CONDITIONAL HETEROSCEDASTIC AUTOREGRESSIVE MOVING AVERAGE MODELS WITH SEASONAL PATTERNS

by

LAU, Suk Ting

A Thesis for
the Degree of Master of Philosophy

Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong

1999

DECLARATION

I hereby declare that the thesis entitled "Conditional Heteroscedastic Autoregressive Moving Average Models with Seasonal Patterns" is original and has not been submitted for a degree, or other qualifications in this University or any other institutes. It does not contain any material, partly or wholly, published or written previously by others, except those references quoted in the text.

LAU, Suk Ting

ACKNOWLEDGEMENTS

I thank my chief supervisor Dr. H. Wong for his invaluable suggestions and encouragement during my research development. Sincere thanks are also due to the supervisor for comments that led to great improvement of the research.

I would also like to express my sincere gratitude to Prof. W. K. Li for giving me many useful advice on time series methodology.

My thanks go to Dr. L. K. Li who gave me a number of advice on computing and the exchange rate data of USD/JPY.

I wish to convey my thanks to all the staffs in the Department of Applied Mathematics and thank the Research Degree Committee of The Hong Kong Polytechnic University for supporting the MPhil. studies.

Abstract of thesis entitled

'Conditional Heteroscedastic Autoregressive Moving Average Models with Seasonal Patterns'

submitted by LAU, Suk Ting

for the degree of Master of Philosophy

at The Hong Kong Polytechnic University in November 1998

The earlier research in time series mainly concentrated on models that assume a constant one-period forecast variance. In reality, however, the assumption may not be met in all cases, especially in economics and finance. Therefore, much recent work has been directed towards the relaxation of the constant conditional variance assumption, namely allowing the conditional variance to change over time and keeping the unconditional variance constant.

Tsay (1987) proposed the conditional heteroscedastic autoregressive moving average (CHARMA) model. One of the advantages of the model is that it includes the autoregressive conditional heteroscedastic (ARCH) model and the random coefficient autoregressive (RCA) models as its special cases. Both models characterize time series with varying conditional variance in different representations. Therefore, the CHARMA model is more flexible and is able to model data from a wider perspective.

It is also believed that seasonal pattern can be an important phenomenon in the conditional variance and so the purpose of this research is to study seasonal conditional heteroscedasticity and extend the CHARMA model to the seasonal CHARMA model. One of the advantages of our approach is that the relevant time series can be modeled in a parsimonious parameterization.

The invertibility and stationarity conditions for the model are derived. We study all the procedures for building up the model. These include the test for varying conditional variance, estimation of the model parameters by the least squares, and the maximum likelihood method and diagnostic checking methodology for testing the adequacy of the fitted model. Two empirical examples are discussed in detail: the exchange rate of US dollar/Japanese Yen and the money supply (M1) of United States.

In addition, the ability of capturing volatility will be compared among the proposed model and the GARCH family since the GARCH family is widely used in modeling conditional heteroscedasticity.

It is found that the exchange rate and money supply have a clear seasonal volatility. The proposed model can capture this effect and produce good forecasts.

CONTENTS

CHAPTER 1 INTRODUCTION

The earlier research in time series mainly concentrated on models that assume a constant one-period forecast variance. The classic book by Box and Jenkins (1994) summarized important contributions in this area. In reality, however, the assumption may not be met in all cases, especially in economics and finance. Therefore, much recent work has been directed towards the relaxation of the constant conditional variance assumption, i.e., allowing the conditional variance to change over time and keeping the unconditional variance constant.

Kendall (1953) pointed out, in the modeling of economic data, it seems reasonable to let the coefficients of the model change through time as the economy changes. Kendall considered his coefficients to follow quadratic trends, which is a special case of the random coefficient autoregressive (RCA) models (further investigated by Conlisk (1974, 1976), Andel (1976), and Nicholls and Quinn (1982)).

The autoregressive conditional heteroscedastic (ARCH) processes were introduced by Engle (1982) to model the United Kingdom's inflation. The errors of the processes are serially uncorrelated with zero mean and constant unconditional variances, but the conditional variances are changing over time. The conditional variance depends upon variances in the conditioning set such as the past residuals while the unconditional variance is a constant as traditionally assumed. Since then, the model has been highly popular with econometricians, statisticians and finance researchers. Bollerslev (1986) gave a generalization of Engle's model, which is known as GARCH (Generalized Autoregressive Conditional Heteroscedasticity).

GARCH model can be interpreted as an ARMA model for the squared residuals, whereas ARCH model can be regarded as an AR model.

Many generalizations of the GARCH model then appeared. In many applications of the GARCH model, an unit root may be present. In order to achieve stationarity, Engle and Bollerslev (1986) introduced integrated GARCH (IGARCH) model. The IGARCH model is stationary but not covariance stationary. Bollerslev (1987) proposed GARCH-*t* model in which the residuals are conditionally *t*-distributed. Exponential GARCH (EGARCH) model was proposed by Nelson (1991) for tackling with leverage effects. Other than EGARCH, all the coefficents of the conditional variances of the above models are non-negative but EGARCH allows the parameters to be negative. So, the conditional variance not only depend on the value of lagged residuals but also their signs. Engle and Ng (1993) introduced the quadratic GARCH (QGARCH) model to cope with the problem that the error distribution of the time series being skewed to the left.

Tsay (1987) proposed the conditional heteroscedastic autoregressive moving average (CHARMA) model, in which the observed process is the usual ARMA model whereas the innovational models follows a purely random coefficient transfer function model. The model links up the ARCH and the RCA models by their common special feature - varying conditional variance. Hence, the ARCH and the RCA models become special cases of the CHARMA model, which allows a more parsimonious description of the data and has a broader platform to capture the data. Wong and Li (1997) gave the multiplicative version of the CHARMA model.

In many economic series there is a clear presence of seasonality effect. Box and Jenkins proposed general multiplicative seasonal ARMA model to capture this effect. Later in this report, the seasonality of the monthly money supply of U.S. will be discussed. It is widely believed that seasonal patterns can exist in the conditional variance structure, especially for high frequency data. For example, in the financial market, it is expected the volatility of Monday is greater than the others since more information accumulates over the weekend. The length of non-trading day between Friday and Monday is longer than that between other days of the week. Therefore, in daily data, the seasonal heteroscedastic pattern may be present. For intraday series, many traders also report their belief of a difference in the volatility across the day. The volatility is usually higher when the market opens, and then drops around lunch hour. Volatility picks up again near the closure of the market. A possible explanation is as follows: it is known that news will affect the market. The increase in volatility at the open of the markets might reflect the news release at that time and also the accumulated information overnight. Around the lunch hour, there might be less market activity and news release so that the volatility drops. The activity is more active before the closure. This might be due to the long waiting time after the market closure and the openings of the other countries' market. Therefore, the intraday periodic pattern is highly possible and investigation is needed.

Baillie and Bollerslev (1989) worked with the conditional variances of daily spot exchange rate for six currencies. Day-of-week and vacation effects present in the data are captured by introducing seasonal dummy variables in the GARCH model. Dacorogna *et al*., (1993, 1996) modeled foreign exchange (FX) rates, which has daily and weekly seasonal heteroscedasticity, by changing time scales. Bollerslev and Hodrick (1995) observed that there is a strong seasonality in the conditional variance of the monthly New York Stock Exchange (NYSE) dividend growth rates. Florentini and Maravall (1996) found that the conditional variance on the monthly Spanish money supply has a strong seasonal pattern and dealt with it by seasonal adjustment. Bollerslev and Ghysels (1996) proposed a periodic autoregressive conditional heteroscedasticity (P-ARCH) model to model the seasonal heteroscedasticity. The above description shows that seasonal conditional heteroscedasticity deserves more extensive attention. Therefore, in this research, the CHARMA models will be extended to the seasonal CHARMA models to model the seasonality effect from another approach.

The details of the proposed model including the stationary and invertible conditions are stated and discussed in chapter 2. In chapter 3, the procedures for identifying and setting up the model are developed. We discuss the test for varying conditional variance, depict the use of the Akaike's information criterion (AIC), and the Schwarz's information criterion (SIC) to select the order of the innovation equation, use the least squares and the maximum likelihood method for estimating the model and finally perform the diagnostic checking for the model. The exchange rate of the U. S. Dollar / Japanese Yen is an empirical example for these investigations.

In chapter 4, we will compare the performance among the proposed model and the GARCH family (GARCH, P-GARCH, and seasonal GARCH) for the Yen data. In chapter 5, money supply (M1) of the United States from 1954 to 1994 will be used as a further example. It is interesting to observe the existence of seasonal conditional heteroscedasticity in these important economic series. We will draw the conclusion and suggest direction of further research in chapter 6.

CHAPTER 2 SEASONAL CHARMA MODEL AND ITS THEORETICAL PROPERTIES

2.1 THE SEASONAL CHARMA MODEL

2.1.1 The autoregressive model

Let Y_t be a stationary time series, if

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \qquad (2.1)$$

and the white noise a_t is assumed to be normally and independently distributed with zero mean, constant variance σ_a^2 and is independent of Y_{t-k}, $k = 1, 2, \dots, p$, i.e.,

$$E(a_t) = 0$$
$$E(a_t a_s) = \begin{cases} \sigma_a^2 & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}$$
$$E(a_t Y_{t-k}) = 0, \text{ where } k = 1, 2, \dots, p$$

Then (2.1) is called an autoregressive (AR) process of order p, AR(p). It was proposed by Yule (1927) and has been given a detailed account by Box and Jenkins (1994).

The reason for the name 'autoregressive' is that a linear model

$$Y = \phi_1 Y_1 + \phi_2 Y_2 + \dots + \phi_p Y_p + a$$

relating a 'dependent' variable Y to a set of 'independent' variables Y_1, Y_2, ..., Y_p, plus an error term a, is often referred to as a regression model. In (2.1), the variable Y is regressed on previous values of itself, hence the model is autoregressive.

If we define an autoregressive operator of order p by

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p,$$

where B is the backshift operator, such that $BY_t = Y_{t-1}$, then the autoregressive model may be written as

$$\phi(B) Y_t = a_t,$$

where a_t is a white noise sequence.

The stationarity condition is that all the roots of $\phi(B) = 0$ are outside the unit circle.

2.1.2 The autoregressive conditional heteroscedastic model

Engle (1982) proposed the autoregressive conditional heteroscedastic (ARCH) models. The models allow for the conditional variance to depend on the squares of previous innovations. With financial data it captures the tendency for volatility clustering, as written by Mandelbrot (1963), "..., large changes to be followed by large changes - of either sign - and small changes tend to be followed by small changes, ...".

Let Y_t be a time series and a_t is serially uncorrelated with mean zero, but the conditional variance equals to h_t that may be changing through time. The ARCH(q) model can be written as

$$Y_t | \psi_{t-1} \sim N(X_t b, h_t)$$

where $h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_q a_{t-q}^2,$

$$a_t = Y_t - X_t b,$$

ψ_{t-1} is the information set available at time t-1,

and X_t^T be a vector of explanatory variables included in ψ_{t-1}.

To make sure the conditional variance to be positive, the necessary condition is $\alpha_0 > 0$, $\alpha_i \geq 0$, for i = 1, ..., q. The model is covariance stationary if and only if the sum of α_i, for i = 1, ..., q, is less than unity.

2.1.3 The random coefficient autoregressive model

The random coefficient autoregressive (RCA) models allow the autoregressive coefficient to change through time. The conditional variance depends on the squares of previous observations. The RCA(p) model can be written as

$$Y_t = (\phi_1 + b_{1,t})Y_{t-1} + (\phi_2 + b_{2,t})Y_{t-2} + \cdots + (\phi_p + b_{p,t})Y_{t-p} + e_t,$$

which is equivalent to

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + a_t$$

where $a_t = \sum_{i=1}^{p} b_{i,t} Y_{t-i} + e_t$,

$b_{i,t}$ are zero mean random variables,

and e_t is a white noise sequence.

The model generalizes the constant coefficient AR(p) model with a_t being the innovation term for Y_t. The random coefficient vector $B_t = (b_{1,t}, b_{2,t}, \ldots, b_{p,t})^T$ is a sequence of identically and independently distributed (i.i.d.) random vectors with mean zero and constant covariance matrix $\Sigma = (\sigma_{ij})$. B_t and e_t are mutually independent. Thus, it can also be written as

$$Y_t | \psi_{t-1} \sim D(g_t, h_t^*),$$

where g_t is the same as that defined in the ARCH model, and

$$h_t^* = \sigma_e^2 + \sum_{i=1}^{p}\sum_{j=1}^{p}\sigma_{ij} Y_{t-i} Y_{t-j} .$$

2.1.4 The conditional heteroscedastic autoregressive moving average model

If Y_t for the ARCH(q) model is a AR(p) model with varying conditional variance, then the model named as ARCH(p,q) and can be written as follows,

$$Y_t | \psi_{t-1} \sim D(g_t , h_t)$$

where $g_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p}$,

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_q a_{t-q}^2 ,$$

$$a_t = Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i} ,$$

and ψ_{t-1} is the information set available at time t-1.

Tsay (1987) proved that a stationary ARCH(p,q) model is second-order equivalent to a special case of a stationary RCA(p+q) model. Using this connection, Tsay (1987) proposed the CHARMA model that includes ARCH model and RCA model as special cases.

The CHARMA(p,q,r,s) model is defined by

$$\phi(B)(Y_t - \mu) = \theta(B) a_t , \qquad \text{(Observation Equation)}$$

and $\delta_t(B) a_t = \omega_{0,t}[\hat{Y}_{t-1}(1) - \mu] + \omega_t^*(B)(Y_t - \mu) + e_t .$ (Innovation Equation)

In the Observation Equation, which is a traditional ARMA model, $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, and $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$ are constant polynomials in the backshift operator B, and of order p and q, respectively. For model stationarity and invertibility, all of the roots of $\phi(B) = 0$ and $\theta(B) = 0$ are outside the unit circle and have no common roots.

In the Innovation Equation, $\omega_t^*(B) = \omega_{1,t}B + \cdots + \omega_{s,t}B^s$ and $\delta_t(B) = 1 - \delta_{1,t}B - \cdots - \delta_{r,t}B^r$ are purely random coefficient polynomials of degrees s and r. Here e_t is a white noise sequence, $\hat{Y}_{t-1}(1)$ is the one-step-ahead forecast of Y_t. Thus $\hat{Y}_{t-1}(1) = E(Y_t|\psi_{t-1})$ and ψ_{t-1} is the available information at time t-1. A CHARMA(p,0,0,p) with $\omega_{0,t} = 0$ reduces to an RCA model; a CHARMA(p,0,q,0) model with uncorrelated $\delta_{i,t}$ is second-order equivalent to an ARCH(p,q) model.

2.1.5 The seasonal autoregressive moving average model

Since many time series data contain both non-seasonal and seasonal patterns, Box-Jenkins develop the multiplicative seasonal ARMA model, $\text{ARMA}(p,q)\times(P,Q)_s$,

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)a_t$$

where s is the seasonal period,

and a_t is a white noise sequence.

Assumptions for the seasonal ARMA model

1. $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\Phi(B^s) = 1 - \Phi_1 B^s - \ldots - \Phi_P B^{Ps}$, $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$, and $\Theta(B^s) = 1 - \Theta_1 B^s - \ldots - \Theta_Q B^{Qs}$ are constant coefficient polynomials in B of degrees p, P, q and Q, respectively.

2. All of the roots of $\phi(B) = 0$, $\Phi(B^s) = 0$, $\theta(B) = 0$, and $\Theta(B^s) = 0$ are outside the unit circle and $\phi(B)$, $\Phi(B^s)$, $\theta(B)$, and $\Theta(B^s)$ have no common factors.

The autocorrelation function (acf) and partial autocorrelation function (pacf) are useful tool for identifying the tentative model. Here we shall briefly describe and figure out the patterns for the non-seasonal cases (Figure 1) and the similar patterns occur at multiples of lag s for the seasonal cases.

1. For an AR(p) model, the acf dies out while the pacf cuts off at lag p.
2. For an MA(q) model, the acf cuts off at lag q while the pacf dies out.
3. For an ARMA(p,q) model, both acf and pacf die out.

Figure 1. Patterns in acf and pacf for some ARMA models

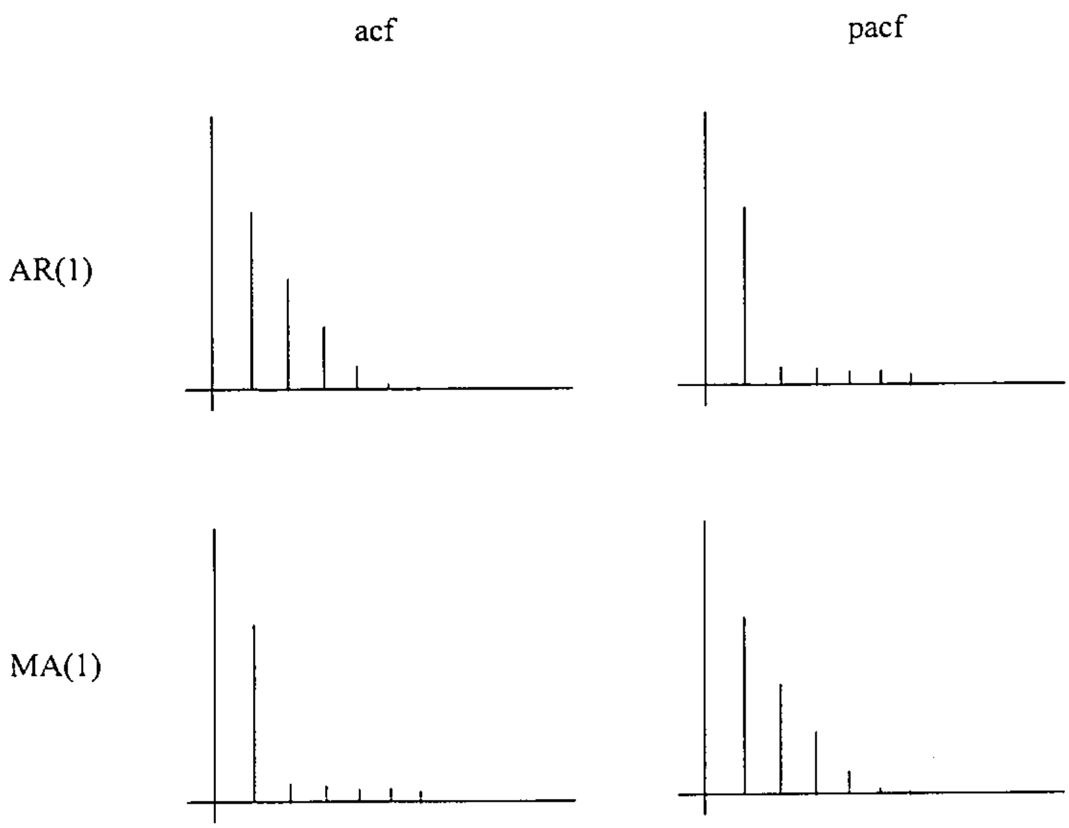
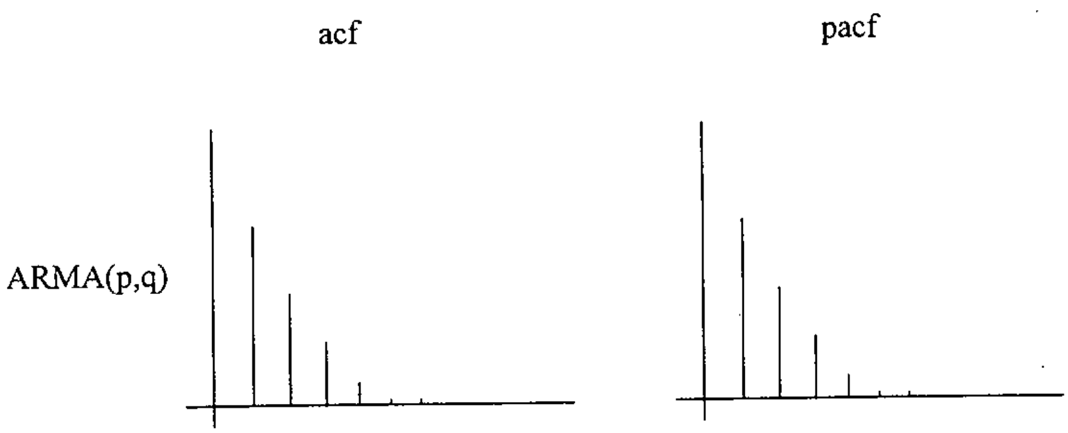


Figure 1(cont.). Patterns in acf and pacf for some ARMA models



2.1.6 The proposed model

We now extend the CHARMA model to the seasonal case. With the convenience in describing the Observation Equation, the model will be extended in a multiplicative way. The properties of models following the Observation Equation, the seasonal ARMA model, have been discussed in Box-Jenkins (1994) and we list some of them in the above section. Properties relating to the Innovation Equation in our model will be described shortly.

Let

$$\phi(B)\Phi(B^d)(Y_t - \mu) = \theta(B)\Theta(B^d)a_t \qquad \text{Observation Equation} \qquad (2.2)$$

$$\delta_t(B)\Delta_t(B^d)a_t = [\omega_t(B)\Omega_t(B^d)-1](Y_t - \mu) + e_t \qquad \text{Innovation Equation} \qquad (2.3)$$

where d is the seasonal period,

μ is the mean level of Y_t,

e_t is a white noise sequence,

and B is the backshift operator, i.e. $B^j Y_t = Y_{t-j}$ where $j > 0$.

In the Observation Equation (2.2), $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\Phi(B^d) = 1 - \Phi_1 B^d - \ldots - \Phi_P B^{Pd}$, $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$, and $\Theta(B^d) = 1 - \Theta_1 B^d - \ldots - \Theta_Q B^{Qd}$.

In the Innovation Equation (2.3), $\delta_t(B) = 1 - \delta_{1,t} B - \ldots - \delta_{r,t} B^r$, $\Delta_t(B^d) = 1 - \Delta_{1,t} B^d - \ldots - \Delta_{R,t} B^{Rd}$, $\omega_t(B) = 1 + \omega_{1,t} B + \ldots + \omega_{s,t} B^s$, and $\Omega_t(B^d) = 1 + \Omega_{1,t} B^d + \ldots + \Omega_{S,t} B^{Sd}$.

To facilitate the discussion of the theoretical properties of the model, the following formulations and observations are useful.

Rewriting (2.3),

$$(1-\sum_{i=1}^{r}\delta_{i,t}B^i)(1-\sum_{j=1}^{R}\Delta_{j,t}B^{jd})a_t = [(\sum_{i=1}^{s}\omega_{i,t}B^i)(\sum_{j=1}^{S}\Omega_{j,t}B^{jd})-1](Y_t-\mu)+e_t\,,$$

$$(1-\lambda_{1,t}B-\ldots-\lambda_{r+Rd,t}B^{r+Rd})a_t = (\gamma_{1,t}B+\ldots+\gamma_{s+Sd,t}B^{s+Sd})(Y_t-\mu)+e_t\,,$$

we have the following representation of the Innovation Equation:

$$\lambda_t(B)a_t = \gamma_t(B)(Y_t-\mu)+e_t \tag{2.4}$$

For further analysis, we put (2.4) in another way

$$\begin{aligned} a_t = {} & \lambda_{1,t}a_{t-1}+\cdots+\lambda_{r+Rd,t}a_{t-(r+Rd)}+\gamma_{1,t}(Y_{t-1}-\mu)+\cdots+ \\ & \gamma_{s+Sd}(Y_{t-(s+Sd)}-\mu)+e_t \end{aligned} \tag{2.5}$$

and let $\lambda_t = (\lambda_{1,t}, \lambda_{2,t}, \ldots, \lambda_{r+Rd,t})^T$,

$\gamma_t = (\gamma_{1,t}, \gamma_{2,t}, \ldots, \gamma_{s+Sd,t})^T$.

Here ^{T} denotes the transpose of a matrix.

