Discussion Paper No. 545

A SURVEY OF RECENT THEORETICAL RESULTS FOR TIME SERIES MODELS WITH GARCH ERRORS*

W. K. Li
Shiqing Ling
and
Michael McAleer

June 2001

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

* To appear in *Journal of Economic Surveys*.

A Survey of Recent Theoretical Results for Time Series Models with GARCH Errors*

W.K. Li[1], Shiqing Ling[2] and Michael McAleer[3]

[1]Department of Statistics and Actuarial Science,
University of Hong Kong
[2]Department of Mathematics,
Hong Kong University of Science and Technology
[3]Department of Economics, University of Western Australia

Revised: June 2001

Abstract

This paper provides a review of some recent theoretical results for time series models with GARCH errors, and is directed towards practitioners. Starting with the simple ARCH model and proceeding to the GARCH model, some results for stationary and nonstationary ARMA-GARCH are summarized. Various new ARCH-type models, including double threshold ARCH and GARCH, ARFIMA-GARCH, CHARMA and vector ARMA-GARCH, are also reviewed.

*The authors are most grateful for the helpful comments and suggestions of a referee and seminar participants at the Institute for Monetary and Economic Studies, Bank of Japan, Bocconi University, Milan, Chinese University of Hong Kong, Curtin University of Technology, Edith Cowan University, Erasmus University Rotterdam, Griffith University, Hiroshima University, Hitotsubashi University, International Foundation for Artificial Intelligence, Tokyo, Kyoto University, Monash University, Murdoch University, Centre for Financial Engineering, National University of Singapore, Osaka School of International Public Policy, Osaka University, Queensland University of Technology, Yokohama National University, and University of Waikato. For financial support, the first author would like to thank the Hong Kong Research Grants Council, the second author wishes to acknowledge the Australian Research Council, and the third author is grateful to the Australian Research Council and the Institute of Social and Economic Research, Osaka University.

1 Introduction

A primary feature of the autoregressive conditional heteroscedasticity (ARCH) model, as developed by Engle (1982), is that the conditional variances change over time. Following the seminal idea, numerous models incorporating this feature have been proposed. Among these models, Bollerslev's (1986) generalized ARCH (GARCH) model is certainly the most popular and successful because it is easy to estimate and interpret by analogy with the autoregressive moving average (ARMA) time series model. Analyzing financial and economic time series data with ARCH and GARCH models has become very common in empirical research, with a huge literature having been established. Several excellent surveys on ARCH/GARCH models are available, such as Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), and Bera and Higgins (1993). More recently, the Stochastic Volatility model of Taylor (1986) offers an alternative to GARCH. Stochastic Volatility models will not be discussed in this paper and interested readers are referred to the review by Shephard (1996). In a series of papers, Nelson has made important contributions to the filtering theory of ARCH processes. His work has been nicely summarized by Ross (1996), and hence will not be the focus of attention in this paper. Gourieroux (1997) provides a summary of some earlier results on GARCH models.

The aim of this paper is to provide a review of some recent theoretical results for time series models with ARCH/GARCH errors, and is directed towards practitioners. The plan of the paper is as follows. We begin with the simple ARCH model in Section 2 and proceed to the GARCH model in Section 3. The stationary ARMA-GARCH model is considered in Section 4, and its nonstationary counterpart in Section 5. Finally, we review some results for other ARCH-type models, including double threshold ARCH, ARFIMA-GARCH, CHARMA, and vector ARMA-GARCH, in Section 6. Concluding marks are given in Section 7.

2 ARCH Models

Engle's (1982) ARCH (r) model can be defined as follows:

$$\varepsilon_t = \eta_t h_t^{1/2}, \ \ h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_r \varepsilon_{t-r}^2, \tag{2.1}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$ $(i = 1, \cdots, r)$ and the η_t are a sequence of independently and identically distributed (i.i.d.) random variables with zero mean and unit variance. Denote by $\mathcal{F}_t$ the $\sigma-$field generated by $\{\eta_t, \eta_{t-1}, \cdots\}$. Then $E(\varepsilon_t^2|\mathcal{F}_{t-1}) = h_t$, that is, the conditional variance of the process ε_t varies over time instead of being constant, as in traditional time series analysis.

2.1 Basic Properties

When a new time series model is proposed, a basic question concerns the conditions under which the model will be stationary. Engle (1982) showed that ε_t is second-order stationary (i.e. $E\varepsilon_t^2 < \infty$) if and only if all the roots of

$$z^r - \sum_{i=1}^{r} \alpha_i z^{r-i} = 0 \tag{2.2}$$

are outside the unit circle. To prove this result, Engle (1982) assumed that ε_t starts infinitely far in the past with finite variance, which is impossible to verify in practice. Using a different method, Milhøj (1985) avoided Engle's (1982) assumption and showed that ε_t is second-order stationary if and only if

$$\alpha_1 + \cdots + \alpha_r < 1. \tag{2.3}$$

In particular, Milhøj (1985) showed that (2.3) is also a sufficient condition for strict stationarity and ergodicity of ε_t. Since α_i is nonnegative for $i = 1, \ldots, r$, conditions (2.2) and (2.3) are equivalent by Lemma 2.1 in Ling (1999b).

For the first-order ARCH model, Engle (1982) showed that, if η_t is normal, the $2m$th moment of ε_t exists if and only if

$$\alpha_1^m \prod_{j=1}^{m} (2j - 1) < 1, \tag{2.4}$$

under the assumption that ε_t starts infinitely far in the past with finite $2m$th moment. Without this assumption, Milhøj (1985) obtained the necessary and sufficient condition for the existence of the $2m$th moment of ε_t. When η_t is normal and $r = 1$, Milhøj's condition is the same as (2.4). A unique drawback is that Milhøj's (1985) condition cannot be given an explicit form when $r > 1$ and $m > 2$.

It should be noted that (2.3) is not necessary for the strict stationarity of model (2.1). The necessary and sufficient condition for the strict stationarity of model (2.1) was established by Bougerol and Picard (1992) in terms of the top Lyapunov exponent (see §3.1). The regions of strict stationarity are, in general, much larger than those of second-order stationarity. As an illustration, for the first-order ARCH model, ARCH(1), the various conditions under normality are summarized as follows:

			Moments		
Variable	ε_t	Strict stationarity	2nd	4th	8th
Coefficient	α_1	(0, 3.56214)	(0, 1)	(0, 0.57735)	(0, 0.31239)

Non-normality reduces the permissible range of the ARCH(1) parameter for the 4th and higher moments. It seems difficult to obtain a closed form expression of strict stationarity in terms of the ARCH(r) parameters for any $r > 1$.

2.2 Sample ACVF and ACF

In time series analysis, the autocovariance function (ACVF) and autocorrelation function (ACF) are important because they usually provide meaningful information about the series. Define the sample ACVF and sample ACF, respectively, by

$$
\begin{aligned}
\gamma_{n\varepsilon}(k) &= \frac{1}{n}\sum_{t=k+1}^{n} \varepsilon_t \varepsilon_{t-k} \ , \\
\rho_{n\varepsilon}(k) &= \frac{\gamma_{n,\varepsilon}(k)}{\gamma_{n,\varepsilon}(0)} \ ,
\end{aligned}
$$

where n is the sample size and $k \geq 0$. Correspondingly, the true values are given by:

$$\begin{aligned}
\gamma_\varepsilon(k) &= E(\varepsilon_0 \varepsilon_k) , \\
\rho_\varepsilon(k) &= \frac{\gamma_\varepsilon(k)}{\gamma_\varepsilon(0)}.
\end{aligned}$$

As the ARCH process ε_t is an uncorrelated white noise sequence, $\gamma_\varepsilon(k) = \rho_\varepsilon(k) = 0$ if $k > 0$. Under the fourth moment condition, Milhøj (1985) showed that $\gamma_{n\varepsilon}(k)$ and $\rho_{n\varepsilon}(k)$ are consistent estimators of $\gamma_\varepsilon(k)$ and $\rho_\varepsilon(k)$, respectively, and $\sqrt{n}[\gamma_{n\varepsilon}(k) - \gamma_\varepsilon(k)]$ and $\sqrt{n}[\rho_{n\varepsilon}(k) - \rho_\varepsilon(k)]$ are asymptotically normal.

