA(1)	-3110.52	-3166.22	-3191.03	-3200.28
ARIMA(2)	-3127.91	-3184.49	-3213.30	-3218.28
ARIMA(3)	-3126.04	-3182.67	-3211.33	-3216.28

Selected Model: ARIMA(2,0) – ARCH(3)   AIC: -3218.28

	1	2	3	σ
a (up)	0.1097	-0.2114	–	0.0113
b (down)	0.1097	-0.2114	–	0.0113

	α_1	α_2	α_3
ARCH coef.	0.24	0.26	0.12

Table 1 (cont.): Nikkei 225 Index (Jun 12, 1992 – Dec. 15, 1994)

(i) Modelling with daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-3541.45	-3565.11	-3591.16	-3604.67
SSARIMA(2)	-3539.49	-3563.46	-3598.44	-3605.82
SSARIMA(3)	-3539.28	-3562.02	-3597.07	-3603.82
ARIMA(0)	-3544.81	-3566.08	-3592.55	-3603.28
ARIMA(1)	-3543.32	-3565.25	-3591.39	-3604.31
ARIMA(2)	-3541.35	-3563.49	-3598.73	-3605.75
ARIMA(3)	-3541.10	-3561.97	-3597.33	-3603.77

Selected Model: SSARIMA(2,0) – ARCH(3) AIC: <u>-3605.82</u>

	1	2	3	σ
a (up)	-0.0920	0.1228	–	0.0138
b (down)	-0.0005	0.1125	–	0.0127

	α_1	α_2	α_3
ARCH coef.	0.10	0.28	0.12

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Daily effects	1	0.62	0.47	0.45	0.31	–

(ii) Modelling without daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-3532.68	-3554.39	-3571.48	-3591.93
SSARIMA(2)	-3530.71	-3552.51	-3577.91	-3592.61
SSARIMA(3)	-3529.69	-3551.01	-3577.14	-3590.61
ARIMA(0)	-3536.31	-3554.74	-3572.29	-3589.20
ARIMA(1)	-3534.36	-3553.42	-3570.62	-3591.18
ARIMA(2)	-3532.38	-3551.49	-3577.72	-3591.97
ARIMA(3)	-3531.29	-3549.89	-3577.05	-3589.98

Selected Model: SSARIMA(2,0) – ARCH(3) AIC: -3592.61

	1	2	3	σ
a (up)	-0.1192	0.0950	–	0.0105
b (down)	-0.0153	0.0861	–	0.0095

	α_1	α_2	α_3
ARCH coef.	0.12	0.19	0.18

Table 2 : S & P 500 Index (Jan. 1, 1990 – Jun 11, 1992)

(i) Modelling with daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4034.48	-4033.51	-4043.41	-4043.50
SSARIMA(2)	-4033.53	-4032.35	-4042.65	-4042.56
SSARIMA(3)	-4031.83	-4030.65	-4040.89	-4041.11
ARIMA(0)	-4036.02	-4035.33	-4046.02	-4045.88
ARIMA(1)	-4036.10	-4035.28	-4045.17	-4045.38
ARIMA(2)	-4035.08	-4034.05	-4044.32	-4044.39
ARIMA(3)	-4033.42	-4032.37	-4042.60	-4042.99

Selected Model: ARIMA(0,0) – ARCH(2) AIC: -4046.02

	1	2	3	σ
a (up)	–	–	–	0.0094
b (down)	–	–	–	0.0094

	α_1	α_2	α_3
ARCH coef.	0.03	0.11	–

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.
Daily effects	1	0.80	0.49	0.75	0.91	–

(ii) Modelling without daily effects

	AIC			
	ARCH(0)	ARCH(1)	ARCH(2)	ARCH(3)
SSARIMA(1)	-4033.19	-4033.18	-4039.60	-4040.22
SSARIMA(2)	-4032.07	-4031.72	-4038.50	-4038.95
SSARIMA(3)	-4030.49	-4030.01	-4036.78	-4037.53
ARIMA(0)	-4033.46	-4034.17	-4041.31	-4041.56
ARIMA(1)	-4034.32	-4034.62	-4040.84	-4041.74
ARIMA(2)	-4033.09	-4033.07	-4039.61	-4040.36
ARIMA(3)	-4031.58	-4031.39	-4037.94	-4039.04

Selected Model: ARIMA(1,0) – ARCH(3) AIC: -4041.74

	1	2	3	σ
a (up)	0.0640	–	–	0.0081
b (down)	0.0640	–	–	0.0081

	α_1	α_2	α_3
ARCH coef.	0.03	0.09	0.08