Properties of the innovational series

Following Tsay (1987), it can be shown that:

1. $E(a_t|\psi_{t-1}) = 0$

2. $E(a_t e_{t-i}) = \begin{cases} \sigma_e^2 & \text{if } i=0 \\ 0 & \text{otherwise} \end{cases}$

3. $var(a_t|\psi_{t-1}) = E(a_t^2|\psi_{t-1})$

$$= A_{t-1}^T \Lambda A_{t-1} + F_{t-1}^T \Gamma F_{t-1} + \sigma_e^2,$$

where $A_{t-1} = (a_{t-1}, \ldots, a_{t-(r+Rd)})^T$,

$F_{t-1} = (Y_{t-1} - \mu, \ldots, Y_{t-(s+Sd)} - \mu)^T$,

$\Lambda = E(\lambda_t \lambda_t^T)$,

and $\Delta = E(\gamma_t \gamma_t^T)$.

4. $E(a_t a_{t-j}) = 0 \qquad \text{if } j > 0$

Assumptions for the Innovation Equation

1. $\delta_t(B)$, $\Delta_t(B^d)$, $\omega_t(B)$, and $\Omega_t(B^d)$ are purely random coefficient polynomials in B of degrees r, R, s and S, respectively.
2. Since the Innovation Equation depends on the available information up to time t-1 only, it does not contain the Y_t term.
3. $\{e_t\}$ is a sequence of i.i.d. random variables with mean zero and finite positive variance σ_e^2.
4. $\{\lambda_t\}$ and $\{\gamma_t\}$ are sequences of i.i.d. random vectors with zero expectations and constant covariance matrices $E(\lambda_t \lambda_t^T) = \Lambda = \Lambda_{ij}$ (i, j = 1, . . ., r+Rd) and $E(\gamma_t \gamma_t^T) = \Gamma = \Gamma_{ij}$ (i, j = 1, . . ., s+Sd), respectively. Both Λ and Γ are nonnegative definite.

5. The three sequences of random variates $\{e_t\}$, $\{\lambda_t\}$, and $\{\gamma_t\}$ are mutually independent.

6. e_t, λ_t, and γ_t are jointly normal.

Under the assumptions above for the Observation Equation and the Innovation Equation, it can be shown that the seasonal CHARMA model is both stationary and invertible.

2.2 INVERTIBILITY AND STATIONARITY CONDITION OF THE MODEL

For ease of exposition, the theoretical properties of the model will be illustrated and discussed in the special case p = q = r = s = P = Q = R = S = 1 and d = 4. The general case follows as a direct extension and will be indicated in the course of the discussion.

Consider the Observation Equation,

$$(1 - \phi_1 B)(1 - \Phi_1 B^4)(Y_t - \mu) = (1 - \theta_1 B)(1 - \Theta_1 B^4)a_t ,$$

which can be written as

$$(1 - c_1 B - c_4 B^4 - c_5 B^5)(Y_t - \mu) = (1 - g_1 B - g_4 B^4 - g_5 B^5)a_t ,$$

where $c_1 = \phi_1$, $c_4 = \Phi_1$, $c_5 = -\phi_1\Phi_1$, $g_1 = \theta_1$, $g_4 = \Theta_1$, $g_5 = -\theta_1\Theta_1$,

i.e.,

$$\begin{aligned} Y_t - \mu = c_1(Y_{t-1} - \mu) + c_4(Y_{t-4} - \mu) + c_5(Y_{t-5} - \mu) + \\ a_t - g_1 a_{t-1} - g_4 a_{t-4} - g_5 a_{t-5} \end{aligned} \tag{2.6}$$

For the Innovation Equation (2.3),

$$(1-\delta_{1,t}B-\Delta_{1,t}B^4+\delta_{1,t}\Delta_{1,t}B^5)a_t = (\omega_{1,t}B+ \Omega_{1,t}B^4+\omega_{1,t}\Omega_{1,t}B^5)(Y_t - \mu) + e_t ,$$

it can be written as

$$(1 - \lambda_{1,t}B - \lambda_{4,t}B^4 - \lambda_{5,t}B^5)a_t = (\gamma_{1,t}B + \gamma_{4,t}B^4 + \gamma_{5,t}B^5)(Y_t - \mu) + e_t ,$$

where $\lambda_{1,t} = \delta_{1,t}$, $\lambda_{4,t} = \Delta_{1,t}$, $\lambda_{5,t} = -\delta_{1,t}\Delta_{1,t}$, $\gamma_{1,t} = \omega_{1,t}$, $\gamma_{4,t} = \Omega_{1,t}$, $\gamma_{5,t} = -\omega_{1,t}\Omega_{1,t}$,

i.e., $$a_t = \lambda_{1,t}a_{t-1} + \lambda_{4,t}a_{t-4} + \lambda_{5,t}a_{t-5} + \gamma_{1,t}(Y_t - \mu) + \gamma_{4,t}(Y_t - \mu) + \gamma_{5,t}(Y_t - \mu) + e_t \qquad (2.7)$$

2.2.1 Stationarity condition

Combining the observation and Innovation Equation, i.e., substituting a_t in (2.7) into (2.6) gives,

$$Y_t - \mu = (c_1 + \gamma_{1,t})(Y_{t-1} - \mu) + (c_4 + \gamma_{4,t})(Y_{t-4} - \mu) + (c_5 + \gamma_{5,t})(Y_{t-5} - \mu)+$$
$$(\lambda_{1,t} - g_1)a_{t-1} + (\lambda_{4,t} - g_4)a_{t-4} + (\lambda_{5,t} - g_5)a_{t-5} + e_t$$

In matrix form the above can be written as,

$$Y_t^* = (D_t + M)\, Y_{t-1}^* + N_t \qquad (2.8)$$

where

$$Y_t^* = \begin{bmatrix} Y_t - \mu \\ Y_{t-1} - \mu \\ Y_{t-2} - \mu \\ Y_{t-3} - \mu \\ Y_{t-4} - \mu \\ a_t \\ a_{t-1} \\ a_{t-2} \\ a_{t-3} \\ a_{t-4} \end{bmatrix},$$

$$D_t = \begin{bmatrix} \gamma_{1,t} & \gamma_{2,t} & \gamma_{3,t} & \gamma_{4,t} & \gamma_{5,t} & \lambda_{1,t} & \lambda_{2,t} & \lambda_{3,t} & \lambda_{4,t} & \lambda_{5,t} \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ \gamma_{1,t} & \gamma_{2,t} & \gamma_{3,t} & \gamma_{4,t} & \gamma_{5,t} & \lambda_{1,t} & \lambda_{2,t} & \lambda_{3,t} & \lambda_{4,t} & \lambda_{5,t} \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix},$$

$$
M = \begin{bmatrix} c_1 & 0 & 0 & c_4 & c_5 & -g_1 & 0 & 0 & -g_4 & -g_5 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},
$$

and

$$
N_t = \begin{bmatrix} e_t \\ 0 \\ 0 \\ 0 \\ 0 \\ e_t \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
$$

Now (2.8) is in the form of a first order Markov Chain.

In general form, model (2.2) and (2.3) can be written as

$$Y_t^* = (D_t + M)\, Y_{t-1}^* + N_t$$

where $Y_t^* = (Y_t - \mu, Y_{t-1} - \mu, \ldots, Y_{t-m_1+1} - \mu, a_t, a_{t-1}, \ldots, a_{t-m_2+1})^T$,

$m_1 = \max(p+Pd, s+Sd)$, $m_2 = \max(q+Qd, r+Rd)$,

D_t is a matrix whose first and (m_1+1)th row are the random coefficient of the Innovation Equation and the rest are zero.

N_t is a $(m_1+m_2)\times 1$ vector whose first and (m_1+1)th elements are e_t and the rest are zero.

$$\text{Let } M = \left(\begin{array}{c} U \\ \hline I \mid 0 \\ \hline 0 \\ \hline 0 \mid I \end{array} \right),$$

where U is a constant coefficient vector of the Observation Equation. The (2,1) and (4,2) block of M is the $(m_1-1)\times(m_1-1)$ identity matrix plus a $(m_1-1)\times 1$ null vector and $(m_2-1)\times(m_2-1)$ identity matrix plus a $(m_2-1)\times 1$ null vector, respectively. The (2,2) and

(4,1) block is the $(m_1-1)\times m_2$ and $(m_2-1)\times m_1$ null matrix, respectively. The third block is a $1\times(m_1+m_2)$ null vector.

The model (2.2) and (2.3) is stationary if all of the eigenvalues of the matrix $(M\otimes M + D_\bullet)$ where $D_\bullet = E(D_t\otimes D_t)$ are less than unity in modulus, where $\otimes$ is the Kronecker product of matrices.

Proof:

The following identities will be used.

1. $\text{vec}(ABC) = (C^T\otimes A)\text{vec}B$

2. $(\prod_{i=0}^{j} A_i)\otimes(\prod_{k=0}^{j} B_k) = \prod_{i=0}^{j}(A_i \otimes B_i)$

Rewrite (2.8) recursively, we can get

$$\begin{aligned}
Y_t^\bullet &= (M + D_t)Y_{t-1}^\bullet + N_t \\
&= N_t + (M + D_t)N_{t-1} + (M + D_t)(M + D_{t-1})Y_{t-2}^\bullet \\
&= N_t + \sum_{j=1}^{\infty}[\prod_{k=0}^{j-1}(M + D_{t-k})]N_{t-j}
\end{aligned}$$

For positive integers j and v, define

$$S_{j-1,t} = \prod_{k=0}^{j-1}(M + D_{t-k}) \text{ and } \Xi_{v,t} = \sum_{j=1}^{v} S_{j-1,t}N_{t-j}.$$

$$\begin{aligned}
\text{vec } E(\Xi_{v,t}\Xi_{v,t}^T) &= \text{vec } E(\sum_{j=1}^{v} S_{j-1,t}N_{t-j})(\sum_{j=1}^{v} S_{j-1,t}N_{t-j})^T \\
&= \text{vec } E(\sum_{j=1}^{v} S_{j-1,t}N_{t-j}N_{t-j}^T S_{j-1,t}^T) \\
&= E[\sum_{j=1}^{v} \text{vec}(S_{j-1,t}N_{t-j}N_{t-j}^T S_{j-1,t}^T)] \\
&= E[\sum_{j=1}^{v}(S_{j-1,t} \otimes S_{j-1,t})\text{vec } N_{t-j}N_{t-j}^T] && \text{Identity 1} \\
&= E\{\sum_{j=1}^{v}[\prod_{k=0}^{j-1}(M + D_{t-k}) \otimes (M + D_{t-k})]\text{vec } N_{t-j}N_{t-j}^T\} && \text{Identity 2} \\
&= \sum_{j=1}^{v}\{\prod_{k=0}^{j-1}[E(M + D_{t-k}) \otimes (M + D_{t-k})]\text{vec } E(N_{t-j}N_{t-j}^T)\} \\
&= \sum_{j=1}^{v}(M \otimes M + D_\bullet)^j \text{vec}(N_\bullet)
\end{aligned}$$

where $D. = E(D_t \otimes D_t)$,

and $N. = E(N_{t-j} N_{t-j}^T)$, which does not depend on t-j.

To prove the model is stationary it suffices to show that $E(\Xi_{v,t} \Xi_{v,t}^T)$ converges as v→∞, i.e., $\sum_{j=1}^{v} (M \otimes M + D.)^j$ converges as v→∞.

If (M⊗M + D.) may be diagonalized as

$$(M \otimes M + D.) = P \Psi P^{-1}$$

where Ψ is a diagonal matrix having the eigenvalues of (M + D.) along its main diagonal, and zero elsewhere, then $(M \otimes M + D.)^j = P \Psi^j P^{-1}$, and it can easily be seen that if the diagonal elements of Ψ are less than unity in modulus, then Ψ^j converges to zero. Hence, $\lim_{v \to \infty} \sum_{j=0}^{v} \Psi^j = (I - \Psi)^{-1}$.

$$\begin{aligned}
\lim_{v \to \infty} \sum_{j=0}^{v} (M \otimes M + D.)^j \mathrm{vec}(N.) &= P(I - \Psi)^{-1} P^{-1} \mathrm{vec}(N.) \\
&= (I - P \Psi P^{-1})^{-1} \mathrm{vec}(N.) \\
&= (I - M \otimes M - D.)^{-1} \mathrm{vec}(N.)
\end{aligned}$$

Therefore, if all the eigenvalues of (M⊗M + D.) are less than unity in modulus, $\sum_{j=1}^{\infty} (M \otimes M + D.)^j$ converges so that the stationarity of the full model will hold.

If M + D. is not diagonalizable, we shall resort to a result of Issacson and Keller (1966). $\sum_{j=0}^{\infty} A^j$ is convergent, if and only if $\rho(A) < 1$, where $\rho(A) \equiv \max_s |\lambda_s(A)|$ and $\lambda_s(A)$ denotes the *s*th eigenvalue of A. Then, applying to our case and from the above, the stationarity of the full model will still hold.

2.2.2 Invertibility condition

Combine Innovation Equation at time t and Observation Equation at time t-1,

$$a_t = (\lambda_{1,t} + \gamma_{1,t})a_{t-1} - \gamma_{1,t}g_1a_{t-2} + \lambda_{4,t}a_{t-4} + (\lambda_{5,t} - \gamma_{1,t}g_4)a_{t-5} - \gamma_{1,t}g_5a_{t-6} +$$

$$\gamma_{1,t}c_1Y_{t-2} + \gamma_{4,t}Y_{t-4} + (\gamma_{1,t}c_4 + \gamma_{5,t})Y_{t-5} + \gamma_{1,t}c_5Y_{t-6} + e_t$$

In matrix form,

$$\underset{A_t^*}{\begin{bmatrix} a_t \\ a_{t-1} \\ a_{t-2} \\ a_{t-3} \\ a_{t-4} \\ a_{t-5} \end{bmatrix}} = \left[\underset{D_t}{\begin{bmatrix} \lambda_{1,t} & 0 & 0 & \lambda_{4,t} & \lambda_{5,t} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}} + \underset{M_t}{\begin{bmatrix} \gamma_{1,t} & -\gamma_{1,t}g_1 & 0 & 0 & -\gamma_{1,t}g_4 & -\gamma_{1,t}g_5 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}} \right] \underset{A_{t-1}^*}{\begin{bmatrix} a_{t-1} \\ a_{t-2} \\ a_{t-3} \\ a_{t-4} \\ a_{t-5} \\ a_{t-6} \end{bmatrix}} + \underset{N_t}{\begin{bmatrix} y_t \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}$$

where $y_t = \gamma_{1,t}c_1Y_{t-2} + \gamma_{4,t}Y_{t-4} + (\gamma_{1,t}c_4 + \gamma_{5,t})Y_{t-5} + \gamma_{1,t}c_5Y_{t-6} + e_t$ (2.9)

In the general case,

$$A_t^* = (D_t + M_t)A_{t-1}^* + N_t$$

where $A_t^* = (a_t, a_{t-1}, \ldots, a_{t-m})^T$

$m = \max(q+Qd, r+Rd)$

D_t = m×m matrix whose first row is λ_t^T and the rest are zero (λ_t^T is first row of D_t in (2.9))

$$M_t = \left(\frac{U_t}{I|0} \right)$$

The (2,1) block of M_t is the identity, the (2,2) block is the null vector and U_t is the coefficient of A_{t-1} other than λ_t^T.

$$
\begin{aligned}
A_t^* &= (M_t + D_t)A_{t-1}^* + N_t \\
&= N_t + (M_t + D_t)N_{t-1} + (M_t + D_t)(M_{t-1} + D_{t-1})A_{t-2}^* \\
&= N_t + \sum_{j=1}^{\infty}[\prod_{k=0}^{j-1}(M_{t-k} + D_{t-k})]N_{t-j}
\end{aligned}
$$

For positive integers j and v, define

$$S_{j-1,t} = \prod_{k=0}^{j-1}(M_{t-k} + D_{t-k}) \text{ and}$$

$$\Xi_{v,t} = \sum_{j=1}^{v} S_{j-1,t} N_{t-j}$$

$$
\begin{aligned}
\text{vec}\, E(\Xi_{v,t}\Xi_{v,t}^T) &= E[\sum_{j=1}^{v}(S_{j-1,t} \otimes S_{j-1,t})\text{vec}\, N_{t-j}N_{t-j}^T] \\
&= E\{\sum_{j=1}^{v}[\prod_{k=0}^{j-1}(M_{t-k} + D_{t-k}) \otimes (M_{t-k} + D_{t-k})]\text{vec}\, N_{t-j}N_{t-j}^T\} \\
&= \sum_{j=1}^{v}\{\prod_{k=0}^{j-1}[E(M_{t-k} + D_{t-k}) \otimes (M_{t-k} + D_{t-k})]\text{vec}\, E(N_{t-j}N_{t-j}^T)\} \\
&= \sum_{j=1}^{v}(M_\bullet + D_\bullet)^j \,\text{vec}(N_\bullet)
\end{aligned}
$$

where $(M_\bullet + D_\bullet) = E(M_t \otimes M_t + D_t \otimes D_t)$,

and $N_\bullet = E(N_{t-j}N_{t-j}^T)$, which does not depend on t-j for stationary Y_t.

Similar to the condition of stationarity, the model (2.2) and (2.3) is invertible if all of the eigenvalues of the matrix $(M_\bullet + D_\bullet) = E(M_t \otimes M_t + D_t \otimes D_t)$ are less than unity in modulus.

CHAPTER 3 IDENTIFICATION, ESTIMATION AND DIAGNOSTIC CHECKING OF THE SEASONAL CHARMA MODEL

3.1 THE BOX–JENKINS 3 – STEPS APPROACH

Box-Jenkins' iterative approach for constructing linear time series model basically consists of three steps:

1. Identification of the preliminary specifications of the model
2. Estimation of the parameters of the model
3. Diagnostic checking of model adequacy

At the identification stage, we examine the autocorrelation function (acf) and the partial autocorrelation function (pacf) of the data and then choose a particular model from the general class of ARMA models, i.e., selects the order of the non-seasonal and seasonal autoregressive and moving average polynomials necessary to represent the model.

After identifying a particular ARIMA model, the parameters of that model are estimated. There are basically two methods available for estimating these parameters. One method is the least squares method and the other is the maximum likelihood method.

Then, by applying various diagnostic checks, such as residual analyses, and fitting extra parameters, we can determine whether or not the model adequately represents the data. If any inadequacies are detected, a new model must be identified and the cycle of steps 1 to 3 repeated. Finally, the best model that passes all the checks is used to generate forecasts.

For building up our model, the procedures are quite similar to the 3–steps approach. At the stage of initial examination, we first identify and estimate the Observation Equation and then perform similar procedure of building the seasonal ARMA model. In the identification stage of the Innovation Equation, we examine the acf of the squared residuals obtained from the Observation Equation — testing whether the residuals possess the varying conditional variance property and tentatively identifying all possible orders for the Innovation Equation.

Then, we estimate the Innovation Equation only and calculate the corresponding criterion function for all possible equations for choosing the suitable one. After determining the order for the Innovation Equation, we estimate the observation and Innovation Equation together. Finally, we also perform diagnostic checking for the standardized squared residuals. As choosing the right order for time series models is a rather difficult problem, we give a brief account of some standard methods in the next section.

3.1.1 Order selection

The Akaike's information criterion (AIC) (Akaike 1974) and the Schwarz's information criterion (SIC) (Schwarz 1978) are widely used in the identification of the order of linear models and the favored model is determined by minimizing the criterion function among various possible models. Both the AIC and SIC criterion consider maximizing the log-likelihood (the first term of equation (3.1) and (3.2)) with a penalty factor for increasing the parameters in the model (the second term of the equations). SIC penalizes for additional parameters more than AIC when n is large. So SIC favors the model with less parameters as compared with AIC.

$$\text{AIC} = -2 \times \text{log-likelihood} + 2 \times k_l \tag{3.1}$$

$$\text{SIC} = -2 \times \text{log-likelihood} + k_l \times \ln(n) \tag{3.2}$$

where k_l is the number of parameters estimated in the model.

3.2 TEST FOR VARYING CONDITIONAL VARIANCE

The residuals defined in the Box-Jenkins model is a strict white noise sequence, the properties of which are described in section 2.1.1. After fitting the Observation Equation, we need to check for the validity of this assumption. However, the standard diagnostic checks only detect whether or not the residuals are autocorrelated and cannot detect for other types of departures from the white noise assumption.

Granger and Andersen (1978) and Miller (1979) suggested that the autocorrelation function of the squared residuals could be useful in identifying non-linear time series. The autocorrelation function of $\hat{a}_t^2$, the squared residuals obtained from a fitted ARMA model, is estimated by

$$r_{aa}(k) = \frac{\sum_{t=k+1}^{n}(\hat{a}_t^2 - \hat{\sigma}^2)(\hat{a}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{n}(\hat{a}_t^2 - \hat{\sigma}^2)^2}, \qquad k = 0, 1, \ldots, n-1,$$

where $\hat{\sigma}^2 = \dfrac{\sum \hat{a}_t^2}{n}$.

McLeod and Li (1983) showed that the squared residual autocorrelations are useful for detecting the presence of non-linearity in the second moment and proposed a test based on them. The statistic is:

$$Q_{aa} = n(n+2)\sum_{k=1}^{M}\frac{r_{aa}^2(k)}{n-k},$$

and is asymptotically $\chi^2(M)$ distributed if the $\hat{a}_t$ are residuals from an ARMA model fitted to data. A condition for the choice of M is that $M/n \to 0$ as $n \to \infty$.

Wong and Li (1995) proposed a rank portmanteau statistic that is a rank version of the McLeod-Li statistic and found that it is more robust than its parametric counterpart against additive outliers.

The rank autocorrelation function at lag k for a_t^2 is defined as

$$\tilde{r}_k = \frac{\sum_{t=k+1}^{n}(R_t - \bar{R})(R_{t-k} - \bar{R})}{\sum_{t=1}^{n}(R_t - \bar{R})^2}, \quad k = 1, 2, \ldots, n\text{-}1,$$

where $R_t = \text{rank}(\hat{a}_t^2)$,

$$\bar{R} = \frac{\sum R_t}{n} = \frac{n+1}{2},$$

and $\sum(R_t - \bar{R})^2 = \dfrac{n(n^2-1)}{12}$.

And the statistics

$$Q_R = \sum_{k=1}^{M}\frac{(\tilde{r}_k - \mu_k)^2}{\tilde{\sigma}_k^2},$$

is $\chi^2(M)$ distributed with $\mu_k = E(\tilde{r}_k)$ and $\tilde{\sigma}_k^2 = v(\tilde{r}_k)$.

Moran (1948) showed that $E(\tilde{r}_k) = -\dfrac{n-k}{n(n-1)}$, and Dufour and Roy (1986) showed that $v(\tilde{r}_k) = \dfrac{5n^4 - (5k+9)n^3 + 9(k-2)n^2 + 2k(5k+8)n + 16k^2}{5(n-1)^2 n^2 (n+1)}$.

By using simulated and real data, it is found that (Wong and Li, 1995) the power of Q_{aa} statistic is slightly greater than Q_R when there are ARCH effects and without outliers in the data. However, when there are outliers, the Q_{aa} statistic may fail to detect ARCH effect but Q_R does not.

3.2.1 Simulation

As discussed in the last section, it is reasonable to believe that examining the acf of the squared residuals plays an important role in the detection of ARCH pattern for non-seasonal case. Thus, it is of interest to find out the use of the acf for determining the form of the proposed model. Our aim is to investigate if the graph of the acf different among various orders and if the acf shows the pattern at seasonal lags for seasonal conditional heteroscedasticity. We simulate twenty time series for five models, each series with 500 observations, then calculate the corresponding acf and plot them in appendices 1 to 5.

Here are the five simulated models:

a.

$$Y_t - 0.4Y_{t-1} = a_t,$$
$$a_t = \Omega_t Y_{t-5} + e_t,$$
$$\Rightarrow h_t = 0.01 + 0.4356Y_{t-5}^2.$$

b.

$$Y_t - 0.4Y_{t-1} = a_t,$$
$$(1 - \Delta_t B^5)a_t = e_t,$$
$$\Rightarrow h_t = 0.01 + 0.4356a_{t-5}^2.$$

c.

$$Y_t - 0.4Y_{t-1} = a_t,$$
$$(1-\Delta B^5)a_t = \Omega_t Y_{t-5} + e_t,$$
$$\Rightarrow h_t = 0.01 + 0.4356Y_{t-5}^2 + 0.36a_{t-5}^2.$$

d.

$$Y_t - 0.4Y_{t-1} = a_t,$$
$$a_t = [(1-\omega_t B)(1-\Omega_t B^5) - 1]Y_t + e_t,$$
$$\Rightarrow h_t = \beta_0 + \beta_1 Y_{t-1}^2 + \beta_2 Y_{t-5}^2 + \beta_3 Y_{t-6}^2,$$
$$\text{where } \beta_0 = 0.01, \beta_1 = 0.4356, \text{and } \beta_2 = 0.36.$$

e.

$$Y_t - 0.4Y_{t-1} = a_t,$$
$$(1-\delta_t B)(1-\Delta_t B^5)a_t = e_t,$$
$$\Rightarrow h_t = \beta_0 + \beta_1 a_{t-1}^2 + \beta_2 a_{t-5}^2 + \beta_3 a_{t-6}^2,$$
$$\text{where } \beta_0 = 0.01, \beta_1 = 0.4356, \text{and } \beta_2 = 0.36.$$

Figure 2. A summary of the patterns of acf for different models

a.