It is natural to ask if Milhøj's results still hold if the fourth moment condition is not satisfied. This is a difficult problem because ARCH processes exhibit a strong heavy-tailed feature when $E\varepsilon_t^4 = \infty$. Using the point process technique, Davis and Mikosch (1998) showed that, if $E\varepsilon_t^2 < \infty$ but $E\varepsilon_t^4 = \infty$, then

$$\begin{aligned}
&n^{1-2/q} L(n)^{-2} \gamma_{n\varepsilon}(k) \to_d V_q(k) , \\
&n^{1-2/q} L(n)^{-2} \rho_{n\varepsilon}(k) \to_d \frac{V_q(k)}{E\varepsilon_t^2} ,
\end{aligned}$$

where $q \in (2, 4)$ is the unique solution to $E(\alpha_1 \eta_t^2)^{q/2} = 1$, $V_q(k)$ is $q/2-$stable in R, and $L(n)$ is some slowly-varying function. From the above results, $\gamma_{n\varepsilon}(k)$ and $\rho_{n\varepsilon}(k)$ are consistent estimators of $\gamma_\varepsilon(k)$ and $\rho_\varepsilon(k)$, respectively, but the convergence rate is slower than the usual $n^{1/2}$. This result is different from those for linear processes with i.i.d. regularly varying noise. Davis and Resnick (1985, 1986) showed that the sample ACF is still asymptotically normal with scaling $n^{1/2}$ if the i.i.d. noise has finite variance but infinite fourth moment.

Furthermore, Davis and Mikosch (1998) showed that, if $E|\varepsilon|^p < \infty$ for $0 < p < 2$ but $E\varepsilon_t^2 = \infty$, then

$$\begin{aligned}
&n^{1-2/q} L(n)^{-2} \gamma_{n\varepsilon}(k) \to_d V_q(k) , \\
&\rho_{n\varepsilon}(k) \to_d \frac{V_q(k)}{V_q(0)} ,
\end{aligned}$$

where $q \in (0,2)$. In this case, the estimator of the ACF is inconsistent. This result is quite different from that for linear processes with i.i.d. regularly varying noise, in which the sample ACF converges to the true ACF with a convergence rate greater than $n^{1/2}$ (see Davis and Resnick 1985, 1986).

The sample ACVF and ACF of ε_t^2 have also been investigated by Davis and Mikosch (1998). Although they considered only the first-order ARCH model, their results can be extended to higher-order ARCH models. de Vries (1991) demonstrated that, under certain conditions, GARCH processes can generate realizations that have a stable distribution unconditionally.

2.3 Parameter Estimation

The parameters of model (2.1) can be estimated by several methods. The simplest method is the least squares estimator (LSE). First, write model (2.1) as

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_r \varepsilon_{t-r}^2 + \xi_t, \tag{2.5}$$

where $\xi_t = \varepsilon_t^2 - h_t$ and ξ_t can now be considered as a martingale difference. Let $\delta = (\alpha_0, \alpha_1, \cdots, \alpha_r)'$ and $\tilde{\varepsilon}_t = (1, \varepsilon_t^2, \cdots, \varepsilon_{t-r+1}^2)'$. Then the LSE of δ is

$$\hat{\delta} = (\sum_{t=2}^{n} \tilde{\varepsilon}_{t-1} \tilde{\varepsilon}_{t-1}')^{-1} (\sum_{t=2}^{n} \tilde{\varepsilon}_{t-1} \tilde{\varepsilon}_t).$$

Weiss (1986) and Pantula (1989) showed that $\hat{\delta}$ is consistent and asymptotically normal. However, their results assume that the $8th$ moment of ε_t exists, which is a strong condition.

In general, maximum likelihood estimation (MLE) is used to estimate the parameter δ. Given observations ε_t, $t = 1, \cdots, n$, the conditional log-likelihood can be written as

$$L(\delta) = \sum_{t=1}^{n} l_t, \quad l_t = -\frac{1}{2} \ln h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}, \tag{2.6}$$

where h_t is treated as a function of ε_t. Assume that $\delta \in \Theta$, a compact subset of R^{r+1}, and that the true value of δ is δ_0. Define

$$\hat{\delta} = \text{argmax}_{\delta \in \Theta} L(\delta). \tag{2.7}$$

Since the conditional error η_t is not assumed to be normal, $\hat{\delta}$ is called the quasi-maximum likelihood estimator (QMLE). Under the fourth moment condition, Weiss (1986) and Pantula (1989) showed that the QMLE $\hat{\lambda}$ is consistent and asymptotically normal. Ling and McAleer (1999b) proved that the QMLE of δ is consistent and asymptotically normal under only the second moment condition. It is expected that, when ε_t is strictly stationary but $E\varepsilon_t^2 = \infty$, the QMLE will still be consistent and asymptotically normal. The BHHH algorithm is often used to determine $\hat{\delta}$. However, Mak, Wong and Li (1997) suggested that the BHHH algorithm has a convergence problem if the starting values are not sufficiently close to the solutions and that a full Newton-Raphson procedure should instead be used.

When η_t is not normal, the QMLE is not efficient, that is, its asymptotic covariance matrix is not minimal in the class of asymptotically normal estimators. In order to obtain an efficient estimator, one needs to know or estimate the density function of η_t and use an adaptive estimation procedure. This was considered by Linton (1993) and Drost, Klaassen and Werker (1995), who proved that the ARCH model belongs to the locally asymptotically normal (LAN) family. After suitable re-parameterisation, they also constructed adaptive estimators for the parameters of interest.

3 GARCH Models

Bollerslev (1986) extended the ARCH model to the generalized autoregressive conditional heteroscedasticity (GARCH (r, s)) model:

$$\varepsilon_t = \eta_t \sqrt{h_t}\,, \tag{3.1}$$

$$h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \tag{3.2}$$

where $\alpha_0 > 0, \alpha_i \geq 0$, $\beta_i \geq 0$, and η_t is defined as in (2.1).

3.1 Basic Properties

Bollerslev (1986) showed that the necessary and sufficient condition for the second-order stationarity of models (3.1)-(3.2) is:

$$\sum_{i=1}^{r} \alpha_i + \sum_{i=1}^{s} \beta_i < 1. \tag{3.3}$$

For the GARCH(1,1) model, Nelson (1990) obtained the necessary and sufficient condition for strict stationarity and ergodicity as follows:

$$E(\,\ln(\alpha_1 \eta_t^2 + \beta_1)) < 0. \tag{3.4}$$

Condition (3.4) allows $\alpha_1 + \beta_1$ to be 1, or slightly larger than 1, in which case $E\varepsilon_t^2 = \infty$. For the general model (3.1)-(3.2), the necessary and sufficient condition for strict stationarity and ergodicity was established by Bougerol and Picard (1992) and Nelson (1990). Ling and Li (1997c) proved that, under (3.3), there exists a unique $\mathcal{F}_t$-measurable and second-order stationary solution to model (3.1)-(3.2), and that the solution is strictly stationary and ergodic, with the following causal representation:

$$h_t = \alpha_0 + \sum_{j=1}^{\infty} c' \,(\prod_{i=1}^{j} A_{t-i})\, \xi_{t-j} \ a.s., \tag{3.5}$$

where $\xi_t = (\alpha_0 \eta_t, 0, \cdots, 0, \alpha_0, 0, \cdots, 0)_{(r+s)\times 1}$, with the first component $\alpha_0 \eta_t$ and $(r+1)$-th component α_0, $c = (\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s)'$, and

$$A_t = \left(\begin{array}{ccc|ccc} \alpha_1 \eta_t & \cdots & \alpha_r \eta_t & \beta_1 \eta_t & \cdots & \beta_s \eta_t \\ & I_{(r-1)\times(r-1)} & O_{(r-1)\times 1} & & O_{(r-1)\times s} & \\ \hline \alpha_1 & \cdots & \alpha_r & \beta_1 & \cdots & \beta_s \\ & O_{(s-1)\times r} & & & I_{(s-1)\times(s-1)} & O_{(s-1)\times 1} \end{array} \right). \tag{3.6}$$

Bollerslev (1986) provided the necessary and sufficient condition for the existence of the $2m$th moment of the GARCH(1,1) model, and the necessary and sufficient condition for the fourth-order moments of the GARCH(1,2) and GARCH(2,1) models. Using a similar method as in Bollerslev (1986), He and Teräsvirta (1999a) provided the moment conditions of a family of GARCH(1,1) models. Ling and

McAleer (1999d) derived the sufficient condition for the existence of the stationary solution of this family of GARCH(1,1) models, showed that He and Terävirta's (1999a) condition is necessary but not sufficient, and provided the sufficient moment condition. He and Teräsvirta (1999b) and Karanasos (1999) examined the fourth moment structure of the GARCH(p, q) process. From the proof in Karanasos (1999), it can be seen that the condition is necessary but not sufficient. He and Teräsvirta (1999b) stated that their condition is necessary and sufficient. Ling and McAleer (1999c) showed that the necessary condition for the existence of the fourth moment is incomplete, that the condition is not sufficient for the existence of the fourth moment, and also derived the necessary and sufficient conditions for the existence of all the moments.

