MLE for Change-Point in ARMA-GARCH Models with a Changing Drift

Shiqing Ling

Department of Mathematics, Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Abstract

This paper investigates the maximum likelihood estimator (MLE) of structure -changed ARMA-GARCH models. The convergent rates of the estimated change-point and other estimated parameters are obtained. After suitably normalized, it is shown that the estimated change-point has the same asymptotic distribution as that in Picard(1985) and Yao (1987). Other estimated parameters are shown to be asymptotically normal. As special cases, we obtain the asymptotic distributions of MLEs for structure-changed GARCH models, structure-changed ARMA models with structure-unchanged GARCH errors, and structure-changed ARMA models with i.i.d. errors, respectively.
Key words : ARMA-GARCH model, Brownian motion, change-point, rate of convergence, limiting distribution, maximum likelihood estimator.

1 Introduction

Consider the following autoregressive moving-average (ARMA) model with the generalized autoregressive conditional heteroscedasticity (GA RCH) errors:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t, \tag{1.1}$$

$$\varepsilon_t = \eta_t \sqrt{h_t}, \tag{1.2}$$

$$h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \tag{1.3}$$

where p, q, r and s are known positive integers and η_t are independent and identically distributed (i.i.d.). Equations (1.1)-(1.3) is called the ARMA-GARCH model and denoted by $M(\lambda)$, where $\lambda = (m', \delta')'$ with $m = (\phi_1, \cdots, \phi_p, \psi_1, \cdots, \psi_q)'$ and $\delta = (\alpha_0, \alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s)'$. Denote $Y_i^k = (y_i, \cdots, y_k)'$. $Y_i^k \in M(\lambda_0)$ means that $y_i, \cdots, y_k$ are generated by model (1.1)-(1.3) with the true parameter $\lambda = \lambda_0$. We say that Y_1^n follows a structure-changed ARMA-GA RCH model if there exist $k_0 \in [1, n-1]$, $\lambda_0 \in \Theta$ and $\lambda_{01} \in \Theta$ with $\lambda_0 \neq \lambda_{01}$ so that

$$Y_1^{k_0} \in M(\lambda_0) \text{ and } Y_{k_0+1}^{n} \in M(\lambda_{01}). \tag{1.4}$$

This structure-changed ARMA- GARCH model is denoted by $M(k_0, \lambda_0, \lambda_{01})$. k_0 is called the change-point of this structure-changed model. $Y_1^n \in M(k_0, \lambda_0, \lambda_{01})$ means that (1.4) holds. The focus of this paper is to investigate the maximum likelihood estimator (MLE) of model $M(k_0, \lambda_0, \lambda_{01})$.

Structural change has been recognized to be an important issue in econometrics, engineering, and statistics for a long time. The literature in this area is extensive. The earliest references can go back to Quandt (1960) and Chow (1960). Many approaches have been developed to detect whether or not structural change exists in a statistical model. Examples are the weighted likelihood ratio test in Picard (1985), and Andrew and Ploberger (1994), Wald and Lagrange multiplier tests in Hansen (1993), Andrews (1993), and Bai and Perron (1998), the exact likelihood ratio test in Davis, Huang and Yao (1993), the empirical methods in Bai (1996), and the sequential test in Lai (1995). A general theory for exact testing change-points in time series models was established by Ling (2002a). Empirically, we want to know not only that structural change exists, but also the location of change-point.

Hinkley (1970) and Hinkley and Hinkley (19 70) investigated the MLE of change-points in a sequence of i.i.d. Gaussian random variables and the binomial model, respectively. Their changed parameters are fixed and the limiting distributions of the estimated change-points seem not to be useful in practice. Picard (1985) allowed the difference between parameters before and after the change-point in AR models to tend to zero but not too fast when the sample size tends to infinity, and obtained a nice limiting distribution for the estimated change-point. This distribution can be used to construct the confidence intervals of the change-point, and hence it is very useful in applications, as these confidence intervals indicate the degree of estimation accuracy. Yao (1987) used a similar idea for independent data and obtained the same limiting distribution as Picard's. Picard's method has been developed for the regression models by Bai (1994, 1995, 1997). In particular, Bai, Lumsdaine and Stock (1998) used Picard's method for the structure-changed multivariate AR model and cointegrating time series model, and derived the asymptotic distributions of the change-points in these models. Chong (2001) developed a comprehensive theorem for the structure-changed AR(1) model. A general theory for estimating change-points in time series models with a fixed drift was established by Ling (2002b).