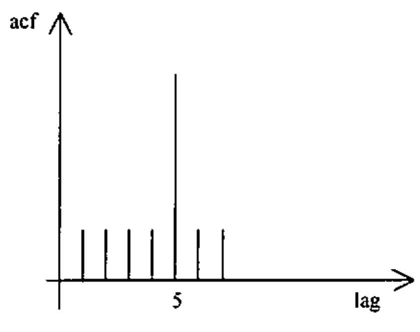
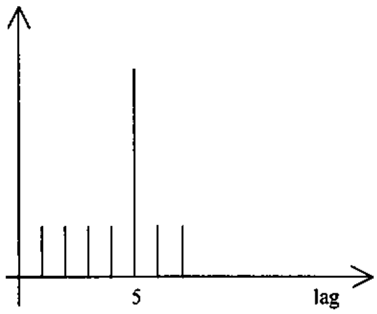
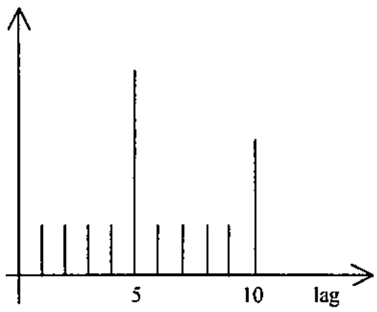


Figure 2(cont.). A summary of the patterns of acf for different models

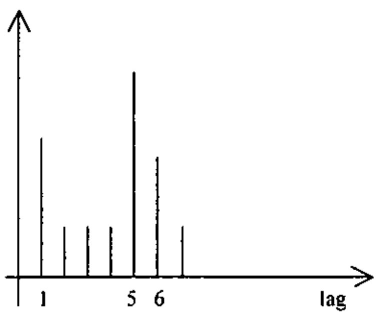
b.



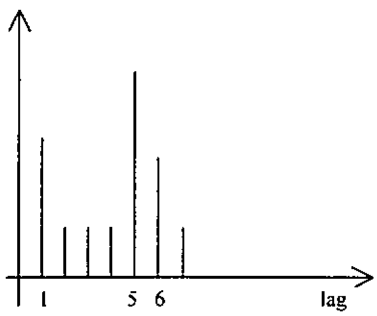
c.



d.



e.



In Figure 2, we summarize the patterns of the autocorrelation functions for the corresponding models. In the Box-Jenkins model, the patterns of the acf are clearly different between the AR model and the MA model (Figure 1). Although the acf of the squared residuals cannot show the same difference in this simulation, it can be seen that there is also a spike at the seasonal lag or multiples of seasonal lags, which can serve as indicators that the time series have seasonal conditional heteroscedastic variance.

3.2.2 Empirical example

As stated in the introduction, many economic series may have seasonal pattern in the conditional variance because of information accumulation. Friedman and Vandersteel (1982) examined the statistical properties of daily foreign exchange spot rates from the period 1 Jun 1973 to 14 Sept 1979 for nine currencies relative to U.S. Dollar: German Mark, Swiss Franc, Pound Sterling, Japanese Yen, Dutch Guilder, French Franc, Canadian Dollar, Belgian Franc and Italian Lira. It was found that all spot rates are leptokurtotic, with massive tails and a sharper central peak than normal. Moreover, it suggested that the leptokurtotic phenomenon might be caused by the changes of the process over time and μ_t and σ_t should be estimated as functions of time-varying economic and institutional variables.

Hsieh (1988) examined ten daily exchange rates against U.S. Dollar, from 1974 to 1983 and found all of them are not independent and identically distributed. Moreover, variances changing over time appeared to be the strongest characterisitic in the data. Hsieh set up the ARCH model with daily and holiday dummy variables to describe the data. The model only fitted the data for Swiss Franc, Canadian

Dollar, and Deutsche Mark, but not for Japanese Yen, and British Pound. It can be seen from his model that the variances of daily exchange rates are larger whenever the trading days span a weekend or a holiday.

Baillie and Bollerslev (1989) found that the autocorrelations of the squared residuals of daily exchange rate are significant around the seasonal lags and set up a GARCH model with daily dummy and holiday dummy variables. From the model, it confirmed that distinctive daily seasonality and holiday effects are present in the conditional variance.

From the above, it confirms that seasonality in the conditional variance present in the daily exchange rate. Although there are many papers in the study of foreign exchange rates, it is hard to say there is a model that can fully characterize the seasonality picture of the behavior of the exchange rate volatility. Foreign exchange is an important market in the financial world, therefore, we would like to study the weekend effect in the exchange rate again using the seasonal CHARMA model. Our aim is to see whether seasonal heteroscedasticity can be characterized in a parsimonious way as compared with dummy variables. In addition, we like to see how well the seasonal CHARMA characterizes the distribution of volatility.

Daily closing exchange rate of the U.S. Dollar / Japanese Yen (USD/JPY) from 5 May 86 until 5 Jun 95 for a total of 2369 observations excluding weekends, Sundays, and bank holidays is considered. The missing data (because of holidays) are handled by assuming they have the same rate as the previous day. The time series plot of the exchange rate X_t is given in Figure 3.

Figure 3. The time series plot of the exchange rate of USD/JPY

Figure 4. The plot of acf for Yen

Autocorrelations: YEN

Lag	Auto-Corr.	Stand. Err.	-1 -.75 -.5 -.25 0 .25 .5 .75 1	Box-Ljung	Prob.
1	-.011	.021	.*.	.282	.596
2	-.018	.021	.*.	1.086	.581
3	-.013	.021	.*.	1.516	.679
4	.007	.021	.*.	1.630	.803
5	-.015	.021	.*.	2.195	.821
6	-.028	.021	*).	4.074	.667
7	.012	.021	.*.	4.421	.730
8	.012	.021	.*.	4.763	.783
9	.013	.020	.*.	5.191	.817
10	.058	.020	.)*	13.230	.211
11	-.001	.020	.*.	13.231	.278
12	.022	.020	.*.	14.376	.277
13	.018	.020	.*.	15.176	.296
14	.007	.020	.*.	15.291	.359
15	.044	.020	.)*	20.013	.171
16	.015	.020	.*.	20.561	.196

Plot Symbols: Autocorrelations * Two Standard Error Limits .
Total cases: 2369 Computable first lags: 2368

Figure 5. The plot of pacf for YEN

```
Partial Autocorrelations:    YEN

   Pr-Aut- Stand.
Lag  Corr.    Err. -1   -.75   -.5 -.25    0    .25  .5    .75   1
                   -----------------------------------------------
  1  -.011   .021                           .*.
  2  -.019   .021                           .*.
  3  -.014   .021                           .*.
  4   .006   .021                           .*.
  5  -.016   .021                           .*.
  6  -.028   .021                           *'.
  7   .011   .021                           .*.
  8   .011   .021                           .*.
  9   .014   .021                           .*.
 10   .059   .021                           .'*
 11   .001   .021                           .*.
 12   .024   .021                           .*.
 13   .022   .021                           .*.
 14   .009   .021                           .*.
 15   .049   .021                           .'*
 16   .020   .021                           .*.

Plot Symbols:       Autocorrelations *     Two Standard Error Limits .
Total cases:  2369      Computable first lags:  2368
```

We consider the first differences of the series' natural logarithms: $Y_t = 100\times\ln(X_t / X_{t-1})$. None of the autocorrelations and partial autocorrelations (Figures 4 and 5) for Y_t are significant at the 5% level, we can conclude that the Observation Equation of the series is a white noise model, i.e.

$$Y_t = a_t,$$

where a_t is the white noise.

In a series of daily data, from Monday to Friday, it can be postulated that the seasonal lags are multiples of 5. For the residuals the Ljung and Box (1978) portmanteau test statistic Q(5) = 2.195, which is insignificant at the 5% level compared with the critical value $\chi^2(5) = 11.070$.

Next, we investigate whether there is any evidence of non-linearity in the second moments. By examining the autocorrelation coefficients of squared residuals of the model (Figure 6), the large acf at lag 1, 3 and 5 indicate that non-linear structure may be present in the residuals. Employing the conditional heteroscedasticity test, $Q_{aa}(5) = 44.3301$, the squared residuals are highly significant at the 5% level. There is strong evidence of conditional heteroscedasticity.

Figure 6. The plot of acf for the squared residuals

```
Autocorrelations:    SQRES

     Auto-  Stand.
Lag  Corr.   Err. -1  -.75  -.5 -.25   0   .25  .5   .75   1   Box-Ljung  Prob.
                  ----------------------------------------
  1   .097   .021                    .³.*                         22.426   .000
  2   .022   .021                    .*.                          23.626   .000
  3   .057   .021                    .³*                          31.474   .000
  4   .028   .021                    .³*                          33.392   .000
  5   .068   .021                    .³*                          44.320   .000
  6   .053   .021                    .³*                          51.062   .000
  7   .004   .021                    .*.                          51.096   .000
  8   .058   .021                    .³*                          58.971   .000
  9   .034   .020                    .³*                          61.770   .000
 10   .022   .020                    .*.                          62.953   .000
 11   .041   .020                    .³*                          66.919   .000
 12   .034   .020                    .³*                          69.701   .000
 13   .009   .020                    .*.                          69.890   .000
 14   .081   .020                    .³.*                         85.732   .000
 15   .009   .020                    .*.                          85.922   .000
 16  -.010   .020                    .*.                          86.153   .000

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:   2369     Computable first lags:  2368
```

3.3 ORDER SELECTION IN THE INNOVATION EQUATION

3.3.1 Initial estimates for parameters of the Innovation Equation

Before calculating the criterion function, we need to estimate the covariance matrices of λ_t, γ_t and e_t. Let vech(G) be the half-column stacking vector of a symmetric matrix G by using elements either on or below the main diagonal, and vech2(G) = vech(G^*), where G^* is obtained from G by multiplying all of the off-diagonal elements by 2.

From properties 3 of the innovation series described in section 2.1.6, the conditional expectation of a_t^2 can be written as

$$
\begin{aligned}
h_t &= E(a_t^2 | \psi_{t-1}) \\
&= (A_{t-1} \otimes A_{t-1})^T \text{vec}(\Lambda) + (F_{t-1} \otimes F_{t-1})^T \text{vec}(\Gamma) + \sigma_e^2 \\
&= X_{1,t-1}^T \text{vech2}(\Lambda) + X_{2,t-1}^T \text{vech2}(\Gamma) + \sigma_e^2 \\
&= W_t^T \beta + \sigma_e^2
\end{aligned}
$$

where $X_{1,t-1} = \text{vech}(A_{t-1} A_{t-1}^T)$,

$X_{2,t-1} = \text{vech}(F_{t-1} F_{t-1}^T)$,

$\beta = (\text{vech2}(\Lambda)^T, \text{vech2}(\Gamma)^T)^T$,

and $W_t = (X_{1,t-1}^T, X_{2,t-1}^T)^T$.

The conditional log-likelihood function of $\{a_t\}$ is

$$
L_n = -\frac{1}{2} \sum_{t=k}^{n} [\ln(h_t) + \frac{a_t^2}{h_t}], \tag{3.3}
$$

where $k = \max(r+Rd, s+Sd)$,

and $h_t = E(a_t^2 | \psi_{t-1}) = W_t^T \beta + \sigma_e^2$.

Let $\upsilon = (\sigma_e^2 \;\; \beta^T)^T$ be the vector that we want to estimate, the likelihood estimating equation is given by setting $f(\upsilon) = \frac{\partial L_n}{\partial \upsilon}$ to zero. We have

$$
\frac{\partial L_n}{\partial \sigma_e^2} = -\frac{1}{2} \sum_{t=k}^{n} \frac{1}{h_t} + \frac{1}{2} \sum_{t=k}^{n} \frac{a_t^2}{h_t^2},
$$

$$
\frac{\partial L_n}{\partial \beta(i)} = -\frac{1}{2} \sum_{t=k}^{n} \frac{W_t(i)}{h_t} + \frac{1}{2} \sum_{t=k}^{n} \frac{a_t^2 W_t(i)}{h_t^2},
$$

where $W_t(i)$ is the ith element in W_t^T.

The Hessian for the Newton's method can be computed as follows,

$$f'(\upsilon) = \frac{\partial^2 L_n}{\partial \upsilon \partial \upsilon^T} ,$$

$$\frac{\partial^2 L_n}{\partial \sigma_e^2 \partial \sigma_e^2} = \frac{1}{2}\sum_{t=k}^{n} \frac{1}{h_t^2} - \sum_{t=k}^{n} \frac{a_t^2}{h_t^3} ,$$

$$\frac{\partial^2 L_n}{\partial \sigma_e^2 \partial \beta(j)} = \frac{1}{2}\sum_{t=k}^{n} \frac{W_t(j)}{h_t^2} - \sum_{t=k}^{n} \frac{a_t^2 W_t(j)}{h_t^3} ,$$

$$\frac{\partial^2 L_n}{\partial \beta(i) \partial \beta(j)} = \frac{1}{2}\sum_{t=k}^{n} \frac{W_t(i) W_t(j)}{h_t^2} - \sum_{t=k}^{n} \frac{a_t^2 W_t(i) W_t(j)}{h_t^3} ,$$

and $\upsilon_n = \upsilon_{n-1} - [f'(\upsilon)]^{-1} f(\upsilon)$.

3.3.2 Model selection criteria

In section 3.1.1, we have described how to choose the suitable order among various tentative linear models by the Akaike's information criterion (AIC) and the Schwarz's information criterion (SIC). Not only useful in the linear models, both criteria are also used heavily in identification of the order of ARCH models. Nevertheless, it is worthwhile to notice that the theoretical performance and statistical properties of the criterion function in the ARCH models are not widely studied.

Recall the definition,

$$AIC = -2\times\ln(\text{maximum likelihood}) + 2k_l \quad (3.4)$$

and $$SIC = -2\times\ln(\text{maximum likelihood}) + k_l\times\ln(n) \quad (3.5)$$

where k_l denotes the number of parameters estimated in the model.

Applying to the Innovation Equation, from (3.3)

$$AIC_A = \sum_{t=k}^{n}[\ln(\hat{h}_t) + a_t^2 / \hat{h}_t] + 2k_1 ,$$

and $$SIC_A = \sum_{t=k}^{n}[\ln(\hat{h}_t) + a_t^2 / \hat{h}_t] + k_1 \times \ln(n) ,$$

where a_t is the residual from the Observation Equation (2.2), $\hat{h}_t$ is the (conditional) maximum likelihood estimate from section 3.3.1 and k_1 is the number of parameters in the Innovation Equation (2.3).

3.3.3 Using mean square error (MSE)

It is natural to consider the mean square error as another loss function instead of the likelihood function, thus similar to AIC_A and SIC_A, some authors like to consider

$$AIC_B = \ln(MSE) + 2k_1 / n,$$

$$SIC_B = \ln(MSE) + k_1 \times \ln(n) / n,$$

and $$MSE = \frac{1}{n}\sum_{t=k}^{n}\{a_t^2 - \hat{h}_t\}^2 .$$

Here the error is defined as the difference between the squared residuals from the Observation Equation (2.2), a_t^2, and the conditional variances in the Innovation Equation (2.3), $\hat{h}_t$. Both criteria are normalized by the sample size n. The order of the equation is again determined by the minimum criteria. Tsay (1987) investigated the performance of these criterion functions by simulations for the CHARMA models and the results were reasonable.

3.3.4 Simulation

In this section, we shall carry out a simulation experiment to test the ability of the various criteria for choosing the correct order of the Innovation Equation for the seasonal cases and assume that all a_t are coming from seasonal ARMA models. Three models are generated and the numbers of observations of each time series are 200, 500 and 800 respectively. For each combination of model and sample size, due to the large amount of computer time involved, we only consider 100 replications. Since the number of replications is only 100, we understand the results can only serve as general guides for model selection.

Recall the Innovation Equation (2.3) of the proposed model is,

$$\delta_t(B)\Delta_t(B^d)a_t = [\omega_t(B)\Omega_t(B^d)-1](Y_t - \mu) + e_t$$

and we symbolize the equation as $(r,s)\times(R,S)_d$, where r, s, R, and S is the order of $\delta_t(B)$, $\Delta_t(B^d)$, $\omega_t(B)$, and $\Omega_t(B^d)$, respectively, and d is the seasonal period.

All the models have the same Observation Equation and the only difference is in the Innovation Equation. The three models are as follows and the right hand side of the parentheses is the specified order for the Innovation Equation.

Model 1 (M1, $(0,1)_5$)

$$(1-0.4B)Y_t = a_t$$

$$a_t = \Omega_t Y_{t-5} + e_t$$

$$\Rightarrow h_t = 0.01 + 0.64Y_{t-5}^2$$

Model 2 (M2, $(1,0)\times(0,1)_5$)

$$(1-0.4B)Y_t = a_t$$

$$(1-\delta_t B)a_t = \Omega_t Y_{t-5} + e_t$$

$$\Rightarrow h_t = 0.01 + 0.64a_{t-1}^2 + 0.25Y_{t-5}^2$$

Model 3 (M3, $(0,1)\times(0,1)_5$)

$$(1-0.4B)Y_t = a_t$$

$$a_t = [(1-\omega_t B)(1-\Omega_t B^5) - 1]Y_t + e_t$$

$$\Rightarrow h_t = \beta_0 + \beta_1 Y_{t-1}^2 + \beta_2 Y_{t-5}^2 + \beta_3 Y_{t-6}^2$$

where $\beta_0 = 0.01, \beta_1 = 0.64, \text{and } \beta_2 = 0.25$

Table 1. The AIC criterion based on the mean square error

n		(0,1)	$(0,1)_5$	(1,0)	$(1,0)_5$	$(0,2)_5$	$(2,0)_5$	$(1,0)\times(0,1)_5$	$(0,1)\times(1,0)_5$	$(0,1)\times(0,1)_5$	$(1,0)\times(1,0)_5$
	M1	1	53	8	24	0	1	6	4	1	2
200	M2	11	14	20	12	0	3	21	6	7	6
	M3	14	15	9	14	0	1	6	16	20	5
	M1	7	60	1	16	0	3	1	7	3	2
500	M2	4	11	14	5	0	0	37	7	5	17
	M3	9	9	4	7	0	3	8	5	42	13
	M1	3	65	5	13	0	2	5	5	0	2
800	M2	5	10	9	2	0	2	32	9	14	17
	M3	6	8	6	4	0	2	3	3	50	18

Table 2. The SIC criterion based on the mean square error

n		(0,1)	$(0,1)_5$	(1,0)	$(1,0)_5$	$(0,2)_5$	$(2,0)_5$	$(1,0)\times(0,1)_5$	$(0,1)\times(1,0)_5$	$(0,1)\times(0,1)_5$	$(1,0)\times(1,0)_5$
	M1	1	57	8	26	0	1	3	2	0	2
200	M2	11	13	26	15	0	1	20	6	5	3
	M3	21	18	11	11	0	2	4	15	13	5
	M1	7	61	2	21	0	3	1	4	0	1
500	M2	7	12	20	3	0	0	36	5	6	11
	M3	10	11	4	7	0	3	8	6	38	13
	M1	2	69	6	14	0	2	1	4	0	2
800	M2	6	12	9	2	0	0	35	9	13	14
	M3	7	10	7	4	0	0	4	7	46	15

Table 3. The AIC criterion based on the log-likelihood

n		(0,1)	$(0,1)_5$	(1,0)	$(1,0)_5$	$(0,2)_5$	$(2,0)_5$	$(1,0)\times(0,1)_5$	$(0,1)\times(1,0)_5$	$(0,1)\times(0,1)_5$	$(1,0)\times(1,0)_5$
200	M1	0	90	0	2	0	2	5	0	1	0
	M2	0	0	2	0	0	0	91	0	2	5
	M3	0	1	0	0	0	0	3	18	73	5
500	M1	0	93	0	0	0	0	7	0	0	0
	M2	0	0	0	0	0	0	99	0	1	0
	M3	0	0	0	0	0	0	0	3	97	0
800	M1	0	98	0	0	0	1	1	0	0	0
	M2	0	0	0	0	0	0	100	0	0	0
	M3	0	0	0	0	0	0	0	0	100	0

Table 4. The SIC criterion based on the log-likelihood

n		(0,1)	$(0,1)_5$	(1,0)	$(1,0)_5$	$(0,2)_5$	$(2,0)_5$	$(1,0)\times(0,1)_5$	$(0,1)\times(1,0)_5$	$(0,1)\times(0,1)_5$	$(1,0)\times(1,0)_5$
200	M1	0	95	0	2	0	2	1	0	0	0
	M2	0	0	4	0	0	0	93	0	2	1
	M3	1	1	0	0	0	0	5	5	63	25
500	M1	0	97	0	0	0	0	3	0	0	0
	M2	0	0	0	0	0	0	100	0	0	0
	M3	0	0	0	0	0	0	0	6	94	0
800	M1	0	99	0	0	0	1	0	0	0	0
	M2	0	0	0	0	0	0	100	0	0	0
	M3	0	0	0	0	0	0	0	1	99	0

First, we use the maximum likelihood estimation to fit the Innovation Equation of various orders for the three models and then calculate their corresponding AIC and SIC criteria. Tables 1 to 4 report the results for selecting the order of the Innovation Equation by the AIC and SIC criteria. The numbers in the Tables indicate how many of the replications are chosen by the corresponding criterion. Except for the cases of model 2 and 3 with n = 200 by SIC criterion using the mean square error as a loss function, all the criteria are able to choose the correct models most of the time as compared with other possibilities. In both undesirable cases, the criterion prefers the model without seasonal factor.

It is clear from all the Tables that the ability of choosing the correct model is higher as the value of n increases. When comparing Tables 1 and 2 with Tables 3 and 4, it is immediately seen that the criteria based on log-likelihood choose the correct order much more than that based on the mean square error. Using the log-likelihood as a loss function, most of the cases get over 90% for choosing the right order. In addition, the criteria based on the mean square error are not quite reliable for model 2 and 3. They are less than 50% in choosing the right order.

Comparing Tables 1 and 3 with Tables 2 and 4, SIC performs better than AIC when the model parameters are less (model 1) but AIC is better when the number of parameters increases (models 2 and 3). There seems to be a tendency that SIC underestimates the order on large parameterizations.

3.3.5 Empirical example

Returning to the exchange rate of USD/JPY, to capture the conditional heteroscedasticity of the residuals detected in section 3.2.2, the suggested Innovation Equations are

Model 1

$$(1 - \delta_{1,t}B)(1 - \Delta_{1,t}B^5)a_t = e_t$$

$$\Rightarrow h_t = \beta_1 a_{t-1}^2 + \beta_3 a_{t-5}^2 + \beta_4 a_{t-6}^2 + \sigma_e^2$$

Model 2

$$(1 - \delta_{1,t}B - \delta_{2,t}B^3)(1 - \Delta_{1,t}B^5)a_t = e_t$$

$$\Rightarrow h_t = \beta_1 a_{t-1}^2 + \beta_2 a_{t-3}^2 + \beta_3 a_{t-5}^2 + \beta_4 a_{t-6}^2 + \beta_5 a_{t-8}^2 + \sigma_e^2$$

As the Observation Equation is a white noise model, the Innovation Equation is the same for selecting order of Y_t^2 or a_t^2. The top part of Table 5 shows the maximum likelihood estimates for model 1 and 2 and at the bottom part of the Table we provide the criteria for the corresponding model. Both AIC_A and SIC_A of the log-likelihood favor model 2 and AIC_B using mean square error prefers model 1 while SIC_B prefers model 2. Hence, we choose model 2 for modeling the exchange rates.