Based on Theorem 2.1 in Ling and Li (1997c) and Theorem 2 in Tweedie (1988), Ling (1999b) showed that a sufficient condition for the existence of the $2m$th moment of model (3.1)-(3.2) is

$$\rho[E(A_t^{\otimes m})] < 1, \tag{3.7}$$

where $\rho(A) = \max\{\text{eigenvalues of a matrix } A\}$. Ling's result does not need to assume that the GARCH(r, s) process starts infinitely far in the past with finite $2m$th moment, as is required in Bollerslev (1986) and He and Teräsvirta (1999a, b), and has a far simpler form as compared with that of Milhøj (1985). Ling and McAleer (1999c) further showed that condition (3.7) is also necessary for the existence of the $2m$th moment. Thus, the moment structure of the GARCH(r, s) model in (3.1)-(3.2) has now been established completely. Bera, Higgins and Lee (1996) considered a random coefficient formulation of GARCH processes. An asymptotic theory for the sample autocorrelations and extremes of a GARCH(1,1) process is provided in Mikosch and Stărică (2000). As an extension of the GARCH(r, s) process, Ling and McAleer (1999c) also derived the necessary and sufficient moment conditions of the asymmetric power GARCH(r, s) model of Ding et al. (1993).

3.2 Quasi-Maximum Likelihood Estimation

The GARCH model is usually estimated by the quasi-maximum likelihood method. However, the properties of the QMLE are not completely clear. Consider the simple but important GARCH(1,1) model. In this case, the likelihood can be written as

$$L(\delta) = \sum_{t=1}^{n} l_t, \;\; l_t = -\frac{1}{2}\ln h_t - \frac{1}{2}\frac{\varepsilon_t^2}{h_t}, \tag{3.8}$$

where h_t is treated as a function of ε_t, and the parameter $\delta = (\alpha_0, \alpha_1, \beta_1)'$ and h_t are calculated through the following recursion:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, h_0 = \text{a positive constant}. \tag{3.9}$$

Lee and Hansen (1994) and Lumsdaine (1996) proved that the local QMLE is consistent and asymptotically normal, assuming that $E(\ln(\alpha_1\eta_t^2 + \beta_1)) < 0$, which is the necessary and sufficient condition for strict stationarity. However, Lee and Hansen (1994) required that all the conditional expectations of $\eta_t^{2+\kappa} < \infty$ uniformly with $\kappa > 0$, while Lumsdaine (1996) required that $E\eta_t^{32} < \infty$. In addition, Lee and Hansen (1994) showed that the global QMLE is consistent if ε_t is second-order stationary. Lee and Hansen (1994) and Lumsdaine (1996) stated that their methods are valid only for the simple GARCH(1,1) model and cannot be extended to more general cases.

For the general order GARCH(r, s) model, Ling and Li (1997b) proved that the local QMLE is consistent and asymptotically normal if $E\varepsilon_t^4 < \infty$. Based on uniform convergence as a modification of a theorem in Amemiya (1985, page 116), Ling and McAleer (1999b) proved the consistency of the global QMLE under only the second-order moment condition. They also derived the asymptotic normality of the global QMLE under the $6th$ moment condition.

When η_t is not normal, the QMLE is inefficient. Drost and Klaassen (1997) investigated adaptive estimation of the GARCH(1,1) model. This method was extended to nonstationary ARMA models with higher-order GARCH(r, s) errors by

Ling and McAleer (1999a). Francq and Zakoïan (2000) consider the estimation of weak GARCH representations (Drost and Nijman, 1993) characterized by an ARMA structure for the squared error terms.

4 Stationary ARMA-GARCH Models

The ARCH process is a non-independent white noise sequence, which first appeared in the regression model of Engle (1982). Engle's original motivation seems to have been that an ARCH structure provides improved statistical inference for the mean of the regression model, such as confidence intervals and forecasting. Over the last decade, there has been a tendency to employ the ARCH/GARCH model to analyze the volatilities of financial and economic data, while ignoring the specification and estimation of the conditional mean. However, if the conditional mean is not specified adequately, then it may not be possible to construct consistent estimates of the true ARCH process, for which statistical inference and empirical analysis regarding the ARCH component might be misleading. Thus, even though the primary interest might be on the volatilities in the data, the specification and estimation of the conditional mean are still important.

The conditional mean is typically given as an AR or ARMA model. However, since the conditional variances of the white noise are not constant, the generating mechanism of the AR or ARMA model is quite different from the traditional AR or ARMA model with i.i.d. errors, or martingale differences with a constant conditional variance. As a number of statistical properties of the traditional AR or ARMA model cannot be extended to the present case, it is necessary to have a thorough investigation of these types of models.

We define the ARMA-GARCH model by the following equations:

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t, \tag{4.1}$$

$$\varepsilon_t = \eta_t \sqrt{h_t},\ h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}. \tag{4.2}$$

There is no paper which is especially devoted to the ARMA-GARCH model, although it is a special case of Ling and Li (1997c, 1998) and Ling and McAleer (1999b). When $s = 0$, the ARMA-GARCH model reduces to the ARMA-ARCH model, which is a special case of the ARMA-ARCH model of Weiss (1986). When $q = 0$, $s = 0$ and $r = 1$, the AR-ARCH(1) model was investigated by Pantula (1988). The properties of the ARMA-GARCH model appear in Ling and Li (1997c). When all the roots of $\phi(z) = z^p - \sum_{i=1}^{p} \varphi_i z^{p-i}$ lie outside the unit circle, y_t is strictly stationary if ε_t is strictly stationary, and y_t is $2m$th order stationary if ε_t is $2m$th stationary. Thus, in this section, we consider estimation of only the ARMA-GARCH model.

The parameters in (4.1)-(4.2) consist of two sets: one set includes the parameters of the conditional mean, denoted by m, and another set includes the parameters of the conditional variance h_t, denoted by δ. In practice, m is first estimated and then the residuals from the estimated conditional mean are calculated. When the residuals have been obtained, δ can be estimated using the methods in Sections 2-3. Furthermore, the estimated h_t is used to obtain a more efficient estimator of m. If the density function of η_t is symmetric, the MLE of m and δ can be obtained through a separate iteration procedure without loss of asymptotic efficiency. The following section examines the estimation of m when δ is assumed to be known.

4.1 Least Squares Estimation

Denote the true value of m by m_0. Given observations $y_1, \cdots, y_n$, the LSE of m_0, $\hat{m}$, is defined as the values in Θ which minimize

$$S_n = \sum_{t=1}^{n} \varepsilon_t^2. \tag{4.3}$$

For the ARMA-ARCH model, Weiss (1986) showed that $\hat{m}$ is consistent for m_0 and

$$\sqrt{n}(\hat{m} - m_0) \longrightarrow_{\mathcal{L}} N(0, A), \tag{4.4}$$

with

$$A = E^{-1}\left[\frac{\partial \varepsilon_t}{m}\frac{\partial \varepsilon_t}{m'}\right] E\left[\varepsilon_t^2\frac{\partial \varepsilon_t}{m}\frac{\partial \varepsilon_t}{m'}\right] E^{-1}\left[\frac{\partial \varepsilon_t}{m}\frac{\partial \varepsilon_t}{m'}\right]_{m=m_0}.$$

Pantula (1989) also obtained the asymptotic distribution of the LSE for the AR model with ARCH(1) errors, and gave an explicit form for A. The results in Weiss (1986) and Pantula (1989) require that y_t has finite fourth moment. As yet, no one seems to have considered the LSE of m_0 for the ARMA-GARCH model. However, the result in Weiss (1986) for the LSE can be easily extended to the ARMA-GARCH model. When GARCH reduces to an i.i.d. white noise process, the LSE is equivalent to the MLE of m_0.

There is presently no asymptotic theory for the LSE of the ARMA-GARCH model when the fourth moment condition is not satisfied. From the results of Davis and Mikosch (1998), it would be expected that the LSE is inconsistent if the variance of ε_t is infinite, but is consistent but with a slower convergence rate than $\sqrt{n}$ if ε_t has finite variance and infinite fourth moment. In such cases, the results would be different from those in Davis and Resnick (1985, 1986).

4.2 Quasi-Maximum Likelihood Estimation

Although the LSE is consistent and asymptotically normal if the fourth moment is finite, it is inefficient for ARMA-ARCH/GARCH models. In such cases, it is standard to use MLE. The maximum likelihood method was first used by Engle (1982) for both the AR-ARCH model and a fixed design regression with ARCH errors. First, the log-likelihood function can be written as

$$L(m) = \sum_{t=1}^{n} l_t, \quad l_t = -\frac{1}{2}\ln h_t - \frac{1}{2}\frac{\varepsilon_t^2}{h_t}, \tag{4.5}$$

where h_t is treated as a function of y_t and m, and is calculated through the following recursion:

$$h_t = \alpha_0 + \sum_{i=1}^{r}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{s}\beta_i h_{t-i}, \quad h_0 = \text{a positive constant}. \tag{4.6}$$

Define $\hat{m} = \max_{m\in\Theta} L(m)$. Since η_t is not assumed to be normal, $\hat{m}$ is referred to as the QMLE of m. For the ARMA-ARCH model, Weiss (1986) showed that the QMLE is consistent and asymptotically normal under a finite fourth moment condition. From Ling and Li (1997c), there exists a locally consistent and asymptotically normal QMLE for the ARMA-GARCH model if it has finite fourth moment. When η_t is normal, the asymptotic covariance matrix of $\sqrt{n}(\hat{m} - m_0)$ is

$$B = E\left[\frac{1}{h_t}\frac{\partial \varepsilon_t}{\partial m}\frac{\partial \varepsilon_t}{\partial m'} + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial m}\frac{\partial h_t}{\partial m'}\right]^{-1}_{m=m_0}. \tag{4.7}$$

Engle (1982) demonstrated that the MLE is more efficient than the LSE through a simple fixed design regression model and a first-order ARCH process. Pantula (1989) also showed that the MLE is more efficient than the LSE for the AR model with ARCH(1) errors. In fact, it can be shown that $A \geq B$ for the general ARMA-GARCH case.