In this paper, we use Picard's method to model $M(k_0, \lambda_0, \lambda_{01})$. The convergent rates of the estimated change-point and other estimated parameters are obtained. After suitably normalized, it is shown that the estimated change-point has the same asymptotic distribution as that in Picard(1985) and Yao (1987). Other estimated parameters are shown to be asymptotically normal. As special cases of model $M(k_0, \lambda_0, \lambda_{01})$, this paper obtains the asymptotic distributions of MLEs for structure-changed GARCH models, structure-changed ARMA models with structure-unchanged GARCH errors, and structure- changed ARMA models with i.i.d. errors, respectively.

2 Main Results

As common practice, we parameterize the change -point as $k_0 = [n\tau_0]$, where $\tau_0 \in (0,1)$ and $[x]$ represents the integer part of x. We assume that $\lambda_{01} = \lambda_{0n}$ is changed over n with $d_n = \lambda_{0n} - \lambda_0 \to 0$ as $n \to \infty$. It is reasonable to allow the changed parameters to have such a small shift. Some arguments on this can be found in Picard (1985) and Bai et al. (1998). Let $Y_1^n \in M([n\tau_0], \lambda_0, \lambda_{0n})$ and the corresponding unknown parameter model be $M([n\tau], \lambda, \lambda_1)$, wh -ere $(\tau, \lambda, \lambda_1) \in (0,1) \times \Theta^2$ and Θ be a compact subset of R^l with $l = p + q + r + s + 1$. Suppose that $(\tau_0, \lambda_0, \lambda_{0n})$ is an interior point in $(0,1) \times \Theta^2$ and, for each $\lambda \in \Theta$, it follows that

Assumption 1. *All the roots of $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\psi(z) = 1 + \psi_1 z + \cdots + \psi_q z^q$ are outside the unit circle and have no common root, $\phi_p \neq 0$ and $\psi_q \neq 0$;*

Assumption 2. $0 < \underline{\alpha}_0 \leq \alpha_0 \leq \bar{\alpha}_0$, $\alpha_i \geq 0$, $i = 1, \cdots, r-1$, $\alpha_r \neq 0$, $\beta_j \geq \underline{\beta} > 0$, $j = 1, \cdots, s$, $\sum_{i=1}^r \alpha_i + \sum_{j=1}^s \beta_j < 1$ *and* $\sum_{i=1}^r \alpha_i z^i$ *and* $1 - \sum_{i=1}^s \beta_i z^i$ *have no common root.*]

Assumption 3. $Ey_t^4 < \infty$.

The necessary and sufficient condition for $Ey_t^4 < \infty$, see Ling (1999) and Ling and McAleer (2002b). Conditional on $Y_0 = (y_0, y_{-1}, \cdots)$, the log-likelihood function (ignoring a constant) can be written as

$$L_n(\tau, \lambda, \lambda_1) = \Big[\sum_{t=1}^{[n\tau]} l_t(\lambda) + \sum_{t=[n\tau]+1}^{n} l_t(\lambda_1)\Big] - \Big[\sum_{t=1}^{[n\tau_0]} l_t(\lambda_0) + \sum_{t=[n\tau_0]+1}^{n} l_t(\lambda_{0n})\Big], \tag{2.1}$$

where $l_t(\lambda) = -\varepsilon_t^2(\lambda)/2h_t(\lambda) - 2^{-1}\log h_t(\lambda)$, and $\varepsilon_t(\lambda)$ and $h_t(\lambda)$ are defined as ε_t and h_t, respectively, but now they are the functions of Y_1^n, Y_0 and λ. Since we do not assume that η_t is normal, (2.1) is called the quasi-likelihood function, and its maximizer on the parameter space $(0,1) \times \Theta^2$, denoted by $(\hat{\tau}_n, \hat{\lambda}_n, \hat{\lambda}_{1n})$, is called the quasi-maximum likelihood estimator (QMLE) of $(\tau_0, \lambda_0, \lambda_{0n})$. In practice, the initial value Y_0 is not available and can be replaced by any constant. This does not affect the asymptotic behavior of the QMLE, see Ling and Li (1997).