Table 5. The AIC and SIC of the specified order for the Innovation Equation

	Model 1	Model 2
σ_e^2	0.34853	0.29284
β_1	0.11602	0.10714
β_2	--	0.09754
β_3	0.09696	0.09494
β_4	0.07710	0.07921
β_5	--	0.03170
AIC_A	559.23708	**523.43274**
SIC_A	582.31798	**558.05408**
AIC_B	0.23161	**0.23126**
SIC_B	**0.24136**	0.24587

3.4 ESTIMATING THE FULL MODEL

3.4.1 Least squares estimation

The first step is to estimate the parameters in the Observation Equation by the least squares estimation and calculate the residuals a_t. The second step is to estimate the parameters in the Innovation Equation by using a_t from the first step.

Then the least squares estimates of the parameters in the Innovation Equation are

$$\hat{\beta} = \{\sum_t (W_t - \overline{W})(W_t - \overline{W})^T\}^{-1}\{\sum_t (W_t - \overline{W})a_t^2\} ,$$

and $\hat{\sigma}_e^2 = (n-k)^{-1}\sum_t a_t^2 - \hat{\beta}^T\overline{W}$,

where t sums from k+1 to n,

$k = \max(r+Rd, s+Sd)$,

and $\overline{W}$ is the sample mean of W_t.

Recall that $W_t = (X_{1,t-1}^T, X_{2,t-1}^T)^T$, $X_{1,t-1} = \text{vech}(A_{t-1}A_{t-1}^T)$, and $X_{2,t-1} = \text{vech}(F_{t-1}F_{t-1}^T)$.

We have the following result:

Theorem

Let $X_t = (1, W_t^T)^T$, $\hat{\upsilon} = (\hat{\sigma}_e^2\ \hat{\beta}^T)^T$, and $\eta_t = a_t^2 - h_t$. If $E(X_tX_t^T)$ is nonsingular, then

(a) $\hat{\upsilon}$ is a strongly consistent estimate of υ if $E(Y_t^4) < \infty$.

(b) $n^{1/2}(\hat{\upsilon} - \upsilon) \rightarrow_d N(0, C)$ if $E(Y_t^8) < \infty$ where $\rightarrow_d$ denotes convergence in distribution and $C = E(X_tX_t^T)^{-1}E(X_t\eta_t^2X_t^T)E(X_tX_t^T)^{-1}$.

It can be proved along the same lines as Tsay (1987, p.595, theorem 4).

3.4.2 Maximum likelihood estimation

Expanding equation (2.2) and re-define the coefficients of Y_t, and a_t as ζ_i, and ξ_i, respectively, then

$$(Y_t - \mu) - \sum_{i=1}^{p+Pd} \zeta_i (Y_{t-i} - \mu) = a_t - \sum_{i=1}^{q+Qd} \xi_i a_{t-i} ,$$

whence we put

$$a_t = (Y_t - \mu) - \sum_{i=1}^{p+Pd} \zeta_i (Y_{t-i} - \mu) + \sum_{i=1}^{q+Qd} \xi_i a_{t-i} ,$$

into the conditional log-likelihood function.

The conditional log-likelihood function of $\{a_t\}$ becomes

$$L_n = -\frac{1}{2}\sum_{t=k}^{n}[\ln(h_t) + \frac{(Y_t - \mu - b_t)^2}{h_t}], \tag{3.6}$$

where $k = \max(r+Rd, s+Sd)$,

$h_t = W_t^T\beta + \sigma_e^2$,

$b_t = \sum_{i=1}^{p+Pd} \zeta_i (Y_{t-i} - \mu) - \sum_{i=1}^{q+Qd} \xi_i a_{t-i}$,

and $\upsilon = (\zeta_i, \xi_i, \sigma_e^2, \beta^T)$.

We use the maximum likelihood estimation algorithm proposed in Mak (1993) and Mak *et al.* (1997) to estimate υ. As shown in Mak *et al.* (1997), this algorithm's convergence is quite robust to the choice of initial values and is faster in reaching convergence as compared with the BHHH (Berndt *et al.* (1974)) algorithm. Mak's algorithm is as follows: the maximum likelihood estimate $\hat{\upsilon}$ of υ is obtained from solving

$$f(y,\upsilon) = \frac{\partial L_n}{\partial \upsilon} = -\frac{1}{2}\sum \frac{\partial h_t}{\partial \upsilon}\{\frac{1}{h_t} - \frac{(y_t - \mu - b_t)^2}{h_t^2}\} + \sum \frac{\partial b_t}{\partial \upsilon}\frac{(y_t - \mu - b_t)}{h_t} = 0.$$

Let $g(\tilde{\upsilon},\upsilon) = E_y[f(y,\upsilon)|\tilde{\upsilon}]$ and $\upsilon_{(r)}$ be the rth iteration estimates, then

$g(\upsilon_{(r+1)}, \upsilon_{(r)}) = f(y, \upsilon_{(r)})$,

and $\upsilon_{(r)} \rightarrow \hat{\upsilon}$ (the maximum likelihood estimate), as $r \rightarrow \infty$.

We then calculate

$$\frac{\partial g(\tilde{\upsilon},\upsilon)}{\partial \tilde{\upsilon}}\Big|_{\tilde{\upsilon}=\upsilon} = \frac{1}{2}\sum \frac{1}{h_t^2}\left(\frac{\partial h_t}{\partial \upsilon}\right)\left(\frac{\partial h_t}{\partial \upsilon}\right)^T + \sum \frac{1}{h_t}\left(\frac{\partial b_t}{\partial \upsilon}\right)\left(\frac{\partial b_t}{\partial \upsilon}\right)^T.$$

Finally, the iteration equation is

$$\upsilon_{(r+1)} = \upsilon_{(r)} + \left\{ \left[\frac{\partial g(\tilde{\upsilon}, \upsilon)}{\partial \tilde{\upsilon}} \Big|_{\tilde{\upsilon}=\upsilon} \right]^{T} \right\}^{-1}_{\upsilon=\upsilon_{(r)}} f(y, \upsilon_{(r)})$$

Also it is worthwhile to point out that the Fisher's Information matrix can be obtained as,

$$I = \frac{\partial g(\tilde{\upsilon}, \upsilon)}{\partial \tilde{\upsilon}} \Big|_{\tilde{\upsilon}=\upsilon}$$

and the covariance matrix of υ is given by I^{-1}.

3.4.3 Simulation

In this section, the reliability of the least squares estimation and the maximum likelihood estimation are investigated and compared by simulations. The following model is considered,

$$(1 - \phi B)Y_t = a_t$$

$$a_t = [(1 + \omega_{1,t}B)(1 + \Omega_{1,t}B^4) - 1]Y_t + e_t$$

where $\phi = 0.4$, $\sigma_e^2 = 0.01$,

$$\text{var}(a_t) = h_t = \beta_1 Y_{t-1}^2 + \beta_2 Y_{t-4}^2 + \beta_3 Y_{t-5}^2,$$

$$\text{var}(\omega_{1,t}) = \beta_1 = 0.4356,$$

$$\text{var}(\Omega_{1,t}) = \beta_2 = 0.36,$$

and $\text{var}(\omega_{1,t}\Omega_{1,t}) = \text{var}(\omega_{1,t})\ \text{var}(\Omega_{1,t}) = \beta_3 = 0.156816.$

The parameter of interest is $\upsilon = (\phi, \sigma_e^2, \beta_1, \beta_2, \beta_3)^T$. We consider various length of realization with n = 100, 200 and 500. For each case, there are 500 independent replications. The initial value υ_0 is $(0.1, 0.005, 0.1, 0.1, 0.1)^T$ for all cases. The average estimate of ϕ, σ_e^2, β_1, β_2 and β_3 and their average standard errors (in parentheses) are shown in Table 6. The conclusion of the simulations is obvious. For each n, the maximum likelihood estimates are more accurate than the least squares estimation and are much closer to the true values. Also, it can be seen that the estimates for the maximum likelihood method generally improve as n becomes larger. Hence, we will use the maximum likelihood estimation to fit the model. Since the initial values affect the existence and consistency of interested parameters in iterative numerical methods, the least squares estimates will be used as the initial values of the maximum likelihood method.

Table 6. The average estimates of ϕ, σ_e^2, β_1, β_2, β_3 and the corresponding average standard errors

n	*estimation*	ϕ	σ_e^2	β_1	β_2	β_3
100	Least squares	0.31531 (0.09618)	0.05789 (0.03973)	0.16995 (0.08775)	0.18020 (0.09254)	0.13386 (0.09547)
	Maximum likelihood	0.39121 (0.11591)	0.01194 (0.00502)	0.35734 (0.15285)	0.31576 (0.14985)	0.20414 (0.13364)
200	Least squares	0.33433 (0.06691)	0.07142 (0.04355)	0.18980 (0.05963)	0.19930 (0.06246)	0.12022 (0.06464)
	Maximum likelihood	0.39272 (0.08300)	0.01089 (0.00326)	0.39500 (0.11252)	0.33156 (0.10690)	0.16736 (0.08565)
500	Least squares	0.35034 (0.04184)	0.09064 (0.05330)	0.21456 (0.03613)	0.20070 (0.03734)	0.11549 (0.03858)
	Maximum likelihood	0.39500 (0.05323)	0.01053 (0.00202)	0.42325 (0.07367)	0.35224 (0.06877)	0.15735 (0.05228)

3.4.4 Empirical example

Since the Observation Equation is a white noise model, we do not need to estimate the full model again. The final model is

$$Y_t = a_t$$

$$(1 - \delta_{1,t}B - \delta_{2,t}B^3)(1 - \Delta_{1,t}B^5)a_t = e_t$$

$$\Rightarrow h_t = \beta_1 a_{t-1}^2 + \beta_2 a_{t-3}^2 + \beta_3 a_{t-5}^2 + \beta_4 a_{t-6}^2 + \beta_5 a_{t-8}^2 + \sigma_e^2$$

It is found that $\hat{\sigma}_e^2 = 0.2928$, $\beta_1 = 0.1071(0.0240)$, $\beta_2 = 0.0975(0.0236)$, $\beta_3 = 0.0949(0.0235)$, $\beta_4 = 0.0792(0.0226)$, and $\beta_5 = 0.0317(0.0174)$.

The coefficients with standard error in parentheses are estimated by the maximum likelihood method. All the t-ratios of the variances are significantly different from zero. The large coefficient for a_{t-5}^2, the seasonal lag, has sound financial interpretation as discussed in previous sections. It supports the presence of weekend effect in the Yen exchange rates.

3.5 MODEL DIAGNOSTIC CHECKING

Li and Mak (1994) proposed a general class of squared residual autocorrelations and derived their asymptotic distribution that can be used for diagnostic checking general conditional heteroscedastic models.

Let $\hat{\varepsilon}_t$ be the estimated residual from the Observation Equation, $\hat{h}_t$ be the estimated conditional variance from the Innovation Equation. Then, the lag-k-squared standardized residual autocorrelation is defined as

$$\tilde{r}_k = \frac{\sum_{t=k+1}^{n}(\hat{a}_t^2/\hat{h}_t - \bar{a})(\hat{a}_{t-k}^2/\hat{h}_{t-k} - \bar{a})}{\sum_{t=1}^{n}(\hat{a}_t^2/\hat{h}_t - \bar{a})^2}, \quad k = 1, 2, \ldots$$

where $\bar{a} = n^{-1}\sum \frac{\hat{a}_t^2}{\hat{h}_t}$.

If the model is correct, it may be shown that $\bar{a}$ converges in probability to 1. Thus, $\tilde{r}_k$ can be replaced by

$$\hat{r}_k = \frac{\sum_{t=k+1}^{n}(\hat{a}_t^2/\hat{h}_t - 1)(\hat{a}_{t-k}^2/\hat{h}_{t-k} - 1)}{\sum_{t=1}^{n}(\hat{a}_t^2/\hat{h}_t - 1)^2}, \quad k = 1, 2, \ldots .$$

Also, the lag-k squared standardized residual autocovariance is defined as

$$\hat{C}_k = \frac{1}{n}\sum(\frac{\hat{a}_t^2}{\hat{h}_t} - 1)(\frac{\hat{a}_{t-k}^2}{\hat{h}_{t-k}} - 1),$$

and an overall test statistics is

$$Q(M) = n\hat{r}^T\hat{V}^{-1}\hat{r} \sim \chi^2(M).$$

Here $\hat{r}^T = (\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_M)$, $\hat{V} = I_M - \frac{1}{4}XG^{-1}X^T$, and I_M = order M identity matrix. Furthermore, $G = -E(n^{-1}\frac{\partial^2 L_n}{\partial\upsilon\partial\upsilon^T})$, L_n is the log-likelihood as defined in (3.6), and X is a matrix of the partial derivatives of the residual autocovariances $\hat{C}_k$ with respect to υ.

For ARCH(q) model,

$$Q(M) = n\sum_{i=r+1}^{M}\hat{r}_i^2 \sim \chi^2(M\text{-}r),$$

since $E[\frac{1}{h_t}(\frac{\partial h_t}{\partial\upsilon})(\frac{a_{t-k}^2}{h_{t-k}} - 1)] = 0$, if $k > r$.

3.5.1 Empirical example

Table 7 gives the first ten autocorrelation coefficients for the squared standardized residuals. They are within the standard error bound ($\approx \pm 2/\sqrt{n} = \pm 0.04109$). Q(5) in Li-Mak test is 1.58761, which is insignifiant at the 5% level. There is no heteroscedasticity in the squared standardized residuals. The autocorrelation coefficients at lag 1, 3 and 5 are comparatively smaller after considering the Innovation Equation (Table 8). There are good improvements if the

Innovation Equation is fitted, i.e., the model successfully captured the seasonal non-linearities.

Table 7. The autocorrelation coefficients of squared standardized residuals up to order 10

lag	1	2	3	4	5	6	7	8	9	10
acf	0.0025	0.0181	-0.0072	0.0157	-0.0062	-0.0142	-0.0182	-0.0153	0.0126	0.0094

Table 8. The comparison of acf before and after the incorporation of Innovation Equation

	lag 1	lag 3	lag 5
before fitting the Innovation Equation	0.097	0.057	0.068
after fitting the Innovation Equation	0.0025	-0.0072	-0.0062

3.6 SUMMARY OF MODEL BUILDING PROCEDURES

We summarize the procedures for building a seasonal CHARMA model as follows:

Step 1 Start with the traditional ARMA models. Identify the order $(p,q)\times(P,Q)_d$ for the Observation Equation. If the series is non-stationary, transform to stationarity by standard techniques like the Box-Cox transformation and differencing.

Step 2 Estimate the parameters ϕ_i's, θ_j's, Φ_k's and Θ_l's by the conditional least squares or conditional maximum likelihood method and obtain the residuals $\hat{a}_t$.

Step 3 Perform diagnostic checking using say, the Ljung and Box test. If any inadequacies are detected, go back to step 1 and repeat the procedure again until the model passes all the checks.

Step 4 Examine autocorrelation function of squared residuals for varying conditional variance and perform the test described in section 3.2. If the test fails to indicate any model discrepancy, the model obtained in step 3 is regarded as adequate. Otherwise, tentatively specify the order r, R, s and S of the Innovation Equation by treating a_t as observable rather than as residuals from the linear model.

Step 5 Estimate all of the parameters simultaneously in the observation and Innovation Equations by using the conditional maximum likelihood method.

Step 6 Perform diagnostic checking for the full model. Go back to step 4 if necessary.

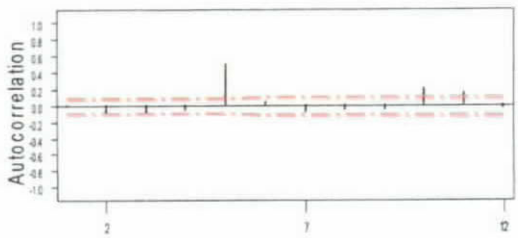
APPENDIX TO CHAPTER 3

Appendix 1 – Model a

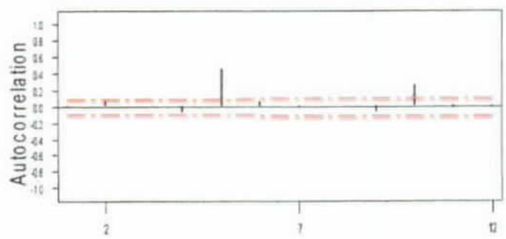
$$h_t = \beta_0 + \beta_1 y_{t-5}^2$$

Note: a_j^2 denotes squared residuals of the jth set data, j = 1, …, 20

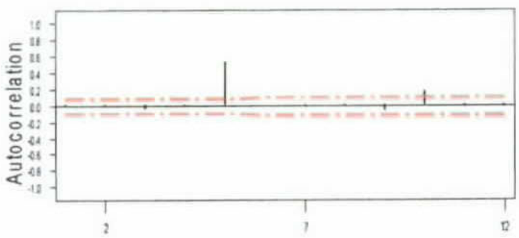
Autocorrelation Function for a_1^2



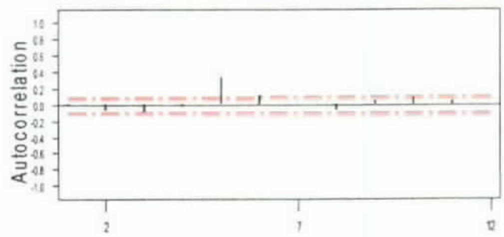
Autocorrelation Function for a_2^2



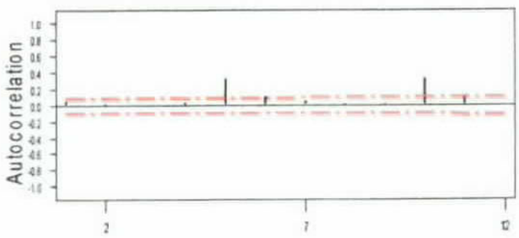
Autocorrelation Function for a_3^2



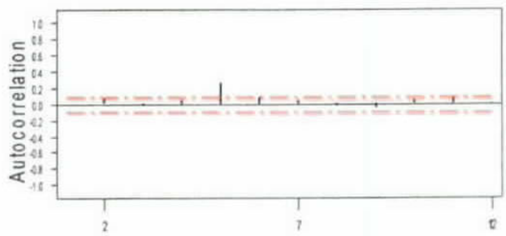
Autocorrelation Function for a_4^2



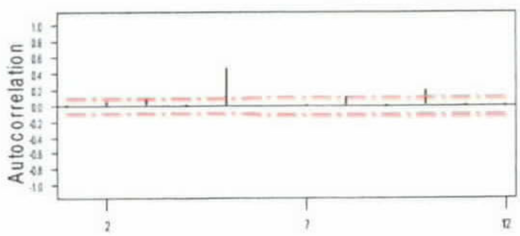
Autocorrelation Function for a_5^2



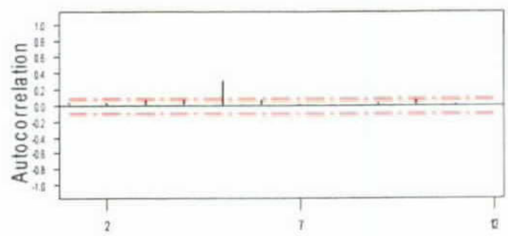
Autocorrelation Function for a_6^2



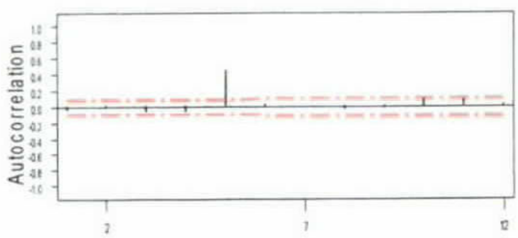
Autocorrelation Function for a_7^2



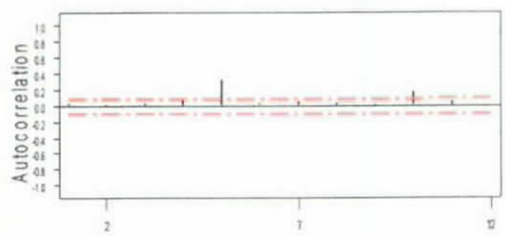
Autocorrelation Function for a_8^2



Autocorrelation Function for a_9^2

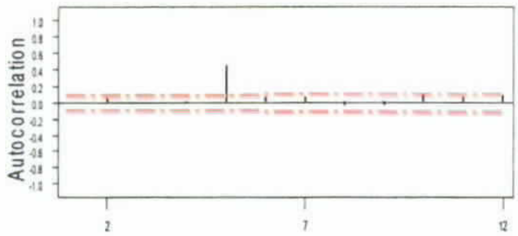


Autocorrelation Function for a_10^2

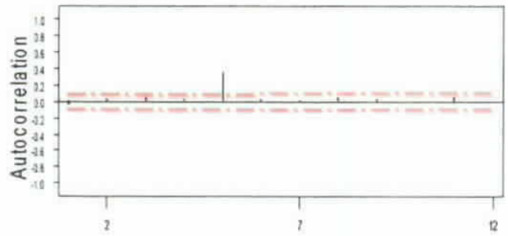


Appendix 1 – Model a (cont.)