Under the existence of the second moment, Ling and McAleer (1999b) showed that the global QMLE is consistent. However, in order to derive the asymptotic normality of the global QMLE, the model must satisfy the sixth moment condition. For the ARMA-GARCH$(1, q)$ model, it is possible to show that the global QMLE of m_0 is consistent and asymptotically normal, even if the fourth moment condition is not satisfied.

4.3 Adaptive Estimation

The QMLE of m_0 in the stationary ARMA-GARCH model is efficient only if η_t is normal. When η_t is not normal, adaptive estimation is useful for obtaining efficient estimators. A comprehensive account of the theory and method of adaptive estimation can be found in Bickel (1982) and Bickel, Klaassen, Ritov and Wellner (1993), with valuable surveys available in Robinson (1988) and Stoker (1991).

In the time series context, Kreiss (1987a) investigated the stationary ARMA model with i.i.d. errors. He proved the local asymptotic normality (LAN) property of the model and constructed adaptive estimators of m_0. Unlike Bickel (1982), Kreiss' adaptive procedure avoids the split sample technique, and hence is quite useful for practical applications. Jeganathan (1995) and Koul and Schick (1996) constructed adaptive estimators without splitting the sample for some nonlinear AR time series with i.i.d. noise. Koul and Schick (1996) also showed through simulation that the adaptive estimator without splitting the sample is superior to those based on the split sample technique.

Lee and Tse (1991) and Engle and González-Rivera (1991) are among the first to have used a semiparametric approach for models (4.1)-(4.2), but they did not obtain any theoretical results. Koul and Schick (1996) investigated adaptive estimation for a random coefficient AR model, which is an ARCH-type time series model. Jeganathan (1995) and Drost, Klaassen and Werker (1997) developed general frameworks suitable for stationary ARCH-type times series. The results in Ling and McAleer (1999a) include the development of the adaptive method for stationary ARMA-GARCH models and the conditions required for adaptive estimation.

5 Nonstationary ARMA-GARCH Models

Nonstationary time series have now been extensively investigated for the last two decades. Some important results for nonstationary AR models can be found in Fuller (1976), Dickey and Fuller (1979), Phillips (1987), Chan and Wei (1987, 1988), Tsay and Tiao (1990) and Jeganathan (1995), among many others. However, research on nonstationary time series is almost always limited to innovations with constant conditional variances. Under the framework of Phillips and Durlauf (1986) and Phillips (1987), the long-run variance and the innovation variances are equal in the presence of heteroscedasticity, but it does not include conditional heteroscedastic processes as defined in (3.1)-(3.2).

The ARMA-GARCH model is called nonstationary if the characteristic polynomial $\phi(z)$ has a root on the unit circle. Consider the simple AR(1) case:

$$y_t = \phi y_{t-1} + \varepsilon_t \tag{5.1}$$

where $\phi = 1$, and ε_t follows the GARCH(1, 1) process, that is,

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \tag{5.2}$$

When $\beta_1 = 0$, in which case ε_t follows a first-order ARCH process, Pantula (1989) derived the asymptotic distribution of the LSE of the unit root under the fourth moment condition. Ling and Li (1997b) obtained the same result under the second moment condition, namely $\alpha_1 + \beta_1 < 1$. The asymptotic distribution is

$$n(\hat{\phi}_{LS} - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 B(t)dB(t)}{\int_0^1 B^2(t)dt},$$

where $\hat{\phi}_{LS} = (\sum_{t=1}^n y_{t-1}^2)^{-1}(\sum_{t=1}^n y_t y_{t-1})$ and $B(t)$ is a standard Brownian motion. Thus, the Dickey-Fuller test statistic can still be used. However, Peters and Veloce (1988) and Kim and Schmidt (1993) provided simulation results showing that Dickey-Fuller tests based on the LSE are generally not robust.

It should be noted that, for stationary ARMA-GARCH models, the QMLE is more efficient than the LSE. It seems natural to expect this advantage to extend to nonstationary time series, in which case unit root tests based on the MLE in the presence of ARCH/GARCH innovations should be useful. According to standard statistical theory, an efficient estimator will often provide locally most powerful tests [e.g. see Rao(1973, Chapter 7)]. For this reason, unit root tests based on QMLE would be expected to be more powerful than those based on LSE.

Note that Leybourne, McCabe and Tremayne (1996) observed that heteroscedasticity will be present automatically if ϕ is actually a random variable fluctuating about 1. They developed a score test for such a randomized unit root.

5.1 Quasi-Maximum Likelihood Estimation

In this section, we assume that the characteristic polynomial $\phi(z)$ has only a unit root of +1. The general case was investigated in Ling and Li (1998). Since $\varphi(z)$ has a unit root, it can be decomposed as $(1-z)\phi(z)$, where $\phi(z) = 1 - \sum_{i=1}^{p-1} \phi_i z^i$. Let $w_t = (1-B)y_t$, where B is the backshift operator. Model (4.1) can be rewritten as

$$y_t = \gamma y_{t-1} + w_t, \;\; w_t = \sum_{i=1}^{p-1} \phi_i w_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i}, \tag{5.3}$$

where $\gamma = 1$ and ε_t is defined by (4.2). The parameters in model (5.3) are γ and $m = (\phi', \psi')'$, where $\phi = (\phi_1, \cdots, \phi_{p-1})'$ and $\psi = (\psi_1, \cdots, \psi_q)'$. As in the stationary case, we assume that the parameters in (4.2) are known or can be estimated consistently.

Given the observations $y_1, \cdots, y_n$, with initial values $y_i = 0$, or some constants, for $i \leq 0$, the log-likelihood function can be written as

$$L(\lambda) = \sum_{t=1}^{n} l_t, \;\; l_t = -\frac{1}{2} \ln h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}, \tag{5.4}$$

where $\lambda = (\gamma, m')'$, and h_t is treated as a function of y_t and λ. Ling and Li (1998) showed that there exists a locally consistent QMLE such that

$$G_n^{-1}(\hat{\lambda} - \lambda) \longrightarrow_{\mathcal{L}} (\xi_{ML}, N')' \;, \tag{5.5}$$

where

$$\xi_{ML} = \frac{c \int_0^1 w_1(t) dw_2(t)}{F \int_0^1 w_1^2(t) dt}, \tag{5.6}$$

$c = [1 - \phi(1)]^{-1}$, N is a normal random vector independent of ξ_{ML}, F is a constant depending on the GARCH parameters, $\kappa = E\eta_t^4 - 1$, and $(w_1(t), w_2(t))$ is a bivariate Brownian motion with covariance $t\Omega$. When $r = s = 1$,

$$\Omega = \begin{pmatrix} Eh_t & 1 \\ 1 & E(1/h_t) + \kappa\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2 / h_t^2) \end{pmatrix}, \tag{5.7}$$

and when η_t is normal, $\kappa = 2$ and $F = E(1/h_t) + 2\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2 / h_t^2)$. For higher-order GARCH models, the structure of Ω can be found in Ling and Li (1998). Note also that, unlike the least squares case, the moving average parameters do not appear in (5.6) and (5.7).

The above results were derived under the fourth moment condition in Ling and Li (1998). Furthermore, under the second moment condition, Ling and Li (1997b) derived the same result for models (5.1)-(5.2) when $c = 1$. If the second moment condition is not satisfied, the asymptotic distribution for the LSE or QMLE of the unit root is as yet unknown. For the unit root process with i.i.d. errors having infinite variance and in the domain of attraction of an α-stable law, Chan and Tran (1989) and Chan (1990) showed that $n^{-1/\alpha}(\hat{\phi}_{LS} - 1)$ converges to a functional of a Levy process with $\alpha \in (0, 2)$. It is conjectured that there is a similar asymptotic distribution for the LSE or QMLE of the unit root when the GARCH noise has an infinite variance.