When $y_t \in M(\lambda)$, y_t is a fixed function of λ and $\{\eta_t, \eta_{t-1}, \cdots\}$, and hence when $Y_1^\infty \in M(\lambda_{0n})$, it strictly constitutes a triangular array of the type $\{y_{tn} : t = 1, 2, \cdots; n = 1, 2, \cdots\}$. In order not to overburden notation, we simply refer to the time series generated by $M(\lambda_{0n})$ as y_t. Our first result gives the rates of convergence of the QMLE and it plays an important role in the proof of Theorem 2.2.

Theorem 2.1. *Suppose that $Y_1^n \in M([n\tau_0], \lambda_0, \lambda_{0n})$, $E|\eta_t|^{4+\iota} < \infty$ for some $\iota > 0$, and Assumptions 1-3 hold. If $d_n \to 0$ and $\sqrt{n}\|d_n\|/\log n$*

$\to \infty$, *then*

$$\hat{\tau}_n - \tau_0 = O_p\Big(\frac{1}{nd_n' d_n}\Big),$$
$$\hat{\lambda}_n - \lambda_0 = O_p\Big(\frac{1}{\sqrt{n}}\Big),$$
$$\hat{\lambda}_{1n} - \lambda_{0n} = O_p\Big(\frac{1}{\sqrt{n}}\Big).$$

Now, we state our second main result which shows the limiting distribution of the QMLE. In the following, F is the distribution with density f on R:

$$f(x) = \frac{3}{2}e^{|x|}\Phi(\frac{3}{2}\sqrt{|x|}) - \frac{1}{2}\Phi(\frac{\sqrt{|x|}}{2}),$$
$$\Phi(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\exp(-\frac{u^2}{2})du.$$

This distribution was first found by Picard (1985) and Yao (1987). The latter also tabulated its numerical approximation.

Theorem 2.2. *If the conditions in Theorem 2.1 are satisfied, then $\hat{\tau}_n$, $\hat{\lambda}_n$ and $\hat{\lambda}_{1n}$ are asymptotically independent, and*

$$C_n^{-1}n(d_n'\Omega d_n)(\hat{\tau}_n - \tau_0) \longrightarrow_{\mathcal{L}} F,$$
$$\sqrt{n}(\hat{\lambda}_n - \lambda_0) \longrightarrow_{\mathcal{L}} N(0, \frac{1}{\tau_0}\Omega^{-1}\Sigma\Omega^{-1}),$$
$$\sqrt{n}(\hat{\lambda}_{1n} - \lambda_{0n}) \longrightarrow_{\mathcal{L}} N(0, \frac{\Omega^{-1}\Sigma\Omega^{-1}}{1-\tau_0}),$$

where $C_n = (d_n'\Sigma d_n)(d_n'\Omega d_n)^{-1}$, $\Omega = E[\partial^2 l_t(\lambda_0)/\partial\lambda\partial\lambda']$, and $\Sigma = E[(\partial l_t(\lambda_0)/\partial\lambda)\ (\partial l_t(\lambda_0)\ /\partial\lambda')]$.

As mentioned in Section 1, model (1.1)-(1.3) implies some important special cases. We first consider the GARCH model:

$$\varepsilon_t = \eta_t\sqrt{h_t}, \tag{2.2}$$
$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^s \beta_i h_{t-i}. \tag{2.3}$$

Denote model (2.2)-(2.3) by $M(\delta)$ and let $\tilde{\varepsilon}_i^k = (\varepsilon_i, \cdots, \varepsilon_k)'$. We say that $\tilde{\varepsilon}_1^n$ follows a structure-changed GARCH model if there exists $k_0 \in [1, n-1]$ so that $\tilde{\varepsilon}_1^{k_0} \in M(\delta_0)$ and $\tilde{\varepsilon}_{k_0+1}^n \in M(\delta_{0n})$ with $\delta_0 \neq \delta_{0n}$. This structure-changed model is denoted by $M(k_0, \delta_0, \delta_{0n})$. Let $(\hat{\tau}_n, \hat{\delta}_n, \hat{\delta}_{1n})$ be the QMLE of $(\tau_0, \delta_0, \delta_{0n})$. From Theorem 2.2, we immediately obtain the following result.