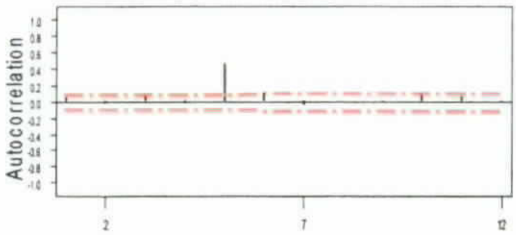
Autocorrelation Function for a_11^2



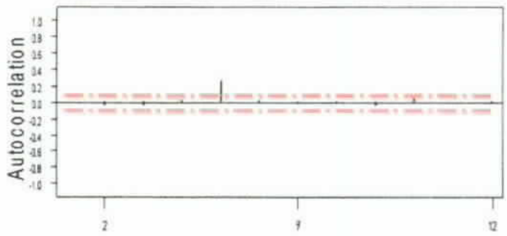
Autocorrelation Function for a_12^2



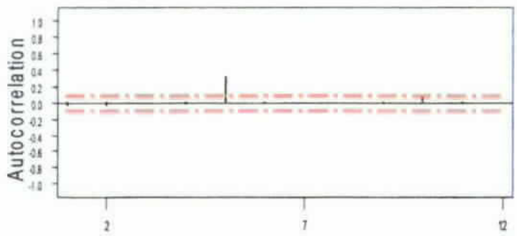
Autocorrelation Function for a_13^2



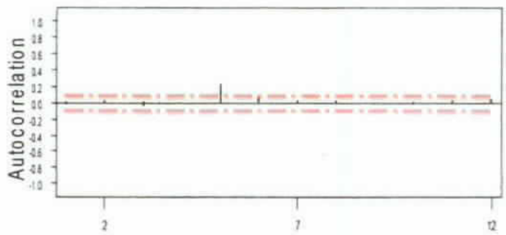
Autocorrelation Function for a_14^2



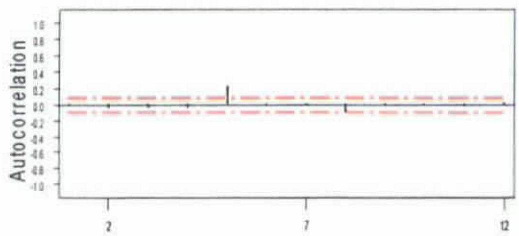
Autocorrelation Function for a_15^2



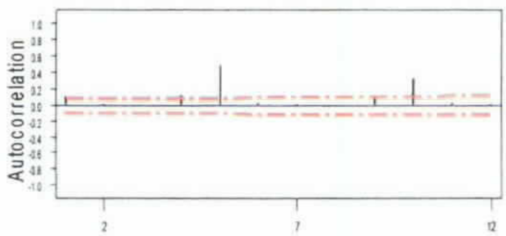
Autocorrelation Function for a_16^2



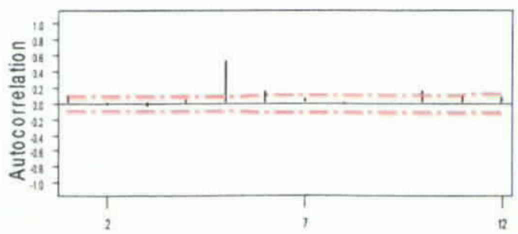
Autocorrelation Function for a_17^2



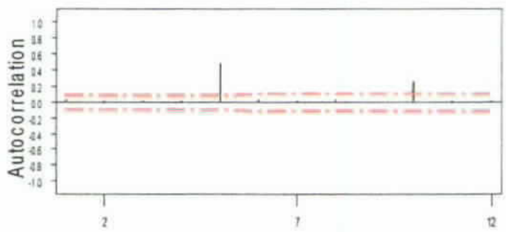
Autocorrelation Function for a_18^2



Autocorrelation Function for a_19^2



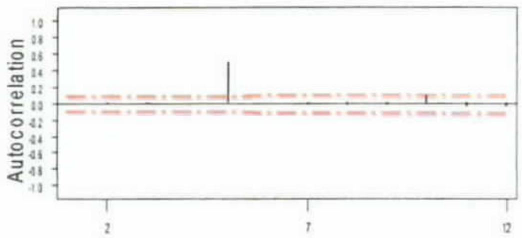
Autocorrelation Function for a_20^2



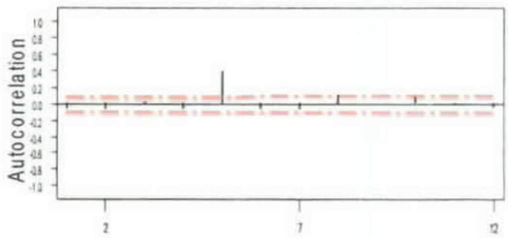
Appendix 2 – Model b

$$h_t = \beta_0 + \beta_1 a_{t-5}^2$$

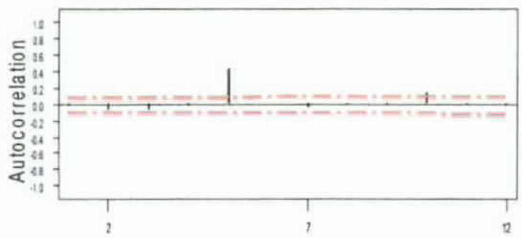
Autocorrelation Function for a_1^2



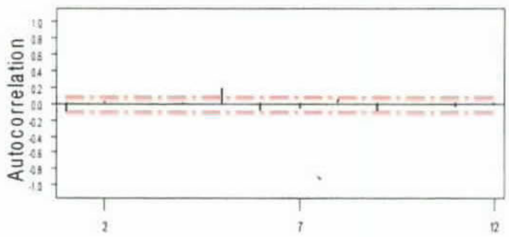
Autocorrelation Function for a_2^2



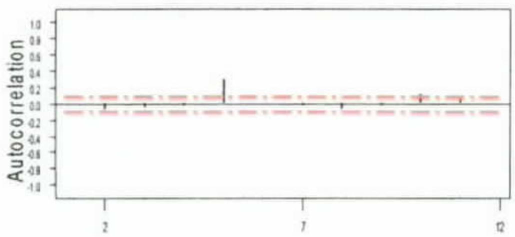
Autocorrelation Function for a_3^2



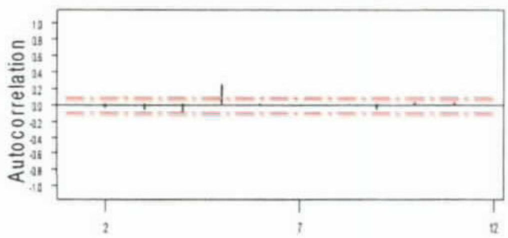
Autocorrelation Function for a_4^2



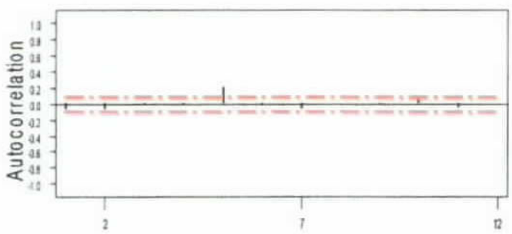
Autocorrelation Function for a_5^2



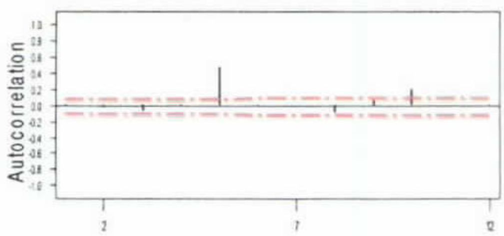
Autocorrelation Function for a_6^2



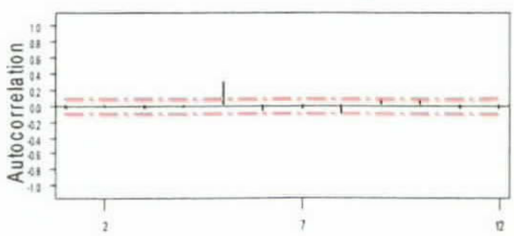
Autocorrelation Function for a_7^2



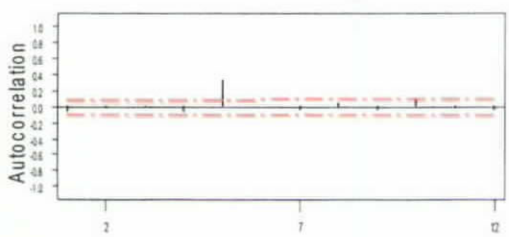
Autocorrelation Function for a_8^2



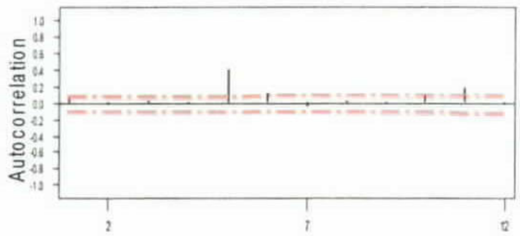
Autocorrelation Function for a_9^2



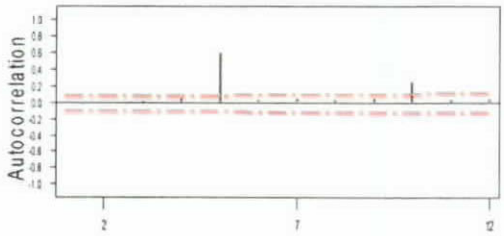
Autocorrelation Function for a_10^2



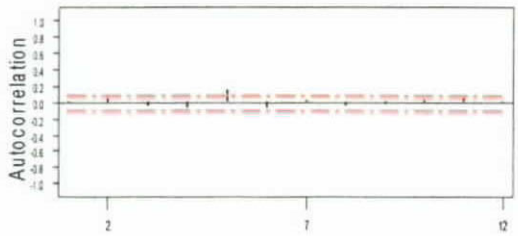
Autocorrelation Function for a_11^2



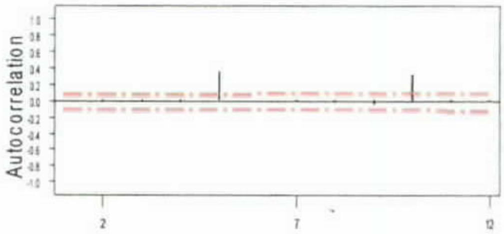
Autocorrelation Function for a_12^2



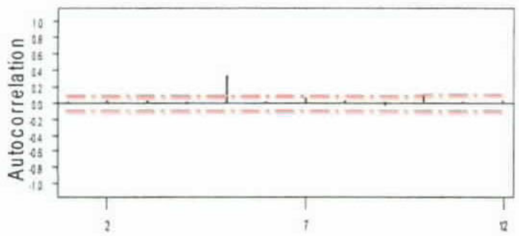
Autocorrelation Function for a_13^2



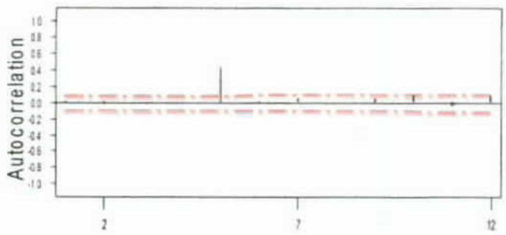
Autocorrelation Function for a_14^2



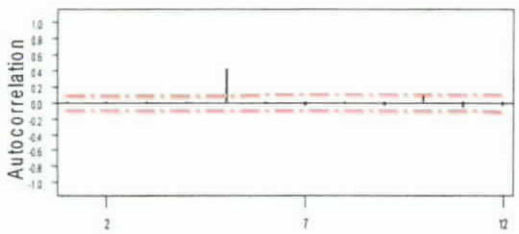
Autocorrelation Function for a_15^2



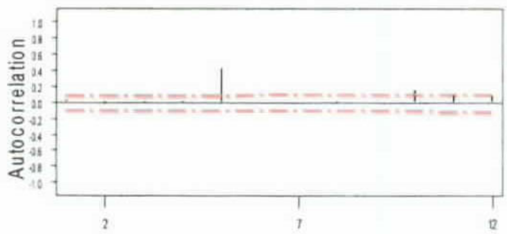
Autocorrelation Function for a_16^2



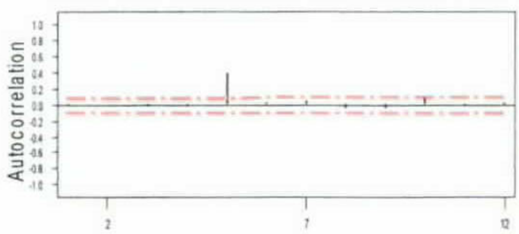
Autocorrelation Function for a_17^2



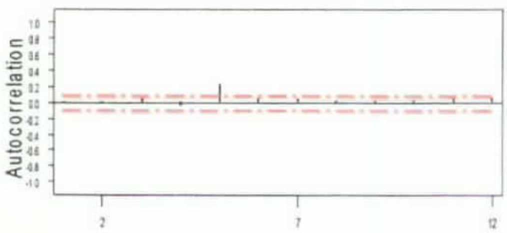
Autocorrelation Function for a_18^2



Autocorrelation Function for a_19^2



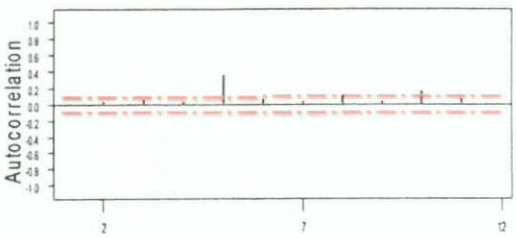
Autocorrelation Function for a_20^2



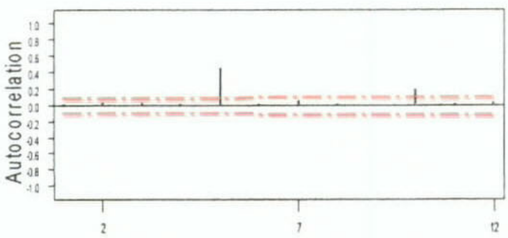
Appendix 3 – Model c

$$h_t = \beta_0 + \beta_1 y_{t-5}^2 + \beta_2 a_{t-5}^2$$

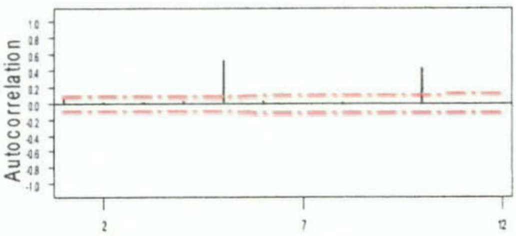
Autocorrelation Function for a_1^2



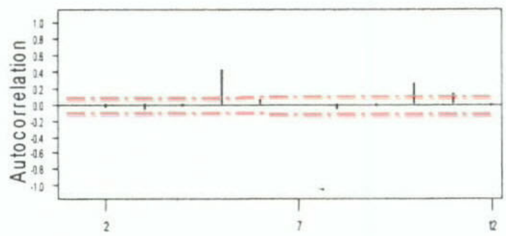
Autocorrelation Function for a_2^2



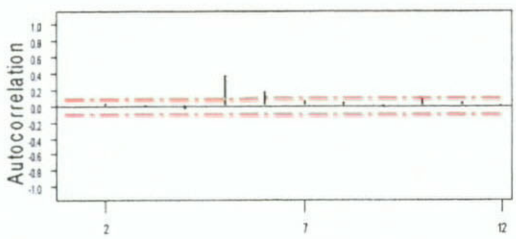
Autocorrelation Function for a_3^2



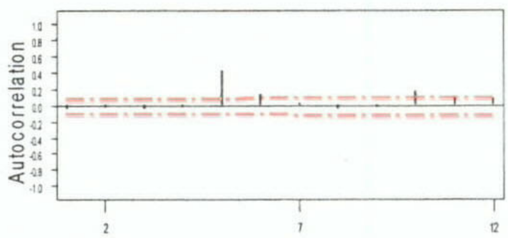
Autocorrelation Function for a_4^2



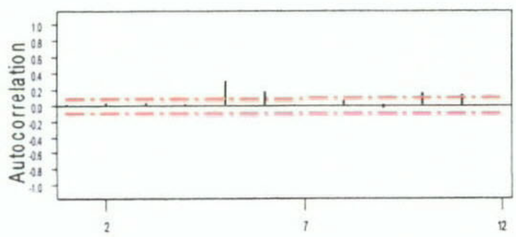
Autocorrelation Function for a_5^2



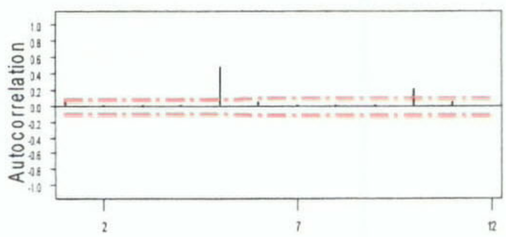
Autocorrelation Function for a_6^2



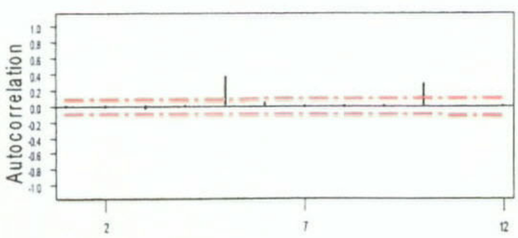
Autocorrelation Function for a_7^2



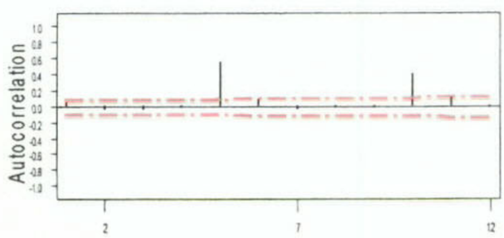
Autocorrelation Function for a_8^2



Autocorrelation Function for a_9^2

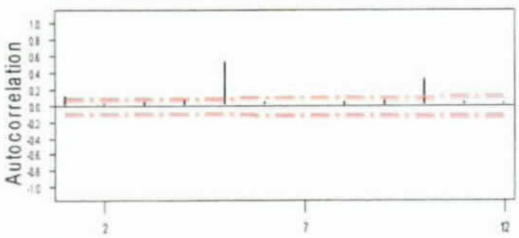


Autocorrelation Function for a_10^2

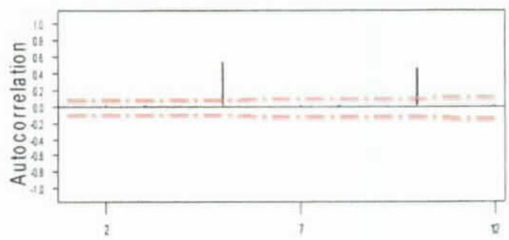


Appendix 3 – Model c (cont.)

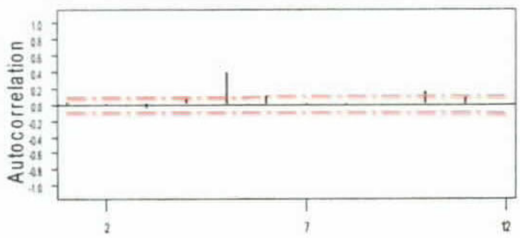
Autocorrelation Function for a_11^2



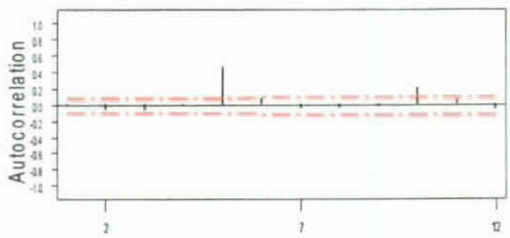
Autocorrelation Function for a_12^2



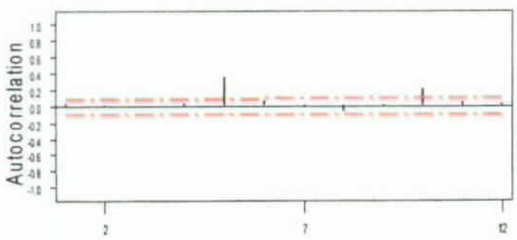
Autocorrelation Function for a_13^2



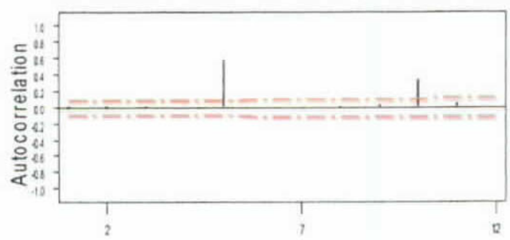
Autocorrelation Function for a_14^2



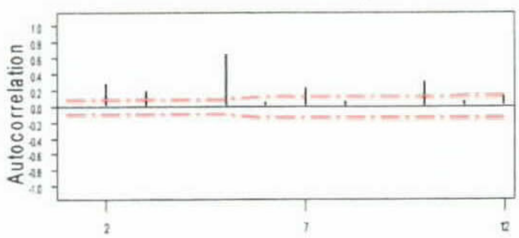
Autocorrelation Function for a_15^2



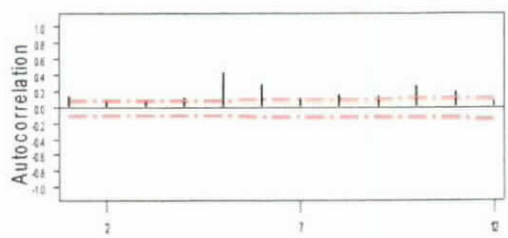
Autocorrelation Function for a_16^2



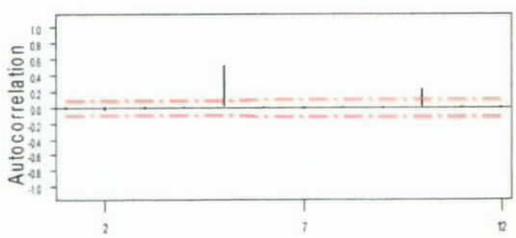
Autocorrelation Function for a_17^2



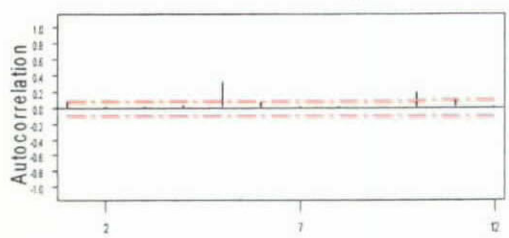
Autocorrelation Function for a_18^2



Autocorrelation Function for a_19^2



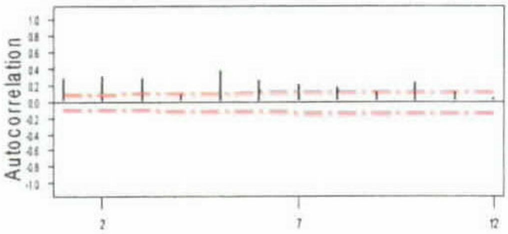
Autocorrelation Function for a_20^2



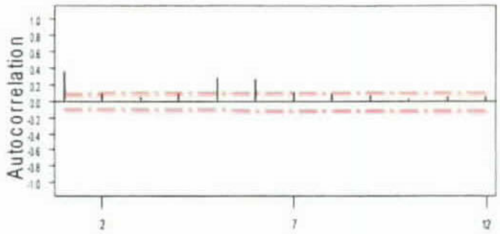
Appendix 4 – Model d

$$h_t = \beta_0 + \beta_1 y_{t-1}^2 + \beta_2 y_{t-5}^2 + \beta_3 y_{t-6}^2$$

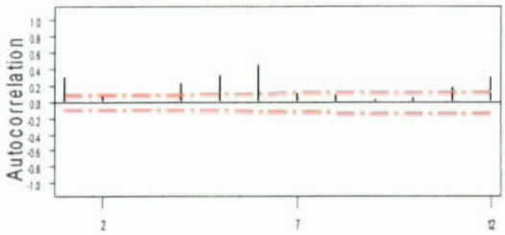
Autocorrelation Function for a_1^2



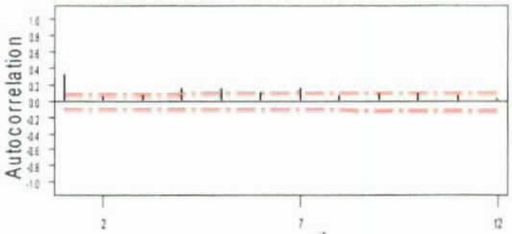
Autocorrelation Function for a_2^2



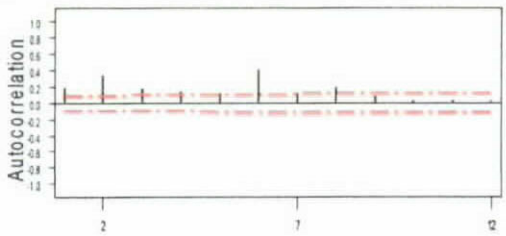
Autocorrelation Function for a_3^2



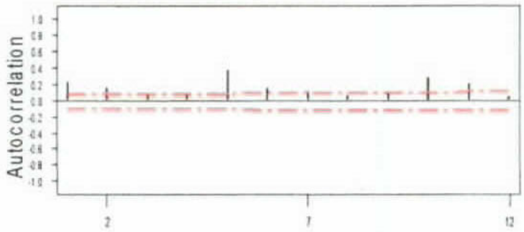
Autocorrelation Function for a_4^2



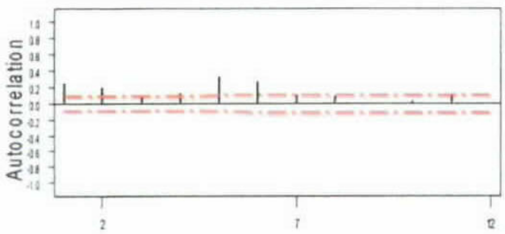
Autocorrelation Function for a_5^2



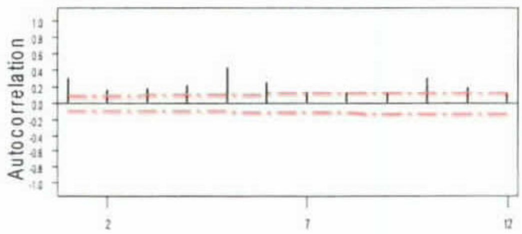
Autocorrelation Function for a_6^2



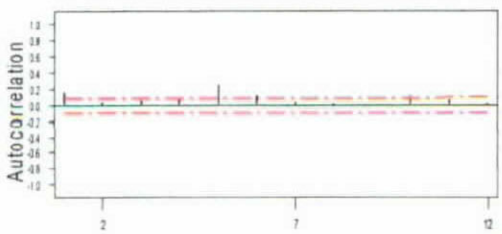
Autocorrelation Function for a_7^2



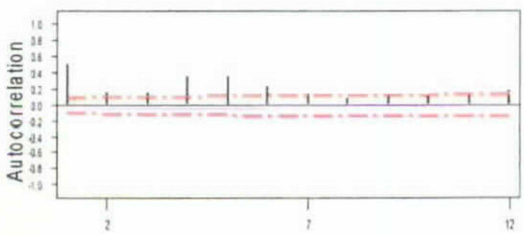
Autocorrelation Function for a_8^2



Autocorrelation Function for a_9^2

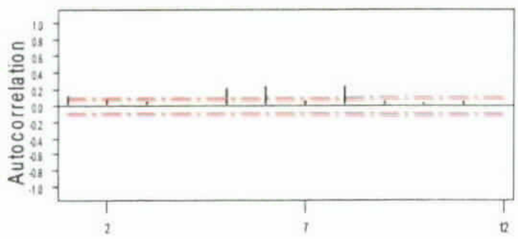


Autocorrelation Function for a_10^2

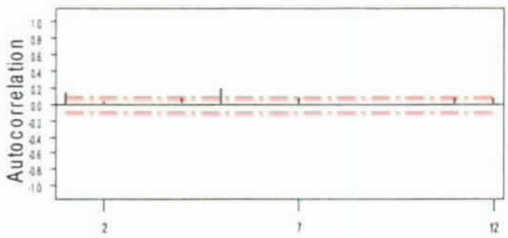


Appendix 4 – Model d (cont.)