5.2 Unit Root Tests Based on QMLE

The asymptotic distribution for the QMLE of the unit root can be used to construct a unit root test. For simplicity, we consider only models (5.1)-(5.2). Denote $\tilde{\phi}_{ML}$ as the QMLE of ϕ, and let

$$B_1(t) = \frac{1}{\sigma} w_1(t) \text{ and } B_2(t) = -\frac{1}{\sigma^2}\sqrt{\frac{\sigma^2}{\sigma^2 K - 1}} w_1(t) + \sqrt{\frac{\sigma^2}{\sigma^2 K - 1}} w_2(t),$$

where $\sigma^2 = Eh_t$ and K is the (2,2)th element of Ω. Then $B_1(t)$ and $B_2(t)$ are two independent standard Brownian motions. As shown in Ling and Li (1998),

$$n(\tilde{\phi}_{ML} - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 B_1(t)dB_1(t)}{\sigma^2 F \int_0^1 B_1^2(t)dt} + \frac{\sqrt{\sigma^2 K - 1}}{\sigma^2 F} \frac{\int_0^1 B_1(t)dB_2(t)}{\int_0^1 B_1^2(t)dt}. \tag{5.8}$$

The second term in (5.8) can be simplified to $[\sqrt{\sigma^2 K - 1}/F\sigma^2]\ (\int_0^1 B_1^2(t)\ dt)^{-1/2}\xi$, where ξ is a standard normal random variable independent of $\int_0^1 B_1^2(t)dt$ (see Phillips, 1989). Thus,

$$n(\hat{\phi}_{ML} - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 B_1(t)dB_1(t)}{\sigma^2 F \int_0^1 B_1^2(t)dt} + \frac{\sqrt{\sigma^2 K - 1}}{\sigma^2 F}(\int_0^1 B_1^2(t)dt)^{-1/2}\xi. \tag{5.9}$$

From (5.8)-(5.9), we see that the asymptotic distribution of $\hat{\phi}_{ML}$ can be represented as a combination of the asymptotic distribution of $\hat{\phi}_{LS}$ and a scale mixture of normals. This property is similar to that of the least absolute deviation estimator of unit roots given in Herce (1996). Ling and Li (1998) showed that the QMLE of ϕ is more efficient than the LSE.

As the asymptotic distribution in (5.9) includes nuisance parameters, we cannot use it directly to test for a unit root. There are two methods to overcome this difficulty. The first is to combine the LSE and QMLE to construct a unit root test, as in Ling and Li (1997b). Let

$$L_\phi = n(\hat{\phi}_{LS} - 1), \quad L_t = (\frac{1}{n^2}\sum_{t=1}^{n} y_{t-1}^2)^{1/2} L_\phi,$$

where $\bar{y} = n^{-1}\sum_{t=1}^{n} y_{t-1}$. Furthermore, define

$$M_\phi = \frac{\hat{\sigma}^2 \hat{F}}{\sqrt{\hat{\sigma}^2 \hat{K} - 1}} \{n(\hat{\phi}_{ML} - 1) - (\hat{F}\hat{\sigma}^2)^{-1}[n(\hat{\phi}_{LS} - 1)]\},$$
$$M_t = (\frac{1}{n^2}\sum_{t=1}^{n} y_{t-1}^2)^{1/2} M_\phi.$$

Ling and Li (1997b) showed that

$$M_\phi \xrightarrow{\mathcal{L}} [\int_0^1 B_1^2(t)dt]^{-1/2}\xi \text{ and } M_t \xrightarrow{\mathcal{L}} \xi,$$

where ξ is a standard normal random variable independent of $\int_0^1 B_1^2(t)dt$.

The limiting distributions of M_ϕ and M_t are the same as those based on the least absolute deviations estimators of Herce (1996). However, the test statistics themselves are quite different. Empirical critical values of these distributions were reported in Ling, Li and McAleer (1999), who showed that M_ϕ and M_t can overcome the excessive sizes, as reported in Peters and Veloce (1988) and Kim and Schmidt (1993), and have power comparable to that of the Dickey-Fuller test.

Another method of overcoming the presence of nuisance parameters is to construct a unit root test without using the LSE, as used in Seo (1999). First, rewrite (5.9) as

$$nc_1(\tilde{\phi}_{ML} - 1) \xrightarrow{\mathcal{L}} \frac{\rho\int_0^1 B_1(t)dB_1(t)}{\int_0^1 B_1^2(t)dt} + \sqrt{1-\rho^2}\frac{\int_0^1 B_1(t)dB_2(t)}{\int_0^1 B_1^2(t)dt}, \tag{5.10}$$

where $c_1 = \dfrac{\sigma F}{\sqrt{K}}$ and $\rho^2 = 1/(\sigma^2 K) \in (0, 1)$. The t-statistic is then given by

$$nc_2(\frac{1}{n^2}\sum_{t=1}^{n} y_{t-1}^2)^{1/2}(\tilde{\phi}_{ML} - 1) \xrightarrow{\mathcal{L}} \frac{\rho\int_0^1 B_1(t)dB_1(t)}{(\int_0^1 B_1^2(t)dt)^{-1/2}} + \sqrt{1-\rho^2}\frac{\int_0^1 B_1(t)dB_2(t)}{(\int_0^1 B_1^2(t)dt)^{-1/2}}, \tag{5.11}$$

where $c_2 = c_1/\sigma$. Seo (1999) tabulated the limiting distribution in (5.11) for different values of ρ. The simulation results in Seo (1999) showed that the unit root test based on (5.11) not only overcomes the size distortion problem, but is also consistently more powerful than tests based on the LSE. These results confirm the expectation that more efficient estimates of unit roots yield more powerful unit root tests.

When the conditional errors η_t are not normal, the estimator of the unit root is not efficient. Ling and McAleer (1999c) investigated adaptive estimation of the non-stationary ARMA model with GARCH errors. They obtained the locally asymptotic quadratic form of the log-likelihood ratio, and showed that it was neither locally asymptotic normal nor locally asymptotic mixed normal. A new efficiency criterion was given for a class of defined M-estimators. When the conditional error density is known, Ling and McAleer (1999c) showed that efficient estimators can be constructed using the kernel estimator for the score function. It is also shown that the adaptive procedure for the parameters in the conditional mean part uses the full sample.

6 Other ARCH-type Models

In this section, some other ARCH-type models are considered, namely double threshold ARCH, ARFIMA-GARCH, CHARMA, and vector ARMA-GARCH.

6.1 Double Threshold ARCH Models

Given the success of Tong's (1978, 1980) threshold model in nonlinear time series, it is natural to consider threshold structures for the conditional variance specification. The use of thresholds to model asymmetries is supported by well known empirical characteristics as to the likely asymmetric behaviour of volatility in the stock market (see, for example, French et al. (1987)).

Li and Li (1996) proposed the double threshold AR conditional heteroskedastic

(DTARCH) time series model:

$$
\begin{aligned}
y_t &= \phi_0^{(j)} + \sum_{i=1}^{p_i} \phi_i^{(j)} y_{t-i} + \varepsilon_t, \quad a_{j-1} < y_{t-b} \leq a_j, & (6.1)\\
\varepsilon_t &= \eta_t h_t^{\frac{1}{2}}, & (6.2)\\
h_t &= \alpha_0^{(k)} + \sum_{i=1}^{r_k} \alpha_i^{(k)} \varepsilon_{t-i}^2, \;\; c_{k-1} < y_{t-d} \leq c_k, & (6.3)
\end{aligned}
$$

where $j = 1, \cdots, v_1$; $k = 1, \cdots, v_2$; and b and $d \geq 1$ are the delay parameters. In (6.1)-(6.3), the threshold parameters satisfy $-\infty = a_0 < a_1 < \cdots < a_{v_1} = \infty$ and $-\infty < c_0 < c_1 < \cdots < c_{v_2} = \infty$, $\phi_i^{(j)}$ and $\alpha_i^{(k)}$ are constants, $\alpha_0^{(k)} > 0$ and $\alpha_i^{(k)} \geq 0$. The model generalizes the threshold AR model of Tong (1978, 1980) to include a threshold ARCH component. Tong (1990) referred to this type of hybrid model as a second generation model. Note that other indicator variables may be used in place of y_{t-b} and y_{t-d}. The threshold variables are typically defined as a linear combination of the lagged values of the observed process, but van Dijk, Teräsvirta and Franses (2000) relaxed this definition of threshold variables to include non-linear combinations of the lags of the observed process as well as of other variables. Li and Lam (1995) combined the threshold autoregressive model with a fixed ARCH specification in studying the asymmetry of a stock index. Extension to a double-threshold GARCH model was considered by Brooks (2001).

Ling (1999b) showed that, if $\sum_{i=1}^{p} \max_j |\phi_i^{(j)}| < 1$ and $\sum_{i=1}^{r} \max_k \alpha_i^{(k)} < 1$, then there exists a strictly stationary solution $\{y_t, \varepsilon_t\}$ satisfying models (6.1)-(6.3), and $E_{\pi_1}(|y_t|)$ and $E_{\pi_2}(\varepsilon_t^2)$ are finite, where π_1 and π_2 are the stationary distributions of $\{y_t\}$ and $\{\varepsilon_t\}$, respectively. However, the uniqueness and ergodicity conditions are as yet unknown. If the second threshold, $c_{k-1} < y_{t-d} \leq c_k$, is replaced by $c_{k-1} < \varepsilon_{t-d} \leq c_k$, the strict stationarity and ergodicity condition has been obtained by Liu , Li and Li (1997).