Corollary 2.1. *Suppose that $\tilde{\varepsilon}_1^n \in M([n\tau_0], \delta_0, \delta_{0n})$, $E|\eta_t|^{4+\iota} < \infty$ for some $\iota > 0$, and Assumption 2 holds. If $d_n = \delta_{0n} - \delta_0 \to 0$ and $\sqrt{n}\|d_n\|/\log n \to \infty$, then $\hat{\tau}_n$, $\hat{\delta}_n$ and $\hat{\delta}_{1n}$ are asymptotically independent, and*

$$n\kappa^{-1}(d_n'Qd_n)(\hat{\tau}_n - \tau_0) \longrightarrow_{\mathcal{L}} F,$$
$$\sqrt{n}(\hat{\delta}_n - \delta_0) \longrightarrow_{\mathcal{L}} N(0, \frac{\kappa}{4\tau_0}R^{-1}),$$
$$\sqrt{n}(\hat{\delta}_{1n} - \delta_{0n}) \longrightarrow_{\mathcal{L}} N(0, \frac{\kappa}{4(1-\tau_0)}R^{-1}).$$

Next, we consider the structure-changed AR-MA model with structure-unchanged GARCH errors. We say that Y_1^n follows such a model if there exists $k_0 \in [1, n-1]$ so that $Y_1^{k_0} \in M(\lambda_0) = M(m_0, \delta_0)$ and $Y_{k_0+1}^n \in M(\lambda_{0n}) = M(m_{0n}, \delta_0)$ with $m_0 \neq m_{0n}$. This structure-changed model is denoted by $M(k_0, m_0, m_{0n}, \delta_0)$. Let $(\hat{\tau}_n, \hat{m}_n, \hat{m}_{1n}, \hat{\delta}_n)$ be the QMLE of $(\tau_0, m_0, m_{0n}, \delta_0)$. From Theorem 2.2, we obtain the following corollary.

Corollary 2.2. *Suppose that $Y_1^n \in M([n\tau_0], m_0, m_{0n}, \delta_0)$, Assumptions 1-3 hold, η_t has a symmetric density and $E|\eta_t|^{4+\iota} < \infty$ for some $\iota > 0$. If $d_n = m_{0n} - m_0 \to 0$ and $\sqrt{n}\|d_n\|/\log n \to \infty$, then $\hat{\tau}_n$, $\hat{m}_n$, $\hat{m}_{1n}$ and $\hat{\delta}_n$ are asymptotically independent, and*

$$C_n^{-1}n(d_n'\Omega_m d_n)(\hat{\tau}_n - \tau_0) \longrightarrow_{\mathcal{L}} F,$$
$$\sqrt{n}(\hat{m}_{1n} - m_{0n}) \longrightarrow_{\mathcal{L}} N(0, \frac{\Omega_m^{-1}\Sigma_m\Omega_m^{-1}}{1-\tau_0}),$$
$$\sqrt{n}(\hat{m}_n - m_0) \longrightarrow_{\mathcal{L}} N(0, \frac{1}{\tau_0}\Omega_m^{-1}\Sigma_m\Omega_m^{-1}),$$
$$\sqrt{n}(\hat{\delta}_n - \delta_0) \longrightarrow_{\mathcal{L}} N(0, \frac{\kappa}{4}R^{-1}),$$

where $C_n = (d_n'\Sigma_m d_n)(d_n'\Omega_m d_n)^{-1}$, $\Omega_m = P + 2Q$, and $\Sigma_m = P + \kappa Q$.