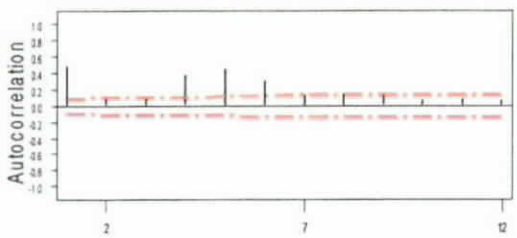
Autocorrelation Function for a_11^2



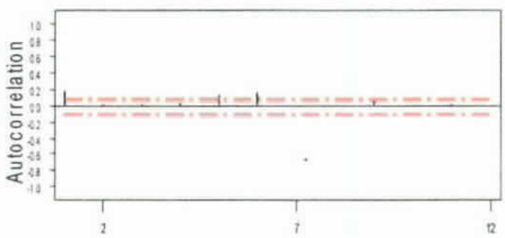
Autocorrelation Function for a_12^2



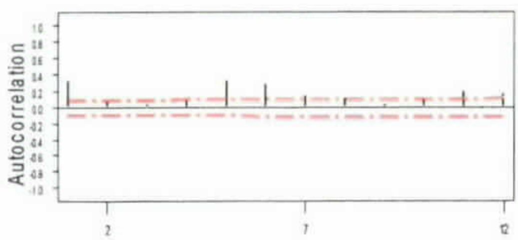
Autocorrelation Function for a_13^2



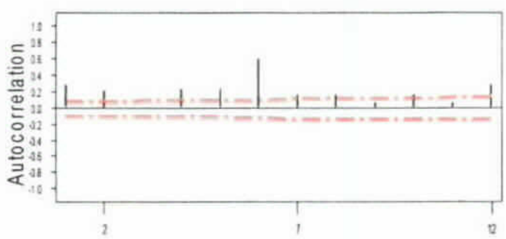
Autocorrelation Function for a_14^2



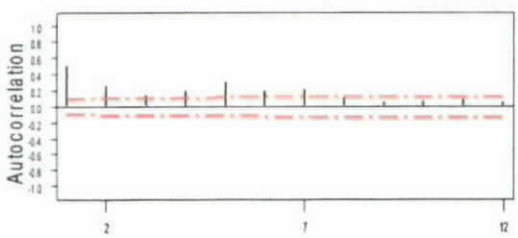
Autocorrelation Function for a_15^2



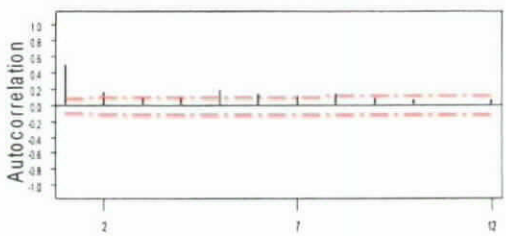
Autocorrelation Function for a_16^2



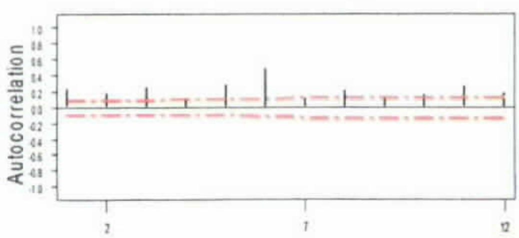
Autocorrelation Function for a_17^2



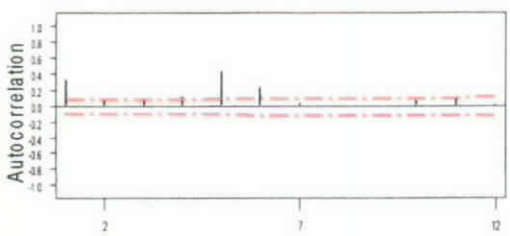
Autocorrelation Function for a_18^2



Autocorrelation Function for a_19^2



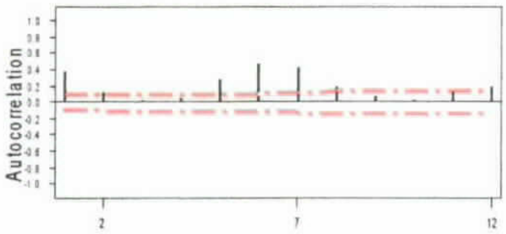
Autocorrelation Function for a_20^2



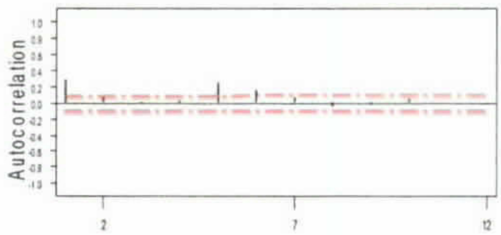
Appendix 5 – Model e

$$h_t = \beta_0 + \beta_1 a_{t-1}^2 + \beta_2 a_{t-5}^2 + \beta_3 a_{t-6}^2$$

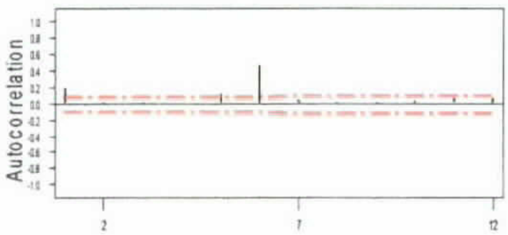
Autocorrelation Function for a_1^2



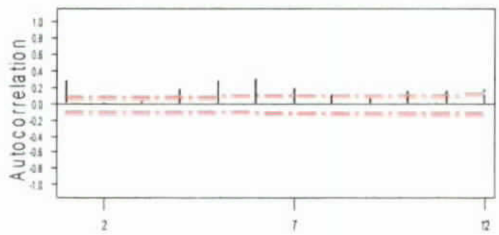
Autocorrelation Function for a_2^2



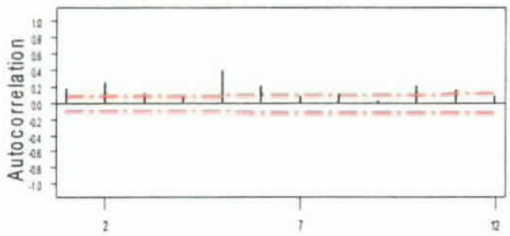
Autocorrelation Function for a_3^2



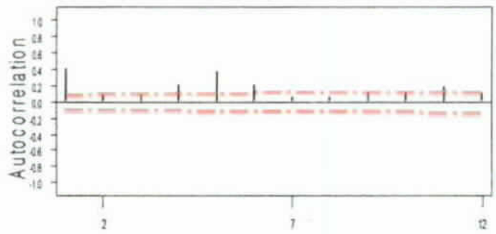
Autocorrelation Function for a_5^2



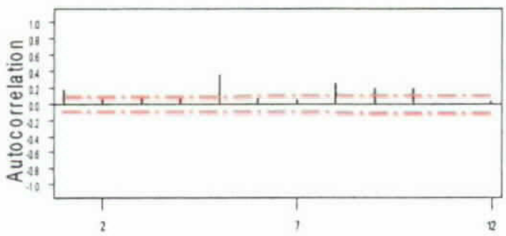
Autocorrelation Function for a_4^2



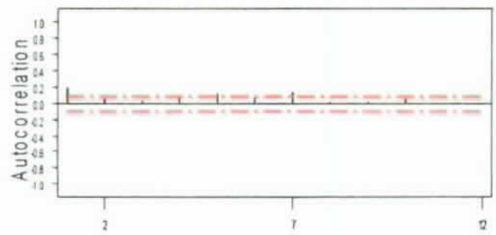
Autocorrelation Function for a_6^2



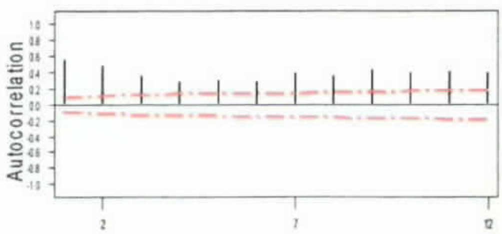
Autocorrelation Function for a_7^2



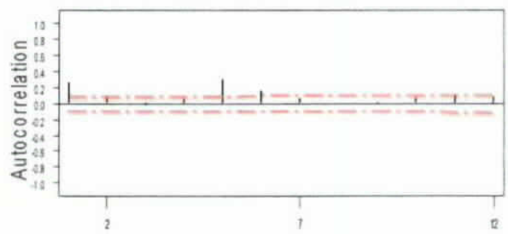
Autocorrelation Function for a_8^2



Autocorrelation Function for a_9^2

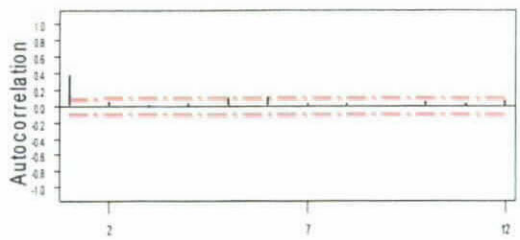


Autocorrelation Function for a_10^2

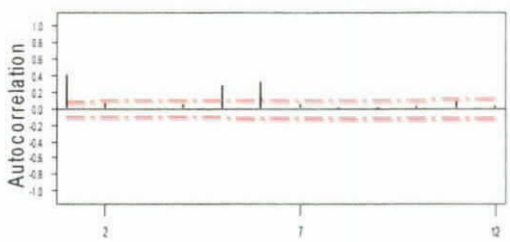


Appendix 5 – Model e (cont.)

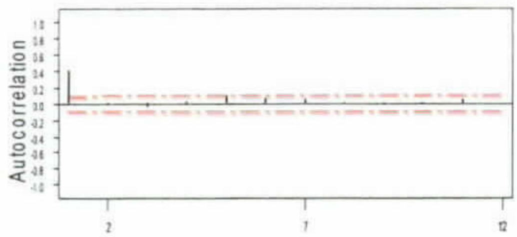
Autocorrelation Function for a_11^2



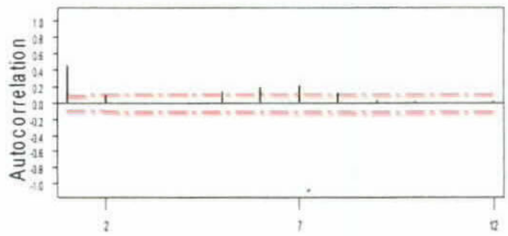
Autocorrelation Function for a_12^2



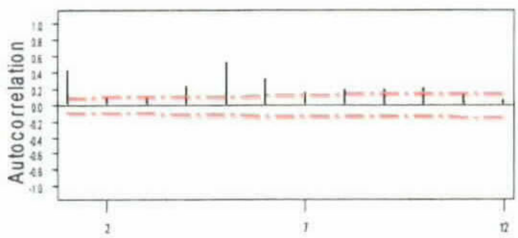
Autocorrelation Function for a_13^2



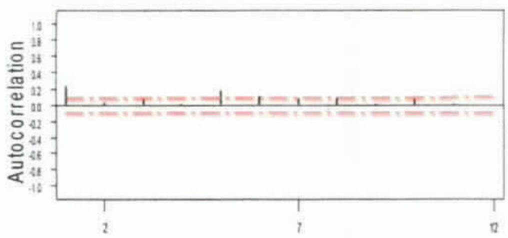
Autocorrelation Function for a_14^2



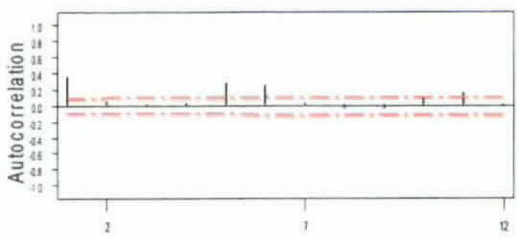
Autocorrelation Function for a_15^2



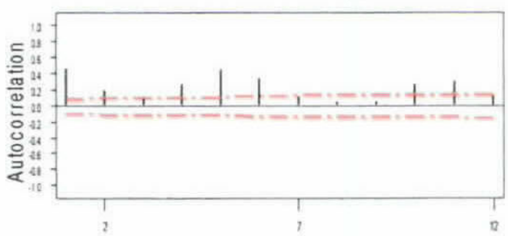
Autocorrelation Function for a_16^2



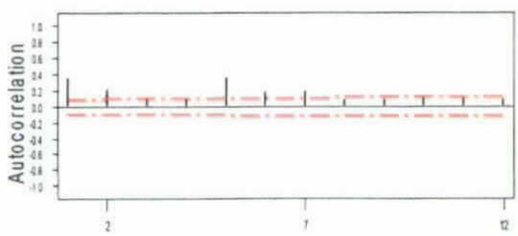
Autocorrelation Function for a_17^2



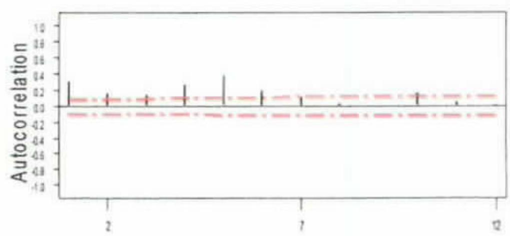
Autocorrelation Function for a_18^2



Autocorrelation Function for a_19^2



Autocorrelation Function for a_20^2



CHAPTER 4 COMPARISON WITH THE GARCH FAMILY

4.1 GARCH MODEL

The ARCH model is very popular and useful in modeling economic and financial data after it is introduced. However, in many applications, the ARCH(q) model calls for a long lag q, thus leading to a large number of parameters to be estimated. Therefore, Bollerslev (1986) extended the ARCH model to the generalized ARCH (GARCH) model in order to provide a more flexible lag structure. The conditional variance of the ARCH model is a linear function of past squared residuals, whereas the GARCH model includes both past squared residuals and lagged conditional variances.

Let $\{y_t\}$ be a time series, a_t denote the innovations obtained from the fitted model, ψ_{t-1} the information set available at time t-1 and L denotes the lag operator. The GARCH(p,q) model is defined as

$$a_t = Y_t - X_t^T b,$$

$$a_t \mid \psi_{t-1} \sim N(0, h_t), \tag{4.1}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{4.2}$$

$$= \alpha_0 + A(L)a_t^2 + B(L)h_t \tag{4.3}$$

where X_t^T be a vector of explanatory variables,

$p \geq 0, q > 0,$

$\alpha_0 > 0, \alpha_i \geq 0, \qquad i = 1, \ldots, q,$

and $\beta_j \geq 0, \qquad j = 1, \ldots, p.$

A(L) and B(L) are polynomials in L and have no common roots. When p = 0, the model reduces to the ARCH(q) model. From (4.3), if all the roots of the polynomial B(x) = 1 lie outside the unit circle, the conditional variance of the GARCH(p,q) model can be represented by past a_t^2 only. Namely,

$$\begin{aligned} h_t &= \alpha_0(1 - B(1))^{-1} + A(L)(1 - B(L))^{-1} a_t^2 \\ &= \alpha_0(1 - \sum_{i=1}^{p} \beta_i)^{-1} + \sum_{i=1}^{\infty} \delta_i a_{t-i}^2 \end{aligned},$$

where $\delta_i = \begin{cases} \alpha_i + \sum_{j=1}^{n} \beta_j \delta_{t-j}, & i = 1, \ldots, q \\ \sum_{j=1}^{n} \beta_j \delta_{t-j}, & i = q+1, \ldots \end{cases}$.

The mean and variance of time series with the GARCH(p,q) model are 0 and $\alpha_0(1 - A(1) - B(1))^{-1}$, respectively. It is stationary if and only if A(1) + B(1) < 1. For example, the stationary condition for GARCH(1,1), the simplest but is found to suffice in many applications, is $\alpha_1 + \beta_1 < 1$.

Rearranging terms in (4.2) as

$$\begin{aligned} a_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j a_{t-j}^2 - \sum_{j=1}^{p} \beta_j v_{t-j} + v_t \\ &= \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \beta_i) a_{t-i}^2 - \sum_{j=1}^{p} \beta_j v_{t-j} + v_t \end{aligned}$$

where $m = \max\{p,q\}$,

and v_t is the serially uncorrelated innovation sequence $\{a_t^2 - h_t\}$.

Thus GARCH(p,q) model can be interpreted as an ARMA(m,p) model for a_t^2 with autoregressive parameters ($\alpha_i + \beta_i$), and moving average parameters $-\beta_j$. From Bollerslev (1988), the Yule-Walker equations for a_t^2 are

$$\rho_n = \sum_{i=1}^{m}(\alpha_i + \beta_i)\rho_{n-i} \text{ , } n > p.$$

The autocorrelations for a_t^2 are

$$\rho_1 = \alpha_1(1-\alpha_1\beta_1-\beta_1^2)/(1-2\alpha_1\beta_1-\beta_1^2),$$

$$\rho_n = (\alpha_1+\beta_1)^{n-1}\rho_1, n > 1.$$

GARCH model is reported to be highly successful in financial modeling, see Bollerslev *et al.* (1992).

4.1.1 Maximum likelihood estimation of the GARCH model

Let $z_t^T = (1,\ a_{t-1}^2,\ \ldots,\ a_{t-q}^2,\ h_{t-1},\ \ldots,\ h_{t-p})$, $\omega^T = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)$, and $\upsilon = (b^T, \omega^T)$. Then (4.2) can be written as

$$h_t = z_t^T\omega .$$

The log-likelihood function of $\{a_t\}$ is

$$L_t = \sum_{t=1}^{n} l_t$$

$$l_t = -\frac{1}{2}\ln(h_t) - \frac{1}{2}(\frac{a_t^2}{h_t})$$

Then,

$$\frac{\partial l_t}{\partial\omega} = \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial\omega}\left(\frac{a_t^2}{h_t}-1\right),$$

and $$\frac{\partial l_t}{\partial b} = \frac{a_t}{h_t}\left(\frac{-\partial a_t}{\partial b}\right) + \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b}\left(\frac{a_t^2}{h_t} - 1\right),$$

where $$\frac{\partial h_t}{\partial \omega} = z_t + \sum_{i=1}^{p}\beta_i \frac{\partial h_{t-i}}{\partial \omega},$$

and $$\frac{\partial h_t}{\partial b} = 2\sum_{i=1}^{n}\alpha_i a_{t-i}\left(\frac{\partial a_{t-i}}{\partial b}\right) + \sum_{i=1}^{n}\beta_i\left(\frac{\partial h_{t-i}}{\partial b}\right).$$

Therefore, the maximum likelihood estimates are calculated from

$$\upsilon_{(r+1)} = \upsilon_{(r)} + \left(\sum_{t=1}^{n}\frac{\partial l_t}{\partial \upsilon}\frac{\partial l_t}{\partial \upsilon^T}\right)^{-1}\left(\sum_{t=1}^{n}\frac{\partial l_t}{\partial \upsilon}\right)$$

with the information matrix $I = \sum_{t=1}^{n}\left[\left(\frac{\partial l_t}{\partial \upsilon}\right)\left(\frac{\partial l_t}{\partial \upsilon^T}\right)\right]^{-1}$.

4.2 P-GARCH MODEL

Bollerslev and Ghysels (1996) introduced the periodic generalized ARCH model (P-GARCH) whose parameters change seasonally for better characterizing the repetitive or seasonal conditional heteroscedasticity patterns. Unlike the multiplicative seasonal ARMA models described in section 2.1.5, in P-GARCH model, the periodic cycles need not be purely repetitive: the periodicity is known but the actual observations may not follow the periodic cycles strictly. For example, the purely repetitive pattern may be interrupted by holidays in the daily data.

Similar to GARCH, the class of strong P-GARCH processes is defined as

$$\tilde{a}_t|\psi_{t-1}^d \sim N(0, \tilde{h}_t) \quad (4.4)$$

and $$\tilde{h}_t = \alpha_{d(t)} + \sum_{i=1}^{q}\alpha_{id(t)}\tilde{a}_{t-i}^2 + \sum_{j=1}^{p}\beta_{jd(t)}\tilde{h}_{t-j}, \quad (4.5)$$

where d(t) is the stage of the period cycle with an upperbound D at time t, where D is the length of the cycle. If the periodic cycle is purely repetitive, then d(t) = t mod D. Note that $\widetilde{a}_t$ may be the residuals from a periodic ARMA model (P-ARMA, Tiao and Grupe (1980)) while a_t in the GARCH model may be from a multiplicative seasonal ARMA model.

Like the GARCH model, the P-GARCH model can also be interpreted as a P-ARMA representation for $\widetilde{a}_t^2$,

$$\widetilde{a}_t^2 = \alpha_{d(t)} + \sum_{i=1}^{m}(\alpha_{id(t)} + \beta_{id(t)})\widetilde{a}_{t-i}^2 - \sum_{j=1}^{p}\beta_{jd(t)}\widetilde{v}_{t-j} + \widetilde{v}_t ,$$

where m = max(p, q), and $\widetilde{v}_t = \widetilde{a}_t^2 - \widetilde{h}_t$.

The conditional log-likelihood function for P-GARCH model is similar to the GARCH model with different seasonal cycles,

$$L_T(\upsilon \mid \psi_{t-1}^d) = \sum_{t=1}^{n} l_t(\upsilon_{d(t)}),$$

where $l_t(\upsilon_{d(t)}) = -\frac{1}{2}[\ln(\widetilde{h}_t(\upsilon_{d(t)})) + \widetilde{a}_t^2(\upsilon_{d(t)}) / \widetilde{h}_t(\upsilon_{d(t)})]$.

4.3 SIMULATION

As GARCH model is so popular and successful in the literature recently, it is of interest to investigate and compare the performance between the proposed model and GARCH model. A simulation experiment is worked out to study their forecasting performance. The true model is an AR-seasonal ARCH model with ϕ = 0.4, $\alpha_0 = 0.01$, $\alpha_1 = 0.4356$, $\alpha_2 = 0.36$, $\alpha_3 = 0.156816$, i.e.,

$$Y_t - \phi Y_{t-1} = a_t$$

$$(1-\delta_t B)(1-\Delta_t B^{12})a_t = e_t$$

$$\Rightarrow \quad h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-12}^2 + \alpha_3 a_{t-13}^2$$

Table 9. The maximum likelihood estimates of the proposed model and GARCH model

		Proposed	GARCH
n = 200	ϕ	0.30786 (0.06565)	0.39687 (0.06787)
	α_0	0.01246	0.01220
	α_1	0.39219 (0.15941)	0.34786 (0.10315)
	α_2	0.30828 (0.15010)	--
	α_3	0.14316 (0.10046)	--
	β_1	--	0.42506 (0.11073)
n = 500	ϕ	0.36048 (0.03961)	0.39819 (0.04129)
	α_0	0.01144	0.01695
	α_1	0.41834 (0.09684)	0.40938 (0.06717)
	α_2	0.33589 (0.08893)	--
	α_3	0.14313 (0.05887)	--
	β_1	--	0.29045 (0.06177)

The proposed and GARCH models are built up and listed at the top of Table 9. The length of the data set, n, is 200 and 500. For each time series, 500 replications are generated. In order to evaluate the performance of forecasting the volatility among different models, the average AIC and SIC described in section 3.3.2 and 3.3.3 are calculated. Besides these criteria, the relative squared error loss function is also considered, namely heteroscedasticity-adjusted MSE (HMSE, Bollerslev and Ghysels (1996)),

$$\text{HMSE} = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\hat{a}_t^2}{h_t} - 1\right)^2 .$$

We use the models listed in Table 9 to do the out-of-sample forecasting. For each time series, 1000 replications are forecasted and we calculated the average of their errors, i.e. the differences between their corresponding a_t^2 estimated from the Observation Equation, and the conditional variance, h_t. In Table 10, we compare the forecast errors of the proposed model with those of the GARCH(1,1) and find that the proposed model prevailed the GARCH (1,1). The GARCH(1,1) model is weaker in modeling the time series as it does not take care of the seasonal conditional heteroscedastic variance.