Under the assumption that y_t is strictly stationary and ergodic, and the threshold parameters a_i and c_i are known, Li and Li (1996) proved that the MLE is consistent and asymptotically normal. In practice, the threshold parameters a_i and c_i are unknown and can be estimated by the maximum likelihood method. However, the asymptotic distributions of the estimators are as yet unknown. For the threshold AR model with i.i.d. errors, Chan (1993) showed that the estimator of the threshold parameter has a convergence rate of n and an asymptotic distribution associated with the compound Poisson process. This method could possibly be used for the DTARCH model.

Pesaran and Potter (1997) considered a floor and ceiling model of US output which may be interpreted as a double threshold ARCH model. Rabemanjara and Zakoïan (1993) examined an asymmetric ARCH model which may be regarded as a special case of the DTARCH model. Fornari and Mele (1997) considered a similar formulation to handle asymmetry in volatility. Lee and Li (1998) developed a smooth transition double threshold model. Lundbergh and Teräsvirta (1998a) used a double smooth AR-GARCH model to analyse some high-frequency exchange rate data. Wong and Li (1997) considered tests for the presence of autoregression under ARCH, while Wong and Li (1999) examined tests for the null of AR-ARCH against the double threshold ARCH model.

In the spirit of threshold nonlinear models Wong and Li (2000), Wong and Li (2001a, b) considered mixtures of autoregressive models and mixtures of autoregressive models with ARCH. Some interesting features of these types of models are that some components of the mixture can be non-stationary while the entire series can be stationary, the predictive distributions can be multimodal, and it is fairly easy to derive the conditions for stationarity and expressions for the autocorrelations.

6.2 Fractional ARIMA Models

Let $\{y_t\}$ satisfy

$$\phi(B)(1-B)^d(y_t-\mu)=\theta(B)\varepsilon_t, \tag{6.4}$$

$$\varepsilon_t \mid F_{t-1} \sim N(0,h_t),\ \ h_t=\alpha_0+\sum_{i=1}^{r}\alpha_i\varepsilon_{t-i}^2+\sum_{i=1}^{s}\beta_i h_{t-i}, \tag{6.5}$$

where $(1-B)^d$ is defined by the binomial series:

$$(1-B)^d = \sum_{k=0}^{\infty} \frac{(k+d-1)!}{k!(d-1)!} B^k. \tag{6.6}$$

The specifications in (6.4)-(6.5) are referred to as the fractional ARIMA-GARCH or equivalently the ARFIMA-GARCH model, which was investigated by Ling and Li (1997c). Baillie, Chung and Tieslau (1995) considered a fractional ARIMA$(0,d,1)$-GARCH(1,1) model for the CPI series of 10 different countries. Note that exact maximum likelihood estimation of (6.4) with h_t = a constant has been considered as early as in 1981 in the University of Western Ontario Ph.D. Thesis by W.K. Li.

Sufficient conditions for stationarity, ergodicity and the existence of higher-order moments of the fractional ARIMA model were derived by Ling and Li (1997c). Under some mild conditions, it is shown that the MLE is locally consistent and asymptotically normal. It is well known that, when $p = q = 0$ so that $(1-B)^d y_t = \varepsilon_t$, the MLE of d converges to $N(0, 6/\pi^2)$ in distribution if ε_t is i.i.d. (see Li and McLeod, 1986). However, when ε_t is a GARCH process, Ling and Li (1997c) showed that the asymptotic variance is

$$\Omega_\gamma = E[\frac{1}{h_t}(\frac{\partial \varepsilon_t}{\partial d})^2 + \frac{1}{2h_t^2}(\frac{\partial h_t}{\partial d})^2],$$

which is no longer independent of d and is less than $6/\pi^2$. Ling and Li (1997c) also examined the large sample distributions of the residual autocorrelations and the squared-residual autocorrelations, and two portmanteau test statistics. Robinson (1991) considered tests for conditional heteroskedasticity in long memory processes. More recently, Beran and Feng (1999) considered local polynomial estimation of a fractional ARIMA model similar to the above.

6.3 CHARMA Models

Tsay (1987) proposed the conditional heteroskedastic autoregressive moving average (CHARMA) model, given by:

$$y_t - \mu = \sum_{i=1}^{p} \psi_i (y_{t-i} - \mu) + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{6.7}$$

$$\varepsilon_t = \sum_{i=1}^{r} \delta_{it}\varepsilon_{t-i} + \sum_{i=1}^{s} w_{it}(y_{t-i} - \mu) + w_{0t}(\hat{y}_{t-1}(1) - \mu) + e_t, \tag{6.8}$$

where the orders p, q, r and s are finite and non-negative integers; μ, ψ_i and θ_i are constant; δ_{it}, w_{it} and e_t are random variables; and $\hat{y}_{t-1}(1) = E(y_t|\mathcal{F}_{t-1})$, where $\mathcal{F}_{t-1}$ is the $\sigma-$field generated by $\{e_{t-i}, w_{t-i}, \delta_{t-i}|i = 1, 2, \cdots\}$, $w_t = (w_{0t}, w_{1t}, \cdots, w_{st})'$, and $\delta_t = (\delta_{1t}, \cdots, \delta_{rt})'$.

The LSE method can be used to estimate the parameters in (6.7). Tsay (1987) proved that the LSE is consistent if $E\varepsilon_t^4 < \infty$, and is asymptotically normal if $E\varepsilon_t^8 < \infty$. Since the model is an extension of the random coefficient AR model, the asymptotic MLE results can be obtained using the method in Nicholls and Quinn (1982). Basic properties such as strict stationarity, ergodicity and the moment structure are given in Ling (1999a).

The CHARMA model has been extended to the multivariate case. Wong and Li (1997) considered a stationary multivariate CHARMA model, and Li, Ling and Wong (1999) investigated a partially nonstationary AR model with conditional heteroscedasticity, as follows:

$$Y_t = \Phi_1 Y_{t-1} + \cdots + \Phi_p Y_{t-p} + \varepsilon_t \tag{6.9}$$

and

$$\varepsilon_t = \alpha_{1t}\varepsilon_{t-1} + \cdots + \alpha_{qt}\varepsilon_{t-q} + e_t, \tag{6.10}$$

where the Φ_i are constant matrices; $\det\{\Phi(z)\} = |I - \Phi_1 z - \cdots - \Phi_p z^p| = 0$ has $d \leq m$ unit roots and other roots outside the unit circle; rank$[\Phi(1)] = m - d$; $\delta_t = (\alpha_{1t}, \cdots, \alpha_{qt})$ is a sequence of i.i.d. matrices with mean zero and nonnegative covariance $E[vec(\delta_t)vec'(\delta_t)] = \Omega$; and e_t is an i.i.d. random vector with mean zero and positive covariance $E(e_t e_t') = G$.

Under the condition for the finite fourth moment, Li, Ling and Wong (1998) derived the asymptotic distributions of the LSE, a full rank MLE, and a reduced rank MLE. When the multivariate ARCH process reduces to the innovation with a constant covariance matrix, these asymptotic distributions are the same as in Ahn and Reinsel (1990). However, in the presence of multivariate ARCH innovations, the asymptotic distributions of the full rank MLE and the reduced rank MLE involve two correlated multivariate Brownian motions, which are different from those given in Ahn and Reinsel (1990). The asymptotic results in Li, Ling and Wong (1998) can be used to construct cointegration tests based on the MLE.

6.4 Vector ARMA-GARCH Models

Ling and McAleer (1999b) proposed the vector ARMA-GARCH model:

$$\Phi(B)(Y_t - \mu) = \Psi(B)\varepsilon_t, \tag{6.11}$$

$$\varepsilon_t = D_t^{1/2}\eta_t, \quad H_t = W + \sum_{i=1}^{r} A_i \tilde{\varepsilon}_{t-i} + \sum_{i=1}^{s} B_i H_{t-i}, \tag{6.12}$$

where $D_t = diag(h_{1t}, \cdots, h_{mt})'$, $H_t = (h_{1t}, \cdots, h_{mt})'$, $\Phi(B) = I - \Phi_1 B - \cdots - \Phi_p B^p$ and $\Psi(B) = I + \Psi_1 B + \cdots + \Psi_q B^q$ are polynomials in B, $\tilde{\varepsilon}_t = (\varepsilon_{1t}^2, \cdots, \varepsilon_{mt}^2)'$, and $\eta_t = (\eta_{1t}, \cdots, \eta_{mt})'$ is a sequence of i.i.d. random vectors with mean zero and covariance Γ.

Ling and McAleer (1999b) obtained the conditions for strict stationarity and ergodicity, and the higher-order moments of the model. The consistency of the global QMLE is proved under the existence of only the second-order moment. In order to derive the asymptotic normality of the global QMLE, the results require the second moment condition for the vector ARCH model, the fourth moment condition for the vector ARMA-ARCH model, and the sixth moment condition for the vector ARMA-GARCH model.