Finally, we consider the ARMA model:

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \psi_i\varepsilon_{t-i} + \varepsilon_t, \tag{2.4}$$

where ε_t are i.i.d. with mean zero and variance σ^2. We denote model (2.4) by $M(m)$. We say that Y_1^n follows a structure-changed ARMA model with i.i.d. errors if there exists $k_0 \in [1, n-1]$ so that $Y_1^{k_0} \in M(m_0)$ and $Y_{k_0+1}^n \in M(m_{0n})$ with $m_0 \neq m_{0n}$. This structure-changed model is denoted by $M(k_0, m_0, m_{0n})$. In this case, maximizing the log-likelihood function (2.1) is equivalent to minimizing:

$$L_n(\tau, m, m_1) = \sum_{t=1}^{[n\tau]} \varepsilon_t^2(m) + \sum_{t=[n\tau]+1}^n \varepsilon_t^2(m_1)$$
$$- \sum_{t=1}^{[n\tau_0]} \varepsilon_t^2(m_0) - \sum_{t=[n\tau_0]+1}^n \varepsilon_t^2(m_{0n}). \tag{2.5}$$

Usually, we call the minimizer of $L_n(\tau, m, m_1)$ the conditional least squares estimator (CLSE), denoted by $(\hat{\tau}_{Ln},\ \hat{m}_{Ln}, \hat{m}_{L1n})$. From Theorem 2.2, we obtain the following result.

Corollary 2.3. *Suppose that $Y_1^n \in M([n\tau_0], m_0, m_{0n})$, Assumption 1 holds and $E|\varepsilon_t|^{2+\iota} < \infty$ for some $\iota > 0$. If $d_n = m_{0n} - m_0 \to 0$ and $\sqrt{n}\|d_n\|/\log n \to \infty$, then $\hat{\tau}_{Ln}$, $\hat{m}_{Ln}$ and $\hat{m}_{L1n}$ are asymptotically independent, and*

$$n(d_n' V d_n)\sigma^{-2}(\hat{\tau}_{Ln} - \tau_0) \longrightarrow_{\mathcal{L}} F,$$

$$\sqrt{n}(\hat{m}_{Ln} - m_0) \longrightarrow_{\mathcal{L}} N(0, \frac{\sigma^2}{\tau_0}V^{-1}),$$

$$\sqrt{n}(\hat{m}_{L1n} - m_{0n}) \longrightarrow_{\mathcal{L}} N(0, \frac{\sigma^2}{1-\tau_0}V^{-1}),$$

where $V = E[(\partial\varepsilon_t(\lambda_0)/\partial m)(\partial\varepsilon_t(\lambda_0)/\partial m')]$.

Remark 2.1. When $Y_1^n \in M([n\tau_0], m_0, m_{0n}, \delta_0)$, the objective function (2.5) can be used to estimate (τ_0, m_0, m_{0n}). Using a similar approach as for Theorems 2.1-2.2, we can show that

$$C_{Ln}^{-1} n(d_n' V d_n)(\hat{\tau}_{Ln} - \tau_0) \longrightarrow_{\mathcal{L}} F, \tag{2.6}$$

where $C_{Ln} = (d_n'\tilde{V}d_n)(d_n'Vd_n)^{-1}$ and $\tilde{V} = E[(\partial\varepsilon_t(\lambda_0)/\partial m)(\partial\varepsilon_t(\lambda_0)/\partial m')\varepsilon_t^2]$. It is interesting to compare the two estimators $\hat{\tau}_n$ in Corollary 2.2 and $\hat{\tau}_{Ln}$ in (2.6). The optimality of estimated change-points has not been established in the literature. We define efficiency as follows:

> *Assume that $\hat{\tau}_{1n}$ and $\hat{\tau}_{2n}$ are two estimators of τ_0 such that $D_{1n}(\hat{\tau}_{1n} - \tau_0) \to_{\mathcal{L}} F$ and $D_{2n}(\hat{\tau}_{2n} - \tau_0) \to_{\mathcal{L}} F$. We say that $\hat{\tau}_{1n}$ is more efficient than $\hat{\tau}_{2n}$ if $D_{1n} \geq D_{2n}$ for all n.*

The reason for this definition is that $\hat{\tau}_{1n}$ can always provide sharper confidence intervals than $\hat{\tau}_{2n}$ at any significant level. Under this definition, it is easy to see that the QMLE $\hat{\tau}_n$ in Corollary 2.2 is more efficient than the CLSE $\hat{\tau}_{Ln}$ in (2.6) if η_t is normal since $(d_n'Vd_n)^2 \leq (d_n'Pd_n)(d_n'\tilde{V}d_n)$ by Chebyshev's inequality. We can show that, even when η_t is not normal, the QMLE is also more efficient than its CLSE.