Table 10. The out-of-sample forecast comparisons

		Proposed	GARCH(1,1)
n = 200	AIC_A	**-452.19237**	-299.77252
	SIC_A	**-435.70078**	-286.57925
	AIC_B	**-4.03586**	-3.05629
	SIC_B	**-3.95340**	-2.99032
	HMSE	**2.66182**	25.24020
n = 500	AIC_A	**-1099.20655**	-667.93969
	SIC_A	**-1078.13351**	-651.08126
	AIC_B	**-3.12117**	-2.11810
	SIC_B	**-3.07903**	-2.08438
	HMSE	**2.80671**	71.98713

4.4 SEASONAL GARCH MODEL

The next natural step is to compare our model with some seasonal heteroscedasticity models. It is worthwhile to look into the seasonal GARCH models.

The seasonal GARCH model is given by

$$a_t \mid \psi_{t-1} \sim N(0, h_t), \tag{4.6}$$

$$\delta_t(B)\Delta_t(B^d)h_t = \alpha_0 + [\omega_t(B)\Omega_t(B^d)-1]a_t \tag{4.7}$$

$$\Rightarrow \quad h_t = \alpha_0 + \sum_{i=1}^{s+Sd} \alpha_i a_{t-i}^2 + \sum_{i=1}^{r+Rd} \beta_i h_{t-i}$$

where $s > 0, r \geq 0,$

$S > 0, R \geq 0,$

$\alpha_0 > 0, \alpha_i \geq 0, i = 1, \ldots, s + Sd,$

and $\beta_i \geq 0, i = 1, \ldots, r + Rd.$

Here a_t are the residuals obtained from the multiplicative seasonal ARMA model and d is the seasonal period, $\delta_t(B)$, $\Delta_t(B^d)$, $\omega_t(B)$, and $\Omega_t(B^d)$ are defined the same as in (2.3). The model reduces to a seasonal ARCH model if r and R = 0 and is also a special case of seasonal CHARMA $(s,0)\times(S,0)_d$.

Rewrite (4.7) as

$$a_t^2 = \alpha_0 + \sum_{i=1}^{s+Sd} \alpha_i a_{t-i}^2 + \sum_{j=1}^{r+Rd} \beta_j a_{t-j}^2 - \sum_{j=1}^{r+Rd} \beta_j \nu_{t-j} + \nu_t$$

$$= \alpha_0 + \sum_{i=1}^{k} (\alpha_i + \beta_i) a_{t-i}^2 - \sum_{j=1}^{r+Rd} \beta_j \nu_{t-j} + \nu_t \tag{4.8}$$

where $\nu_t = a_t^2 - h_t = (\eta_t^2 - 1)h_t$ with $\eta_t \overset{i.i.d.}{\sim} N(0,1)$,

and $\quad k = \max(r + Rd, s + Sd).$

Here ν_t is serially uncorrelated with mean zero, (4.8) can be interpreted as an ARMA(k,r+Rd) in a_t^2. If r, R ≥ s, S, or s, S ≥ r, R, (4.8) can also be regarded as ARMA$(m,r)\times(M,R)_d$, where m = max(r, s) and M = max(R, S).

Multiplying a_{t-n}^2 on both sides of (4.8) and let the covariance function for a_t^2 be denoted as

$$\gamma_n = \gamma_{-n} = \mathrm{cov}(a_t^2, a_{t-n}^2).$$

Then it follows that

$$\gamma_n = \sum_{i=1}^{k}(\alpha_i + \beta_i)\gamma_{n-i}, \ n \geq r + Rd,$$

and we get the Yule-Walker equations

$$\rho_n = \sum_{i=1}^{k}(\alpha_i + \beta_i)\rho_{n-i}, \ n \geq r + Rd.$$

4.5 EMPIRICAL EXAMPLE

4.5.1 GARCH(1,1)

We model the exchange rates by the GARCH model, since it is found to be successful in many applications. The estimates are obtained as follows:

$$Y_t = a_t$$
$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}$$

where $\alpha_0 = 0.0454$,

$\alpha_1 = 0.0998$ (0.0154),

and $\beta_1 = 0.8115$ (0.0258).

α_1 and β_1 are both highly significant at the 5% level. To check for the model adequacy, we employ the acf of the squared standardized residuals up to order 10 and the results are given in Table 11. Although all the acf are within the error bound

(≈±0.04109), the acf at lag 5 is large as compared with the other lags and the acf of the proposed model.

Table 11. The autocorrelation coefficients of squared standardized residuals up to order 10

lag	1	2	3	4	5	6	7	8	9	10
acf	0.0096	-0.0284	0.0072	-0.0139	0.0187	0.0091	-0.0362	-0.0059	-0.0015	-0.0133

4.5.2 Seasonal GARCH

We use the criteria described in section 3.3.2 and 3.3.3 to select the order of seasonal GARCH. Recall the general form for the model can be written as.

$$Y_t = a_t$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-5}^2 + \alpha_3 a_{t-6}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-5}$$

We estimate three GARCH models and the results are shown in Table 12.

Table 12. Order selection of seasonal GARCH

	seasonal GARCH 1	seasonal GARCH 2	seasonal GARCH 3
α_0	0.06866	0.18898	0.12860
α_1	0.09884	--	0.11211
α_2	0.04031	0.10296	0.07535
β_1	0.72524	--	--
β_2	--	0.51074	0.55642
Log-Likelihood	**-259.35744**	-306.72872	-274.84731
AIC_A	**526.71488**	619.45743	557.69462
SIC_A	**549.79577**	636.76810	580.77551
MSE	1.25625	1.26884	**1.25560**
AIC_B	0.23151	0.240635	**0.23099**
SIC_B	0.24125	0.247942	**0.24074**

On comparing the forecasting criteria listed in Table 12, we find that both AIC_A and SIC_A favor model 1, whereas AIC_B and SIC_B favor model 3. And the autocorrelation coefficients of squared standardized residuals are reported in Table

13. By looking at the magnitude of the differences in the criterion functions, we conclude that model 1 is more successful in capturing the seasonal heteroscedasticity.

Table 13. The autocorrelation coefficients of squared standardized residuals up to order 10

lag	1	2	3	4	5	6	7	8	9	10
acf 1	0.0071	-0.0251	0.0147	-0.0071	0.0040	0.0023	-0.0397	-0.0084	-0.0028	-0.0100
acf 3	-0.0015	0.0164	0.0566	0.0280	-0.0212	0.0002	-0.0181	0.0339	0.0147	-0.0173

4.5.3 Comparison

Table 14. The in-sample forecast comparisons of various models for USD/JPY

	Proposed	GARCH	Seasonal GARCH 1	Seasonal GARCH 3
Log-Likelihood	**-255.71637**	-261.00391	-259.35744	-274.84731
AIC_A	**523.43274**	528.00782	526.71488	557.69462
SIC_A	558.05408	**545.31849**	549.79577	580.77551
MSE	**1.25382**	1.25834	1.25625	1.25560
AIC_B	0.23126	0.23233	0.23151	**0.23099**
SIC_B	0.24587	**0.23963**	0.24125	0.24074
HMSE	**4.54536**	4.78313	4.73824	4.64742

In Table 14, we report the evaluated criteria for forecasting volatility of the proposed, GARCH, and seasonal GARCH models. The proposed model produces the largest log-likelihood and the smallest MSE, and HMSE. From log-likelihood and MSE, we can see that GARCH does not do well. However, both SIC_A and SIC_B favor the GARCH model probably because of its parsimony. On the average, the proposed model does quite well but the trade-off is that it has more parameters. We believe seasonal heteroscedasticity exists and cannot be ignored in the modeling of the conditional variance of daily exchange rate for Japanese Yen.

CHAPTER 5 CASE STUDY

5.1 INTRODUCTION

Our case study concerns the money supply M1 of the United States. Money is the medium of exchange because the use of money is more convenient than the system of barter. In addition, money is a store of value – a good maintains some or all of its value over time and a unit of account – the standard unit for quoting prices and measuring value.

There are three monetary aggregates give the definition of money: M1, M2, and M3. M1 can be spent immediately and without restrictions. It is composed of currency (coins and notes in circulation), traveler's checks, demand deposits, and other checkable deposits (i.e., negotiable order of withdrawal [NOW] accounts, and automatic transfer service [ATS] accounts, and credit union share drafts.). M2 is the combination of M1 with short term investment accounts: small time deposits, savings deposits, money market mutual funds, overnight repurchase agreement, and overnight Eurodollars. M3 is combination of M2 with long term investment accounts: large denomination time deposits, money market mutual fund shares (institutional), term repurchase agreements, and term Eurodollars.

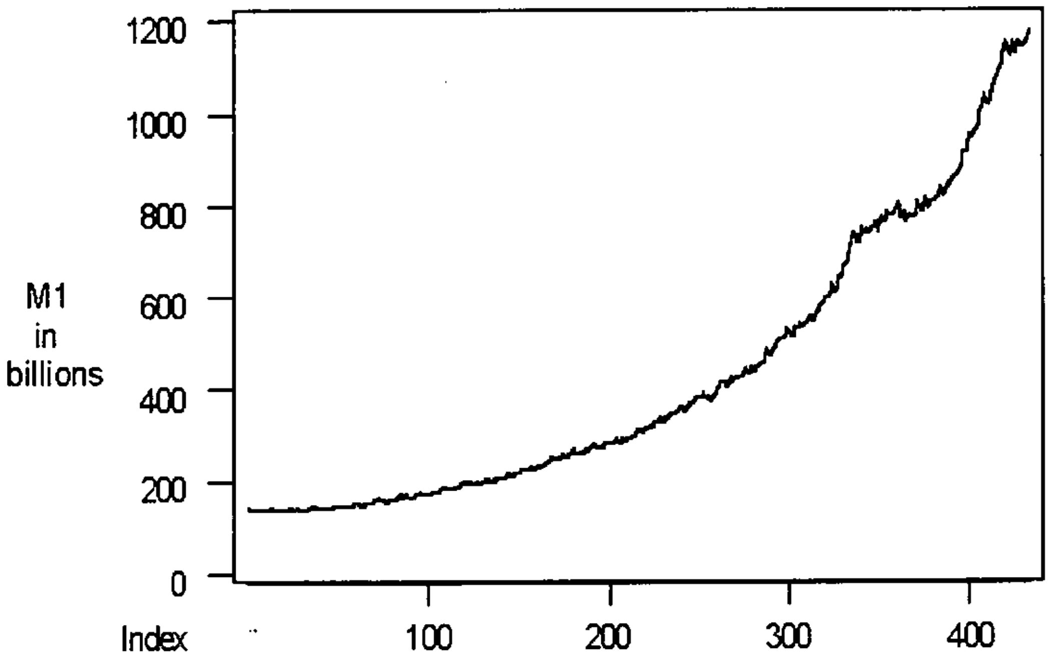
Money supply is complicatedly linked with the other important economic indicators, such as interest rates, and these connections will influence the economy. In the short run, if the money supply increases, then the interest rate will decrease and leads to an increase in the business investment spending. An increase in production will lower the unemployment rate. However, inflation occurs when a nation's central bank attempts to supply a greater quantity of money than the public desires to hold. It is because rising prices help balance the amount of money demanded and supplied. Therefore, in the long run, higher and higher inflation will result in production decreasing and unemployment will be higher again.

We can see that the money supply is an important figure in the economy. Since the United States is the largest economy in the world, it has significant economic and financial effects on other countries. We will use one of its money aggregates, M1, as a case study in this chapter.

5.2 INITIAL EXAMINATION

We consider the money supply (M1) of United States, covering the period Jan 1959 to Dec 1994, collected from Datastream. There are 432 monthly observations in total. The time series plot (in billions) is shown in Figure 7.

Figure 7. The time series plot of the money supply (M1) of United States



To make the series stationary, we take natural logarithm of the data, and then perform nonseasonal and seasonal differencing. By the usual Box-Jenkins analysis (Figures 8 and 9), the suggested model is MA(1)-SMA(1),

$$\underset{}{Y_t} = (1 + \underset{(0.0463)}{0.3139}B)(1 - \underset{(0.0392)}{0.6432}B^{12})a_t$$

where X_t is the logged data of M1,

and $Y_t = 100 \times \nabla^1 \nabla^{12} X_t$.

The values inside the brackets are the standard errors.

However, the residual autocorrelations at lags 3, 6 and 9 are significantly different from zero (Figure 10). Hence, the tentative model is changed to

$$Y_t = (1 + \underset{(0.04540)}{0.3025}B + \underset{(0.04671)}{0.1529}B^3 + \underset{(0.04658)}{0.1556}B^6 + \underset{(0.04673)}{0.1681}B^9)(1 - \underset{(0.04026)}{0.6409}B^{12})a_t$$

Figure 8. The plot of acf for the money supply (M1) of United States

```
Autocorrelations:   M1DATA

     Auto- Stand.
Lag  Corr.   Err. -1  -.75   -.5 -.25   0    .25  .5   .75   1  Box-Ljung  Prob.
                  ----------------------------------------------
  1   .321   .049                    . ³*.****                     43.530   .000
  2   .072   .049                    . ³*.                         45.703   .000
  3   .099   .049                    . ³**                         49.868   .000
  4  -.040   .049                    .*³ .                         50.551   .000
  5   .076   .048                    . ³**                         53.042   .000
  6   .064   .048                    . ³*.                         54.802   .000
  7  -.048   .048                    .*³ .                         55.805   .000
  8   .138   .048                    . ³*.*                        64.013   .000
  9   .133   .048                    . ³*.*                        71.670   .000
 10  -.046   .048                    .*³ .                         72.563   .000
 11  -.122   .048                    **³ .                         79.008   .000
 12  -.401   .048              ******.*³ .                        148.635   .000
 13  -.147   .048                  *.*³ .                        158.059   .000
 14   .076   .048                    . ³**                        160.604   .000
 15   .014   .048                    . * .                        160.690   .000
 16   .032   .048                    . ³*.                        161.127   .000

Plot Symbols:      Autocorrelations *     Two Standard Error Limits .
Total cases:  419     Computable first lags:  418
```

Figure 9. The plot of pacf for the money supply

```
Partial Autocorrelations:   M1DATA

   Pr-Aut- Stand.
Lag  Corr.   Err. -1  -.75  -.5 -.25   0   .25  .5   .75   1
                  ----------------------------------------
  1   .321   .049                      . ³*.****
  2  -.035   .049                      .*³ .
  3   .097   .049                      . ³**
  4  -.112   .049                     **³ .
  5   .142   .049                      . ³*.*
  6  -.018   .049                      . * .
  7  -.055   .049                      .*³ .
  8   .173   .049                      . ³*.*
  9   .043   .049                      . ³*.
 10  -.111   .049                     **³ .
 11  -.140   .049                     *.*³ .
 12  -.365   .049                 *****.*³ .
 13   .142   .049                      . ³*.*
 14   .086   .049                      . ³**
 15   .074   .049                      . ³*.
 16  -.025   .049                      .*³ .

Plot Symbols:       Autocorrelations *     Two Standard Error Limits .
Total cases:  419     Computable first lags:  418
```

Figure 10. The plot of acf for residuals of the model

```
Autocorrelations:   ERR_1       Error for M1DATA from ARIMA, MOD_2 NOCON

     Auto- Stand.
Lag  Corr.   Err. -1  -.75  -.5 -.25   0   .25  .5   .75   1   Box-Ljung  Prob.
                  ----------------------------------------
  1  -.017   .049                      . * .                       .123   .726
  2   .026   .049                      . ³*.                       .405   .817
  3   .224   .049                      . ³*.**                   21.690   .000
  4  -.078   .049                     **³ .                      24.256   .000
  5   .056   .048                      . ³*.                     25.616   .000
  6   .185   .048                      . ³*.**                   40.291   .000
  7  -.104   .048                     **³ .                      44.898   .000
  8   .079   .048                      . ³**                     47.590   .000
  9   .191   .048                      . ³*.**                   63.226   .000
 10  -.124   .048                     **³ .                      69.857   .000
 11   .007   .048                      . * .                     69.877   .000
 12   .047   .048                      . ³*.                     70.833   .000
 13  -.127   .048                     *.*³ .                     77.831   .000
 14   .109   .048                      . ³**                     82.978   .000
 15   .044   .048                      . ³*.                     83.815   .000
 16  -.045   .048                      .*³ .                     84.713   .000

Plot Symbols:       Autocorrelations *     Two Standard Error Limits .
Total cases:  419     Computable first lags:  418
```

5.3 TEST FOR VARYING CONDITIONAL VARIANCE

Inspection of the autocorrelation coefficients (Figure 11) for the residuals do not reveal any mis-specification. However, by examining the plot of the squared residuals (Figure 12), there is a spike (exceeding $2/\sqrt{n}$) at lag 1 and lag 12 which means that there may be a nonseasonal and a seasonal ARCH pattern. To confirm this, we carry out the McLeod and Li test described in section 3.2, $Q_{aa}(12) = 51.265$ is highly significant at the 5% level so the null hypothesis of no ARCH effect can be rejected.

Figure 11. The plot of acf for the residuals for model 2

```
Autocorrelations:    ERR_2        Error for M1DATA from ARIMA, MOD_4 NOCON

     Auto- Stand.
Lag  Corr.   Err. -1  -.75  -.5 -.25   0   .25  .5   .75   1   Box-Ljung  Prob.
                  ----------------------------------------
  1   .010   .049                    . * .                          .042   .838
  2   .040   .049                    . ³*.                          .715   .699
  3   .020   .049                    . * .                          .893   .827
  4   .005   .049                    . * .                          .904   .924
  5   .033   .048                    . ³*.                         1.380   .926
  6   .002   .048                    . * .                         1.383   .967
  7  -.030   .048                    .*³ .                         1.767   .972
  8   .052   .048                    . ³*.                         2.941   .938
  9   .019   .048                    . * .                         3.091   .961
 10  -.052   .048                    .*³ .                         4.279   .934
 11  -.038   .048                    .*³ .                         4.911   .935
 12   .034   .048                    . ³*.                         5.421   .942
 13  -.102   .048                    **³ .                         9.954   .698
 14   .105   .048                    . ³**                        14.786   .393
 15   .032   .048                    . ³*.                        15.231   .435
 16   .017   .048                    . * .                        15.365   .498

Plot Symbols:       Autocorrelations *     Two Standard Error Limits .
Total cases:  419     Computable first lags:  418
```

Figure 12. The plot of acf for the squared residuals

```
Autocorrelations:    SQ

     Auto- Stand.
Lag  Corr.   Err. -1  -.75  -.5 -.25   0   .25  .5   .75   1   Box-Ljung  Prob.
                  ----------------------------------------
  1   .141   .049                    . ³*.*                        8.332   .004
  2   .019   .049                    . * .                         8.489   .014
  3   .047   .049                    . ³*.                         9.417   .024
  4   .105   .049                    . ³**                        14.118   .007
  5   .052   .048                    . ³*.                        15.282   .009
  6   .090   .048                    . ³**                        18.734   .005
  7   .044   .048                    . ³*.                        19.575   .007
  8   .112   .048                    . ³**                        24.970   .002
  9   .050   .048                    . ³*.                        26.032   .002
 10   .023   .048                    . * .                        26.256   .003
 11   .157   .048                    . ³*.*                       36.909   .000
 12   .182   .048                    . ³*.**                      51.267   .000
 13   .073   .048                    . ³*.                        53.564   .000
 14   .009   .048                    . * .                        53.599   .000
 15   .070   .048                    . ³*.                        55.745   .000
 16   .036   .048                    . ³*.                        56.305   .000

Plot Symbols:       Autocorrelations *     Two Standard Error Limits .
Total cases:  419     Computable first lags:  418
```

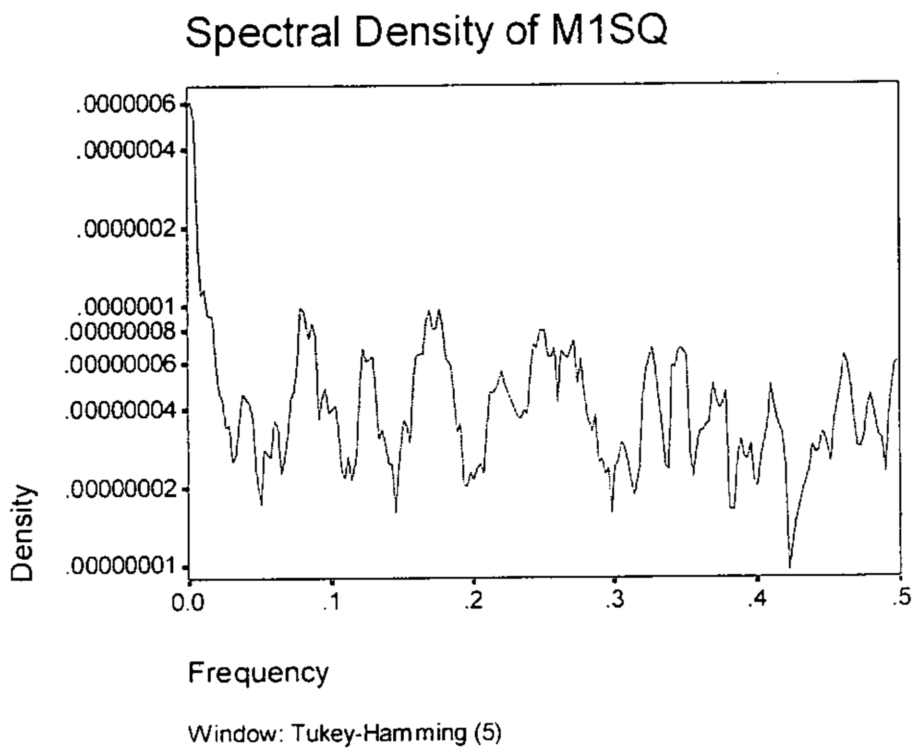
Other than the autocorrelation function, we can also investigate the periodic behavior of time series by spectral analysis. In spectral analysis, we decompose a time series into a function of sines and cosines by Fourier decomposition and calculate the power spectrum. The power spectrum p(f) is defined by

$$p(f) = 2[\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos(2\pi fk)], \quad 0 \le f \le \tfrac{1}{2},$$

where γ_k is the lag k autocovariance.

At the seasonal frequencies, there are peaks in the power spectrum of the seasonal time series. If the length of the seasonal period is denoted as T, then the seasonal frequency f in cycles per unit time is the reciprocal of the period $f = 1/T$. So if we find out the seasonal frequency by plotting the spectrum, we can calculate back the seasonal period.

Figure 13. The spectral density for the squared residuals of M1



We calculate the spectral density function of the squared residuals to support the finding of the seasonality in the conditional variance. In Figure 13, we can see that the frequency of the first major peak is around 0.08. $T = 1/f \approx 12$. This supports that there is a seasonal period of 12 in the conditional variance.

5.4 ESTIMATION AND DIAGNOSTIC CHECKING

5.4.1 The proposed model

Since we detect non-constant conditional variance of the money supply, we build up the innovation equation. Firstly, the orders of equation needed to be determined as described in section 3.3. We detect spikes at lag 1 and lag 12 of the autocorrelation coefficients in the squared residuals, using the same technique as identifying Box-Jenkins model, we guess the innovation equation is multiplicative seasonal model with seasonal period 12. Putting all possible models of order 1 into the general form, we can write $E(a_t^2 \mid \psi_{t-1})$ as

$$E(a_t^2 \mid \psi_{t-1}) = h_t = \sigma_e^2 + \beta_1 a_{t-1}^2 + \beta_2 a_{t-12}^2 + \beta_3 a_{t-13}^2 + \beta_4 y_{t-1}^2 + \beta_5 y_{t-12}^2 + \beta_6 y_{t-13}^2$$

At the top part of Table 15 we report the maximum likelihood estimates with asymptotic standard error in the parentheses of the innovation equation for different order and at the bottom part of the Table we provide the criteria as described in section 3.3.2 and 3.3.3 for model selection.

From the viewpoint of mean square error, both AIC and SIC suggest the order of innovation equation is $(0,1)\times(1,0)_{12}$. However, using log-likelihood as a loss function, AIC favors the order $(1,0)\times(1,0)_{12}$ and SIC favors $(0,1)\times(1,0)_{12}$ because of less parameters. Recall the simulation in section 3.3.4, we found that the criteria based on the log-likelihood favors the right order rather than the mean square error does. And AIC is better than SIC when the "real" model has more parameters. We thus build up both models to see the differences of forecasting performance in the latter section.