7 Conclusion

Most of the theoretical results for GARCH-type processes require that the fourth- or higher-order moments exist. In practice, this condition may not be satisfied. When the fourth moment of the GARCH process is infinite, it exhibits the feature of heavy tails. At present, a theory is lacking for ARMA models derived from this type of GARCH specification, even for ARMA models with i.i.d. heavy-tailed noise (see Resnick (1997)). Since heavy-tailed phenomena are often encountered in finance and economics, an analysis of data exhibiting heavy tails would seem to be an important direction for future research.

Although there have been many contributions to the ARCH/GARCH literature, it seems that until recently very little attention has been paid to model selection. Apart from the diagnostic checking method of Li and Mak (1994) and its extension by Ling and Li (1997a), there would seem to be few formal tools for checking model adequacy. Tse and Zuo (1997) provided a simulation study of the Li–Mak test. More recently, Lundbergh and Teräsvirta (1998b) showed that the Li–Mak test is equivalent to a Lagrange multiplier test of no residual ARCH. Tse (1999) provides a recent review of this literature. A generalization of Li and Mak (1994) is obtained by Horvath and Kokoszka (2001). A robustified version of Li and Mak (1994) against outliers is developed by Jiang, Shao and Hui (2001). All order selection methods for ARMA models, such as those in Hannan (1980), Potscher (1983, 1989), Tsay (1984), and Wei (1992), require that the error processes are i.i.d. or martingale differences with $\sup_t E(\varepsilon_t^2|\mathcal{F}_{t-1}) \leq$ a constant. However, ARCH-type models generally do not satisfy these conditions. It is important to develop a theory for order selection of ARCH, GARCH and ARMA-GARCH models, with Wong and Li (1996) and An, Fong and Li (1999) being two useful attempts in this direction.

REFERENCES

Ahn, S.K. and Reinsel, G.C. (1990) Estimation for partially nonstationary multivariate models. *Journal of the American Statistical Association* **85**, 813-823.

Amemiya, T. (1985) *Advanced Econometrics.* Cambridge, Harvard University Press.

An, H.Z., Fong, P.W. and Li, W.K. (1999) An approach to modelling subset multivariate ARCH model via the AIC principle. Technical Report 223, Department of Statistics and Actuarial Science, The University of Hong Kong.

Baillie, R.T., Chung, C.F. and Tieslau, M.A. (1995) Analyzing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of Applied Econometrics* **11**, 23-40.

Bera, A.K. and Higgins, M.L. (1993) ARCH models: Properties, estimation and testing. *Journal of Economic Surveys* **7**, 305-366; reprinted in L. Oxley et al. (eds.), *Surveys in Econometrics*, Blackwell, Oxford, 1995, pp. 215-272.

Bera, A.K., Higgins, M.L. and Lee, S. (1996) Random coefficient formulation of conditional heteroskedasticity and augmented ARCH models. *Sankhyā B* **58**, 199-220.

Beran, J. and Feng, Y. (1999) Local polynomial estimation with a FARIMA-GARCH error process. Manuscript, University of Konstanz.

Bickel, P.J. (1982). On adaptive estimation. *Annals of Statistics*, **10**, 647-671.

Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J. A. (1993) *Efficient and Adaptive Estimation for Semiparametric Models.* Baltimore: Johns Hopkins University Press.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

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 .

Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994) ARCH models, in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics* **4**, North-Holland, Amsterdam, pp. 2961-3038.

Bougerol, P. and Picard, N.M. (1992) Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**, 115-127.

Brooks, C. (2001) A double-threshold GARCH model for the French France/Deutschmark exchange rate. *Journal of Forecasting* **20** 135-143.

Chan, K.S. (1993) Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics* **21**, 520-533.

Chan, N.H. (1990) Inference for near-integrated time series with infinite variance. *Journal of the American Statistical Association* 85, 1069-1074.

Chan, N.H. and Tran, L.T. (1989) On the first order autoregressive processes with infinite variance. *Econometric Theory* **5**, 354-362.

Chan, N.H. and Wei, C.Z. (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* **15**, 1050-1063.

Chan, N.H. and Wei, C.Z. (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* **16**, 367-401.

Davis, R.A. and Mikosch, M. (1998) The sample autocorrelations of heavy-tailed processs with applications to ARCH. *Annals of Statistics* **26**, 2049-2080.

Davis, R.A. and Resnick, S. (1985) More limit theory for the sample correlation function of moving averages. *Stochastic Processes and Applications* **20**, 257-279.

Davis, R.A. and Resnick, S. (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* **14**, 533-558.

de Vries C.G. (1991) On the relation between GARCH and stable processes. *Journal of Econometrics* **48**, 313-324.

Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**, 427-431.

van Dijk, D., Teräsvirta, T. and Franses, P.H. (2000) Smooth transition autoregressive models - a survey of recent developments. To appear in *Econometric Reviews.*

Ding, Z., Granger, C.W.J. and Engle, R.F. (1973) A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**, 83-106.

Drost, F.C. and Klaassen, C.A.J. (1997) Efficient estimation in semiparametric GARCH models. *Journal of Econometrics* **81**, 193-221.

Drost, F.C., Klaassen, C.A.J. and Werker, B.J.M. (1997) Adaptive estimation in time series models. *Annals of Statistics* **25**, 786-817.

Drost, F.C. and Nijman, T. (1993) Temporal aggregation of GARCH processes. *Econometrica* **61**, 909-927.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987-1008.

Engle, R. F. and González-Rivera, G.-R. (1991) Semiparametric ARCH models. *Journal of Business & Economic Statistics* **9**, 345-359.

Engle, R.F. and Kroner, K.F. (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**, 122-150 .

Fornari, F. and Mele, A. (1997) Sign-and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics* **12**, 49-65.

Francq, C. and Zakoïan, J.-M. (2000) Estimating weak GARCH representations. *Econometric Theory* **16**, 692-728.

French, K.R., Schwert, G.W. and Stambaugh, R.F. (1987) Expected stock returns and volatility. *Journal of Financial Economics* **19**, 3-30.

Fuller, W.A. (1976) *Introduction to Statistical Time Series.* Wiley, New York.

Gourieroux, C. (1997) *ARCH Models and Financial Applications.* Springer-Verlag: New York.

Hannan, E.J. (1980) The estimation of the order of an ARMA model. *Annals of Statistics* **8**, 1071-1081.

He, C. and Teräsvirta, T. (1999a) Properties of moments of a family of GARCH processes. *Journal of Econometrics* **92**, 173-192.

He, C. and Teräsvirta, T. (1999b) Fourth moment structure of the GARCH(p, q) process. *Econometric Theory* **15**, 824-846.

Herce, M.A. (1996) Asymptotic theory of LAD estimation in a unit root process with finite variance errors. *Econometric Theory* **12**, 129-53.

Horvath, L. and Kokoszka, P. (2001) Large sample distribution of weighted sum of ARCH(p) squared residual correlation. *Econometric Theory* **17**, 283-295.

Jeganathan, P. (1995) Some aspects of asymptotic theory with applications to time series models. *Econometric Theory* **11**, 818-887.

Jeganathan, P. (1997) On asymptotic inference in linear cointegrated time series systems. *Econometric Theory* **13**, 692-745.

Jiang, J., Zhao, Q. and Hui, Y.V. (2001) Robust modelling of ARCH models. *Journal of Forecasting* **20**, 111-133.

Karanasos, M. (1999) The second moment and the autocovariance function of the squared errors of the GARCH model. *Journal of Econometrics* **90**, 63-76.

Kim, K. and Schmidt, P. (1993) Unit root tests with conditional heteroskedasticity. *Journal of Econometrics* **59**, 287-300.

Koul, H.L. and Schick, A. (1996) Adaptive estimation in a random coefficient autoregressive model. *Annals of Statistics* **24**, 1025-1052.

Kreiss, J.-P. (1987) On adaptive estimation in stationary ARMA processes. *Annals of Statistics* **15**, 112-133.

Lee, S.-W. and Hansen, B.E. (1994) Asymptotic theory for the GARCH (1,1) quasi-maximum likelihood estimator. *Econometric Theory* **10**, 29-52.

Lee, T.K.Y. and Tse, Y.K. (1991) Term structure of interest rates in the Singapore Asian dollar market. *Journal of Applied Econometrics* **6**, 143-152.

Lee, Y.N. and Li, W.K. (1998) On smooth transition double threshold models. In *Statistics and Finance: An Interface* (editors, W.S. Chan, W.K. Li and H. Tong) pp. 205-225. Imperial College Press: London.

Leybourne, S.J., McCabe, B.P.M. and Tremayne, A.R. (1996) Can economic time series be differenced to stationarity? *Journal of Business and Economic Statistics* **14**, 435-446.