3 Simulation Studies

We first examine the performance of our asymptotic results in the finite samples via some Monte Carlo experiments. The following three models are used:

$$\begin{aligned}
\text{Model 1}: \quad & y_t = \phi_0 y_{t-1} \\
& \qquad \cdot I_{\{t\leq k_0\}} + \phi_{0n}y_{t-1}I_{\{t>k_0\}} + \varepsilon_t, \\
& \varepsilon_t = \eta_t\sqrt{h_t}, \\
& h_t = (\alpha_{00} + \alpha_0\varepsilon_{t-1}^2 + \beta_0 h_{t-1}) \\
& \qquad \cdot I_{\{t\leq k_0\}} + (\alpha_{00n} + \alpha_{0n}\varepsilon_{t-1}^2 \\
& \qquad + \beta_{0n}h_{t-1})I_{\{t>k_0\}} \\
\text{Model 2}: \quad & \varepsilon_t = \eta_t\sqrt{h_t} \text{ and} \\
& h_t = (\alpha_{00} + \alpha_0\varepsilon_{t-1}^2 + \beta_0 h_{t-1}) \\
& \qquad \cdot I_{\{t\leq k_0\}} + (\alpha_{00n} + \alpha_{0n}\varepsilon_{t-1}^2 \\
& \qquad + \beta_{0n}h_{t-1})I_{\{t>k_0\}} \\
\text{Model 3}: \quad & y_t = \phi_0 y_{t-1} \\
& \qquad \cdot I_{\{t\leq k_0\}} + \phi_{0n}y_{t-1}I_{\{t>k_0\}} + \varepsilon_t, \\
& \varepsilon_t = \eta_t\sqrt{h_t}, \\
& h_t = \alpha_{00} + \alpha_0\varepsilon_{t-1}^2 + \beta_0 h_{t-1},
\end{aligned}$$

where $\eta_t \sim$ i.i.d.$N(0,1)$. The true observations are generated through these models with parameters: $\phi_0 = 0.5$, $\alpha_{00} = 1.0$, $\alpha_0 = 0.2$, $\beta_0 = 0.7$, $\phi_{0n} = -0.6$, $\alpha_{00n} = 0.5$, $\alpha_{0n} = 0.57$ and $\beta_{0n} = 0.02$. We use 4000 replications in all the experiments. These experiments are carried out by Fortran 77 and the optimization algorithm from Fortran subroutine DBCOAH in the IMSL library is used.

In Tables 1 and 2, we summarize the empirical means and standard deviations (SD) of the MLEs of λ_0 and λ_{0n}, $\hat{\lambda}_n = (\hat{\phi}_n, \hat{\alpha}_{0n}, \hat{\alpha}_n, \hat{\beta}_n)'$ and $\hat{\lambda}_{1n} = (\hat{\phi}_{1n}, \hat{\alpha}_{01n}, \hat{\alpha}_{1n}, \hat{\beta}_{1n})'$. From the two tables, we see that the SDs of $\hat{\lambda}_n$ and $\hat{\lambda}_{1n}$ are decreased and increased, respectively, as τ_0 is increased from 0.496 to 0.504. This is consistent with our theoretical results in Section 2. As the sample size n is increased from 250 to 400, both the corresponding biases and SDs become smaller. This is the same as the usual results in the structure-unchanged AR-GARCH models. For Models 2 and 3, the Monte Carlo results are similar and hence are not reported here.