Table 15. The AIC and SIC of the specified order for the innovation equation

$(r,s)\times(R,S)$	$(0,0)\times(1,0)_{12}$	$(0,0)\times(0,1)_{12}$	$(0,1)\times(1,0)_{12}$	$(1,0)\times(1,0)_{12}$
σ_e^2	0.24182	0.28222	0.21289	0.19588
β_1	—	—	—	0.13815 (0.06659)
β_2	0.14823 (0.06651)	—	0.15382 (0.06775)	0.15134 (0.06841)
β_3	—	—	—	0.02696 (0.05256)
β_4	—	—	0.05950 (0.03632)	—
β_5	—	0.00432 (0.02214)	—	—
AIC_A	-114.4522	-104.2303	-119.2465	**-119.2940**
SIC_A	-106.3764	-96.1545	**-107.1329**	-103.1425
AIC_B	-1.6293	-1.5950	**-1.6437**	-1.6346
SIC_B	-1.6101	-1.5758	**-1.6148**	-1.5960

Maximum likelihood estimates with standard error reported in parentheses for the joint observation and innovation equations are as follows:

Proposed model A $(0,1)\times(1,0)_{12}$

$$Y_t = (1 + \underset{(0.0413)}{0.2836}B + \underset{(0.0416)}{0.1657}B^3 + \underset{(0.0445)}{0.1572}B^6 + \underset{(0.0455)}{0.1711}B^9)(1-\underset{(0.0371)}{0.6264}B^{12})a_t$$

$$(1 - \Delta_{1,t}B^{12})a_t = \omega_{1,t}Y_{t-1} + e_t$$

$$\Rightarrow \quad h_t = \sigma_e^2 + \beta_1 y_{t-1}^2 + \beta_2 a_{t-12}^2$$

where $\sigma_e^2 = 0.2145$,

$\beta_1 = 0.0587\ (0.0360)$,

and $\beta_2 = 0.1563\ (0.0649)$.

All the t-ratios are significant at the 10% level. Note that the t-ratios of β_2 is 2.41, which is significant at the 5% level. This indicates that the money supply does have annual conditional heteroscedasticity.

Proposed model B $(1,0)\times(1,0)_{12}$

$$Y_t = (1 + \underset{(0.0477)}{0.3037}B + \underset{(0.0408)}{0.1666}B^3 + \underset{(0.0446)}{0.1521}B^6 + \underset{(0.0440)}{0.1692}B^9)(1-\underset{(0.0379)}{0.6234}B^{12})a_t$$

$$(1-\delta_{1,t}B)(1 - \Delta_{1,t}B^{12})a_t = e_t$$

$$\Rightarrow \quad h_t = \sigma_e^2 + \beta_1 a_{t-1}^2 + \beta_2 a_{t-12}^2 + \beta_3 a_{t-13}^2$$

where $\sigma_e^2 = 0.1996$,

$\beta_1 = 0.1382\ (0.0749)$,

$\beta_2 = 0.1490\ (0.0686)$,

$\beta_3 = 0.0210\ (0.0586)$.

Again, the coefficient for a_{t-12}^2 is significantly different from zero at the 5% level, it supports that the money supply has conditional heteroscedasticity as found in model A.

Employing squared standardized residual autocorrelations described in section 3.5, the squared standardized residuals up to order 13 are provided in Table 16. The only large autocorrelations of the squared standardized residuals for model A at lag 11. It was also found the same phenomenon appeared in all seasonal GARCH models. For comparison purpose, we do not pursue this issue further.

Table 16. The autocorrelation coefficients of squared standardized residuals up to order 13

lag	1	2	3	4	5	6	7	8	9	10	11	12	13
acf A	0.047	-0.021	0.073	0.072	0.021	0.052	0.023	0.098	0.001	0.017	0.149	-0.009	-0.024
acf B	-0.154	-0.012	0.082	0.056	0.023	0.043	0.029	0.105	0.009	0.002	0.128	-0.007	-0.017

From Table 17, we can see that the squared residual autocorrelations have been much improved for lag 1 and lag 12, they are within the standard error bounds ($\approx \pm 2/\sqrt{n} = \pm 0.098$) now. Both models remove the annual heteroscedasticity.

Table 17. The acf of squared residual before and after the innovation equations fitted

	lag 1	lag 12
before adding the innovation equation	0.141	0.182
after adding the innovation equation (Proposed_A)	0.0470	-0.0092
after adding the innovation equation (Proposed_B)	-0.0154	-0.0068

We then consider other models.

5.4.2 MA-GARCH model

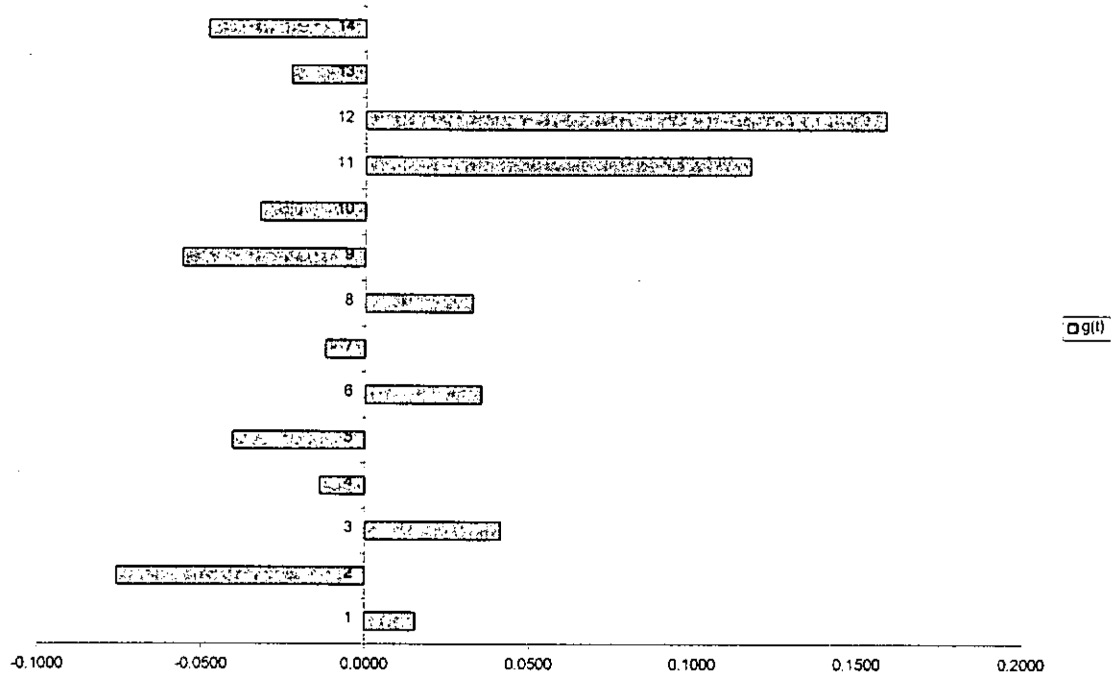
Using the maximum likelihood method described in section 4.4.1 to model the money supply. The result is:

$$Y_t = (1 + \underset{(0.04511)}{0.2961}B + \underset{(0.04779)}{0.1617}B^3 + \underset{(0.04889)}{0.1496}B^6 + \underset{(0.04820)}{0.1724}B^9)(1 - \underset{(0.03502)}{0.5681}B^{12})a_t$$

$$h_t = 0.02876 + \underset{(0.03906)}{0.09154}a_{t-1}^2 + \underset{(0.05281)}{0.81178}h_{t-1}$$

All the coefficients are significantly different from zero at the 5% level. Diagnostic checking results for the whole model is in Figure 14. The autocorrelation coefficients for the squared standardized residuals at lag 11 and 12 are clearly longer than the others. The GARCH(1,1) cannot capture the seasonal ARCH effect.

Figure 14. The acf of the standardized residuals for GARCH model



5.4.3 MA-Seasonal GARCH model

We firstly select the order of seasonal GARCH for the money supply, the general form of seasonal GARCH of order 1 is

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-12}^2 + \alpha_3 a_{t-13}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-12} + \beta_3 h_{t-13}.$$

In fitting the full model, we find that the coefficients of a_{t-13}^2 and h_{t-13} are negative. So we delete the two coefficients in subsequent work. In the top part of Table 18 reports the maximum likelihood estimates of various models with standard errors in parentheses.

Table 18. Order selection of seasonal GARCH

	seasonal GARCH 1	seasonal GARCH 2	seasonal GARCH 3
α_0	0.04222	0.03967	0.12860
α_1	0.07199	--	0.11211
α_2	0.10489	0.06956	0.07535
β_1	0.68127	--	--
β_2	--	0.82094	0.55642
Log-Likelihood	65.48837	65.44123	**68.47438**
AIC_A	-122.97675	-124.88247	**-128.94876**
SIC_A	-106.82526	-112.76886	**-112.79727**
MSE	0.19189	0.18908	**0.18818**
AIC_B	-1.63172	-1.65124	**-1.65128**
SIC_B	-1.59317	**-1.62233**	-1.61274

From the bottom of Table 18, most of the criteria favor model 3 except SIC_B because GARCH model uses fewer parameters than the others. In view of this result, we decided to choose model 3 for building seasonal GARCH for the money supply.

$$Y_t = \underset{}{(1 + \underset{(0.04574)}{0.3141}B + \underset{(0.04253)}{0.1478}B^3 + \underset{(0.04385)}{0.1502}B^6 + \underset{(0.04491)}{0.1649}B^9)(1 - \underset{(0.03855)}{0.6093}B^{12})a_t}$$

$$h_t = 0.06890 + \underset{(0.05074)}{0.05239}a_{t-1}^2 + \underset{(0.04099)}{0.10329}a_{t-12}^2 + \underset{(0.09775)}{0.62655}h_{t-12}$$

However, the coefficient of a_{t-1}^2 is insignificant. So, we drop the parameter and re-estimate the model, i.e. using model 2.

$$Y_t = (1 + \underset{(0.04184)}{0.3090}B + \underset{(0.04251)}{0.1545}B^3 + \underset{(0.04271)}{0.1545}B^6 + \underset{(0.04594)}{0.1631}B^9)(1 - \underset{(0.03845)}{0.6065}B^{12})a_t$$

$$h_t = 0.06533 + \underset{(0.04639)}{0.10463}a_{t-12}^2 + \underset{(0.06279)}{0.69457}h_{t-12}$$

Both the coefficients at lag 12 are significant at the 5% level, it again supports there is seasonal heteroscedasticity in the variance of a_t for the money supply.

5.5 COMPARISON

In order to assess the accuracy of these estimated models, we calculate their in-sample and out-of-sample forecasts errors and also make a comparison among them. The in-sample results are given in Table 19.

Table 19. The in-sample forecast comparisons of various model for money supply

	Proposed_A	Proposed_B	GARCH	Seasonal GARCH
Log-Likelihood	60.94432	61.74985	61.21831	**64.98187**
AIC_A	-105.88865	-105.49970	-106.43662	**-113.96373**
SIC_A	-73.58568	-69.15886	-74.13366	**-81.66076**
MSE	**0.19349**	0.19355	0.20364	0.19371
AIC_B	**-1.60433**	-1.59926	-1.55319	-1.60322
SIC_B	**-1.52723**	-1.51253	-1.47610	-1.52612
HMSE	2.29254	2.24278	2.26215	**2.15042**

According to Table 19, the criteria of the log-likelihood suggest seasonal GARCH model but the proposed model A is slightly better under the criteria based on the mean square error. And seasonal GARCH model gets the smallest HMSE. It is interesting to note that the overall performance of all seasonal models is better than GARCH model. Hence, it is worthwhile to take account of the seasonal pattern in the conditional variance.

That the fitted model fits the historical data well does not imply that it is also good for the future. Therefore, we want to analyze the out-of-sample forecasts. To evaluate the out-of-sample forecasting performance, we forecast the money supply from Jan 1995 to Aug 1998 and calculate their forecasting errors. Table 20 reports the results of one-step-ahead fixed parameter model forecasts. In other words, we use all the models in the previous sections to model the money supply, and update the data set for every forecast but do not change the parameter estimates, i.e., we use the information only up to time t for model building. Let the one-step-ahead forecast be $\hat{h}_{t+1}$, then $\hat{h}_{t+1} = E(Y_{t+1} | \psi_t)$.

Table 20. The out-of-sample one-step-ahead fixed parameter model forecast comparisons for the money supply

	$Proposed_A$	$Proposed_B$	GARCH	Seasonal GARCH
Log-Likelihood	**10.23510**	9.83791	8.96307	8.67448
AIC_A	**-4.47020**	-1.67582	-1.92615	0.65105
SIC_A	**9.80332**	14.38189	12.34737	16.70875
MSE	**0.06897**	0.07369	0.07482	0.08094
AIC_B	**-2.31050**	-2.19885	-2.22901	-2.10500
SIC_B	**-1.98610**	-1.83390	-1.90461	-1.74005
HMSE	**0.76979**	0.828407	1.05624	0.83806

As we can observe from Table 20, all the criteria agree that the proposed model A gives the best volatility forecasts. It is slightly different from the result in in-sample forecasts. In Table 21, we report the criteria for the one-step-ahead updating model forecast errors. We continue to carry out the procedure of updating the time series and re-estimating the parameters of all models. To keep the number of observations constant, we will drop the oldest one when we add the latest one into the series. Then, we use the new model to produce the next volatility forecasts. The forecast result of the fixed model is similar to the updating model. We cannot obtain result of the GARCH model because the coefficients of the model do not converge in some cases. It further suggests that the GARCH model is not suitable for forecasting the money supply.

From the in-sample and out-of-sample forecasts, it can also be seen that the seasonal CHARMA is more successful in capturing the volatility of the money supply than the GARCH, seasonal ARCH (proposed B model), and seasonal GARCH.

Table 21. The out-of-sample one-step-ahead updating parameter model forecast comparisons for the money supply

	$Proposed_A$	$Proposed_B$	GARCH	Seasonal GARCH
Log-Likelihood	**10.31538**	9.89619	--	8.73163
AIC_A	**-4.63076**	-1.79238	--	-1.46326
SIC_A	**9.64276**	14.26532	--	12.81026
MSE	**0.06794**	0.07363	--	0.08263
AIC_B	**-2.32555**	-2.19964	--	-2.12982
SIC_B	**-2.00116**	-1.83469	--	-1.80543
HMSE	**0.75778**	0.82394	--	0.76609

CHAPTER 6 CONCLUSION AND FURTHER WORK

Undoubtedly, the GARCH model is very popular and successful in modeling economic and financial data, since it allows the conditional variance to vary, which is more realistic for the economic world. However, there is another important characteristic in the economic world, that is seasonality, especially for the high frequency data such as daily and intraday series. It can be seen from the real data of the exchange rate USD/JPY, and money supply that there are clear seasonal conditional heteroscedasticity. In a certain sense, the exchange rate is more striking to the statisticians. This is because its linear model is just the white noise, but then it is found that its variance possesses both non-seasonal and seasonal structures.

The GARCH model does not quite capture this phenomenon. Therefore, Bollerslev and Ghysels (1996) proposed the P-GARCH model that can compensate for the weakness of the GARCH and characterize the periodic conditional heteroscedasticity. The advantage of the P-GARCH is that it is flexible to capture the pattern for the seasonality. For instance, the periodic cycles need not be purely repetitive. The model is good for the data that has strong seasonal conditional heteroscedasticity. Nevertheless, the length of the period, d, can bring along a rather serious problem. From (4.5), when the period increases, the number of parameters can also increase dramatically. From the exchange rate and money supply, we show that the seasonal CHARMA is able to fully capture the seasonal conditional heteroscedasticity with parsimonious parameters.

For GARCH (or seasonal GARCH) model, there is a restriction that the conditional variance must use the information of both past squared residuals and lagged conditional variances. Or it can be reduced to ARCH (or seasonal ARCH) model including only past squared residuals but it cannot be reduced to only lagged conditional variance terms. However, the seasonal CHARMA model can use both the past squared residuals and past squared observations, or either one. It is more flexible than the GARCH (or seasonal GARCH) model.

The results in the volatility forecasts of the exchange rate are probably not strong enough to distinguish the proposed model from the others. It does, however, indicates empirically that seasonal conditional heteroscedasticity is an eminent feature in financial time series. In the money supply, both the seasonal CHARMA model and the seasonal GARCH model are much more successful in modeling the data than the GARCH. For the out-of-sample forecasts, the seasonal CHARMA is slightly better than the seasonal GARCH.

In some cases, seasonal ARCH model is enough to capture seasonal conditional heteroscedasticity, such as the exchange rate of USD/JPY. In other cases, both the data and the residuals are used to model the variance. Sometimes, the time series may prefer a model that makes more use of the data, such as the money supply. The interpretation is also natural, as that means the volatility is highly dependent on the level of the data. In this example, the seasonal CHARMA model provides the forecasts that are more suitable than the seasonal GARCH. Hence, the proposed model is valuable in modeling economic or financial data and deserves further investigation.

It is worthwhile to analyze the differences of the properties among the proposed model, the P-GARCH, and the seasonal GARCH with respect to standard tools, such as the pattern of the autocorrelation function. This will help us in the choice of models.

During this research, we have also briefly studied the GDP of Hong Kong, a set of 91 quarterly data. Seasonality was found but the effect is not strong. Thus it was not further pursued. It is well known that the study of ARCH usually needs large data sets. The natural problem is then the study of seasonal conditional heteroscedasticity with short data sets. A possible direction is the Baysian analysis of the seasonal CHARMA model using modern techniques such as the Gibbs sampler.

The availability of high frequency data such as minute-minute, and tick-tick data and the help of high speed computer account for an increase in popularity for modeling intraday series. Many papers have shown that the intraday series have strong seasonality pattern in the conditional variances, such as Andersen and Bollerslev (1997), Müller *et al.* (1990). So, it is interesting to test how well the proposed model can model such series.

In addition, we have only investigated the weekend effect for the daily exchange rate using the seasonal CHARMA model. It will be worthwhile if we can build up the periodic CHARMA model and make a comparison between the periodic and seasonal CHARMA model.

Finally, many economic and financial series do not only depend on its own history but are also influenced by the other variables. Wong and Li (1997) established the multiplicative conditional heteroscedastic model and study the Standard & Poor's 500 (SP500) index and the Sydney All Ordinaries (SAO) index as an example. In this research, the seasonal CHARMA model is only discussed in a univariate case. Therefore, it is worthwhile to extend the model to the multivariate case.

REFERENCES

Akaike, H. (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716-723.

-------- (1977), "On Entropy Maximization Principle," in *Applications of Statistics*, ed. P. R. Krishnaiah, Amsterdam: North-Holland, pp. 27-41.

Andel, J. (1976), "Autoregressive Series with Random Parameters," *Mathematische Operationsforschung und Statistik, Series Statistics*, 7, 735-741.

Andersen, T. G., and Bollerslev, T. (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115-158.

Baillie, R. T., and Bollerslev, T. (1989), "The Message in Daily Exchange Rates: A Conditional-Variance Tale," *Journal of Business & Economic Statistics*, 7, 297-305.

Berndt, E. R., Hall, B. H., Hall, R. E., and Haussman, J. A. (1974), "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, 4, 653-665.

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

-------- (1987), "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, 542-547.

-------- (1998), "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process," *Journal of Time Series Analysis*, 9, 121-131.

Bollerslev, T., and Ghysels, E. (1996), "Periodic Autoregressive Conditional Heteroscedasticity," *Journal of Business & Economic Statistics*, 14, 139-151.

Bollerslev, T., and Hodrick, R. J. (1995), "Financial Market Efficiency Tests," in *Handbook of Applied Econometrics: Macroeconomics*, Oxford, U. K.; Cambridge, Mass: Blackwell, pp. 415-458.

Bollerslev, T., Chou R. Y., and Kroner, K. F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5-59.

Box, G. E. P. and Jenkins, G. M. (1994), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Conlisk, J. (1974), "Stability in a Random Coefficient Model," *International Economic Review*, 15, 529-533.

-------- (1976), "A Further Note on Stability in a Random Coefficient Model," *International Economic Review*, 17, 759-764.

Dacorogna, M. M., Müller, U. A., Nagler, R. J., Olsen, R. B., and Pictet, O. V. (1993), "A Geographical Model for the Daily and Weekly Seasonal Volatility in the FX market," *Journal of International Money and Finance*, 12, 413-438.

Dacorogna, M. M., Gauvreau, C. L., Müller, U. A., Nagler, R. J., Olsen, R. B., and Pictet, O. V. (1996), "Changing Time Scale for Short-term Forecasting in Financial Markets," *Journal of Forecasting*, 15, 203-227.

Dufour, J. M., and Roy, R. (1986), "Generalized Portmanteau Statistics and Tests of Randomness," *Communication in Statistics-Theory and Method*, 15, 2953-2972.

Engle, R. F. (1982), "Autoregressive Conditional Heterocedasticity with Estimates of Variance of United Kingdom Inflations," *Econometrica*, 50, 987-1007.

Engle, R. F., and Bollerslev, T. (1986), "Modelling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1-50, 81-87.

Engle, R. F., and Ng, V. K. (1993), "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749-1778.

Florentini, G., and Maravall, A. (1996), "Unobserved Components in ARCH Models: An Application to Seasonal Adjustment," *Journal of Forecasting*, 15, 175-201.

Friedman, D., and Vandersteel, S. (1982), "Short-Run Fluctuations in Foreign Exchange Rates," *Journal of International Economics*, 13, 171-186.

Granger, C. W. J., and Andersen, A. P. (1978), *An Introduction to Bilinear Time Series Models*, Gottingen: Vandenhoeck and Ruprecht.

Hsieh, D. A. (1998), "The Statistical Properties of Daily Foreign Exchange Rates: 1974-1983," *Journal of International Economics*, 24, 129-145.

Isaacson, E., and Keller, H. B. (1966), *Analysis of Numerical Methods*, New York: Wiley.

Kendall, M. G. (1953), "The Analysis of Economic Series - Part I: Prices," *Journal of the Royal Stastistical Society Series A*, 116, 11-25.

Li, W. K., and Mak, T. K. (1994), "Non-Linear Time Series with Conditional Heteroskedasticity," *Journal of Time Series Analysis*, 15, 627-636.

Ljung, G. M., and G. E. P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297-303.

Mak, T. K., (1993), "Solving Non-linear Estimation Equations," *Journal of the Royal Statistical Society, Series B*, 55, 945-955.

Mak, T. K., Wong H., and Li, W. K. (1997), "Estimation of Nonlinear Time Series with Conditional Heteroscedastic Variances by Iteratively Weighted Least Squares," *Computational Statistics & Data Analysis*, 24, 169-178.

Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394-419.

McLeod, A. I., and Li, W. K. (1983), "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations," *Journal of Time Series Analysis*, 4, 269-273.

Miller, R. B. (1979), Book review on 'An Introduction to Bilinear Time Series Models', by C. W. Granger and A. P. Andersen, *Journal of the American Statistical Association*, 74, 927.

Moran, P. A. P. (1948), "Some Theorems on Time Series. II. The Significance of the Serial Correlation Coefficient," *Biometrika*, 35, 255-260.

Müller, U. A., Dacorogna, M. M., Olsen, R. B., Pictet, O. V., Schwarz, M., and Morgenegg, C. (1990), "Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law, and Intraday Analysis," *Journal of Banking and Finance*, 14, 1189-1208.

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

Nicholls, D. F., and Quinn, B. G. (1982), *Random Coefficient Autoregressive Models: An Introduction*, Springer Verlag: New York.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.

Tiao, G. C., and Grupe, M. R. (1980), "Hidden Periodic Autoregressive Moving Average Models in Time Series Data," *Biometrika*, 67, 365-373.

Tsay, R. S. (1987), "Conditional Heteroscedastic Time Series Models," *Journal of the American Statistical Association*, 82, 590-604.

Wong, H., and Li, W. K. (1995), "Portmanteau Test for Conditional Heteroscedasticity, Using Ranks of Squared Residual," *Journal of Applied Statistics*, 22, 121-134.

-------- (1997), "On a Multivariate Conditional Heteroscedastic Model," *biometrika*, 84, 111-123.

Yule, G. U. (1927), "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers," *Philosophical transactions of the Royal Society of London, Series A*, 226, 267-298.