Li, C.W. and Li, W.K. (1996) On a double threshold autoregressive heteroskedasticity time series model. *Journal of Applied Econometrics* **11**, 253-74.

Li, W.K. and Lam, K. (1995) Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model. *The Statistician* **44**, 333-341.

Li, W.K., Ling, S. and Wong, H. (1998) Estimation for partially nonstationary multivariate autoregressive models with conditional heteroskedasticity. Technical report, Department of Statistics and Actuarial Science, The University of Hong Kong.

Li, W.K. and Mak, T.K. (1994) On the squared residual autocorrelations in nonlinear time series with conditional heteroscedasticity. *Journal of Time Series Analysis* **15**, 627-636.

Li, W.K. and McLeod, A.I. (1986) Fractional time series modelling. *Biometrika* **73**, 217-21.

Ling, S. (1999a) On the stationarity and the existence of moments of conditional heteroskedastic ARMA models. *Statistica Sinica* **9**, 1119-1130.

Ling, S. (1999b) On the probabilistic properties of a double threshold ARMA conditional heteroskedasticity model. *Journal of Applied Probability* **36**, 1-18.

Ling, S. and Li, W.K. (1997a) Diagnostic checking nonlinear multivariate time series with multivariate ARCH errors. *Journal of Time Series Analysis* **18**, 447-464.

Ling, S. and Li, W.K. (1997b) Estimating and testing for unit root processes with GARCH(1,1) errors. Technical report, Department of Statistics, The University of Hong Kong, Hong Kong.

Ling, S. and Li, W.K. (1997c) On fractionally integrated autoregressive moving-average time series models with conditional heteroskedasticity. *Journal of the American Statistical Association* **92**, 1184-1194.

Ling, S. and Li, W.K. (1998) Limiting distributions of maximum likelihood estimators for unstable ARMA models with GARCH errors. *Annals of Statistics* **26**, 84-125.

Ling, S., Li, W.K. and McAleer, M. (1999) Efficient estimation and testing for unit root processes with GARCH(1, 1) errors: Theory and Monte Carlo evidence. To appear in *Econometric Reviews*.

Ling, S. and McAleer, M. (1999a) On adaptive estimation in nonstationary ARMA models with GARCH errors. Tentatively accepted by *Annals of Statistics*.

Ling, S. and McAleer, M. (1999b) Asymptotic theory for a vector ARMA-GARCH model. To appear in *Econometric Theory*.

Ling, S. and McAleer, M. (1999c) Necessary and sufficient moment conditions for the GARCH(r, s) and asymmetric power GARCH(r, s) models. To appear in *Econometric Theory.*

Ling, S. and McAleer, M. (1999d) Stationarity and the existence of moments of a family of GARCH processes. To appear in *Journal of Econometrics.*

Linton, O. (1993) Adaptive estimation in ARCH models. *Econometric Theory* **9**, 535-569.

Liu, J., Li, W.K. and Li, C.W. (1997) On a threshold autoregression with conditional heteroscedastic variances. *Journal of Statistical Planning and Inference* **62**, 279-300.

Lumsdaine, R.L. (1996) Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica* **64**, 575-596.

Lundbergh S. and Teräsvirta T. (1998a) Modelling economic high-frequency time series with STAR-STGARCH models. Working paper **291**, Stockholm School of Economics, Stockholm, Sweden.

Lundbergh S. and Teräsvirta T. (1998b) Evaluating GARCH models. Working paper **292**, Stockholm School of Economics, Stockholm, Sweden.

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.

Mikosch, T. and Stărică C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Annals of Statistics* **28**, 1427-1451.

Milhøj, A.(1985) The moment structure of ARCH processes. *Scandinavian Journal of Statistics* **12**, 281-292.

Nelson, D. B. (1990) Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory* **6**, 318-334.

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

Pantula, S.G. (1989) Estimation of autoregressive models with ARCH errors. *Sankhya B* **50**, 119-38.

Pesaran, M.H. and Potter, S.M. (1997) A floor and ceiling model of US output. *Journal of Economic Dynamics and Control* **21**, 661-695.

Peters, T.A. and Veloce, W. (1988) Robustness of unit root tests in ARMA models with generalized ARCH errors. Manuscript, Brock University, St. Catherine, Canada.

Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica*, **55**, 277-301.

Phillips, P.C.B. and Durlauf, S.N. (1986) Multiple time series regression with integrated processes. *Review of Economic Studies* **LIII**, 473-495.

Potscher, B.M. (1983) Order estimation in ARMA-models by Lagrangian multiplier tests. *Annals of Statistics* **11**, 872-885.

Potscher, B.M. (1989) Model selection under nonstationarity: Autoregression models and stochastic linear regression models. *Annals of Statistics* **17**, 1257-1274.

Rabemanjara, R. and Zakoïan, J.-M. (1993) Threshold ARCH model and asymmetrics in volatility. *Journal of Applied Econometrics* **8**, 31-49.

Rao, C.R. (1978) *Linear Models and Statistical Inference.* Wiley, New York.

Resnick, S.I. (1997) Heavy tail modeling and teletraffic data. *Annals of Statistics* **25**, 1805-1869.

Robinson, P.M. (1988) Semiparametric econometrics: A survey. *Journal of Applied Econometrics* **3**, 35-51.

Robinson, P.M. (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* **47**, 67-84.

Ross, P.E. (1996) *Modelling Stock Market Volatility: Bridging the Gap to Continuous Time.* Academic Press, New York.

Seo, B. (1999) Distribution theory for unit root tests with conditional heteroskedasticity. *Journal of Econometrics* **91**, 113-144.

Shephard, N. (1996) Statistical aspects of ARCH and stochastic volatility, in *Time Series Models: In Econometrics, Finance and Other Fields*, D.R. Cox, D.V. Hinkley and O. Barndorff-Nielsen (eds.), Chapman and Hall, London, pp.1-67.

Stoker, T.M. (1991) *Lectures on Semiparametric Econometrics.* CORE, Louvain-la-Neuve, Belgium.

Taylor, S.J. (1986) *Modelling Financial Time Series.* Wiley, New York.

Tong, H. (1978) On a threshold model. In *Pattern Recognition and Signal Processing.* (editor C.H. Chan) pp.575-586. Sijthoff and Noordhoff, Amsterdam.

Tong, H. (1990) *Non-linear Time Series : A Dynamical System Approach.* Oxford University Press, Oxford.

Tong, H. and Lim, K.S. (1980) Threshold autoregresssion, limit cycles and cyclical data. *Journal of the Royal Statistical Association* **B 42**, 245-292.

Tsay, R.S. (1984) Order section in nonstationary autoregression models. *Annals of Statistics* **12**, 1425-1433.

Tsay, R.S. (1987) Conditional heteroscedastic time series models. *Journal of the American Statistical Association* **82**, 590-604.

Tsay, R.S. and Tiao, G.C. (1990) Asymptotic properties of multivariate nonstationary processes with applications to autoregressions. *Annals of Statistics* **18**, 220-250.

Tse, Y.K. (1999) Residual-based diagnostics for conditional heteroscedasticity models. Manuscript, Department of Economics, National University of Singapore.

Tse, Y.K. and Zuo, X.L. (1997) Testing for conditional heteroscedasticity: Some Monte Carlo results. *Journal of Statistical Computation and Simulation* **58**, 237-253.

Tweedie, R.L. (1988) Invariant measure for Markov chains with no irreducibility assumptions. *Journal of Applied Probability* **25A**, 275-85.

Wei, C.Z. (1992) On predictive least squares principles. *Annals of Statistics* **20**, 1-42.

Weiss, A.A. (1986) Asymptotic theory for ARCH models: Estimation and testing. *Econometric Theory* **2**, 107-131.

Wong, C.S. and Li, W.K. (1996) On the use of the predictive least squares criterion in time series with changing conditional variance. Technical Report 117, The Department of Statistics and Actuarial Science, The University of Hong Kong.

Wong, C.S. and Li, W.K. (1997) Testing for threshold autoregression with conditional heteroscedasticity. *Biometrika* **84**, 407-418.

Wong, C.S. and Li, W.K. (1999) Testing for double threshold autoregressive conditional heteroscedasticity model. To appear in *Statistica Sinica*.

Wong, C.S. and Li, W.K. (2000) On a mixture autoregressive model. *Journal of the Royal Statistical Association B* **62**, 95-115.

Wong, C.S. and Li, W.K. (2001a) On a mixture autoregressive conditional heteroscedastic model. To appear in *Journal of the American Statistical Association*.

Wong, C.S. and Li, W.K. (2001b) On a logistic mixture autoregressive model. To appear in *Biometrika*.

Wong, H. and Li, W.K. (1997) On a multivariate conditional heteroscedasticity model. *Biometrika* **4**, 111-123.

Yap, S.F. and Reinsel, G.C. (1995) Estimation and testing for unit roots in a partially nonstationary vector autoregressive moving average model. *Journal of the American Statistical Association* **90**, 253-267.