We use two methods to estimate Model 3, that is, the MLE and the CLSE as in Corollary 2.3 and Remark 2.4. Model 3 is denoted by Model 3a and Model 3b as it is estimated by the MLE and the CLSE, respectively. In Table 3, we report the 90% range, estimated asymptotic confidence interval (EACI), and asymptotic confidence interval (ACI) of the change - point $k_0 = [n\tau_0]$ with $\tau_0 = 0.500$ and the sample sizes $n = 250, 400$. The empirical mean is the average of $\hat{k}_n$ from the 4000 replications. The 90% range are respectively the 50%-quantile and the range between the 5% and 95% quantiles of the distribution of $\hat{k}_n$. The EACI and ACI are computed, respectively, by the following formulas:

$$\left[\hat{k}_n - [\Delta F_{\omega/2}] - 1, \hat{k}_n - [\Delta F_{\omega/2}] + 1\right]$$
$$\left[k_0 - [\Delta F_{\omega/2}] - 1, k_0 - [\Delta F_{\omega/2}] + 1\right],$$

where $F_{\omega/2}$ is the ωth quantile of the distribution F and $\Delta = (d_n'\Omega d_n)^{-1}$, $(d_n' Q d_n)^{-1}$, $(d_n'\Omega_m d_n)^{-1}$ and $(d_n'\tilde{V}d_n)(d_n' V d_n)^{-2}$ for Model 1, Mo -del 2, Model 3a and Model 3b, respectively. Using the density function $f(x)$ in Section 2, we obtain $F_{0.05} = 7.792$. Ω is estimated by $n^{-1}\sum_{t=1}^{n}\partial^2 l_t(\hat{\lambda}_n)/\partial\lambda\partial\lambda'$. Q, $\tilde{V}$ and V are similarly estimated. Under Assumptions 1-3, these estimators are consistent in probability (see Ling and Li, 1997). From Tables 3, we see that the ACI is exactly the same as EACI. The 90% range is slightly wider than EACI and ACI in all cases. As n is increased from 250 to 400, the EACIs and ACIs of $\hat{k}_n - k_0$ have not been improved. This is because the rate of convergence of $\hat{k}_n - k_0$ is $O_p(1/d_n' d_n)$, which only depends on $d_n' d_n$, while d_n is fixed for $n = 250$ and 400 in our experiments. This finding is similar to those in Bai (1995) for the structure-changed regression model and in Bai et al.(1998) for the structure-changed multivariate AR models and cointegrating time series models. Comparing Model 1 with Model 3a, we can see that both the 90% range and the ACI for Model 1 are, respectively, tighter than those for Model 3a. This means that we can estimate k_0 more precisely when changed coefficients also exist in the GARCH part. Comparing Model 3a with Model 3b, we find that both the 90% range and the ACI for Model 3a are, respectively, tighter than those for Model 3b. This is consistent with our discussion in Remark 2.1.

REFERENCES

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**, 821-856.

Andrews, D.W.K. and Ploberger, W. (1994) Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* **62**, 1383-1414.

Bai, J. (1994) Least squares estimation of a shift in linear processes. *J. Time Ser. Anal.* **15**, 453-472.

Bai, J. (1995) Least absolute deviation estimation of a shift. *Econometric* **11**, 403-436.

Bai, J. (1996) Testing for parameter constancy in linear regressions: an empirical distribution function approach. *Econometrica* **64**, 597-622.

Bai, J. (1997) Estimating multiple breaks one at a time. *Econometric Theory* **13**, 315-352.

Bai, J. and Perron, P. (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* **66**, 47-78.

Bai, J., Lumsdaine, R.L. and Stock, J.H. (1998) Testing for and dating common bre -aks in multivariate time series. *Rev. Eco nom. Stud.* **65**, 395-432.

Birnbaum, Z.W. and Marshall, A.W. (1961) Some multivariate Chebyshew inequalities with extensions to continuous parameter processes. *Ann. Math. Statist.* **32**, 687-703.

Chong, T.T.L. (2001) Structural change in AR(1) models. *Econometric Theory* **17**, 87-155.

Chow, G.C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **28** 591-605.

Davis, R.A. and Huang, D.W. and Yao, Y.C. (1995) Testing for a change in the parameter values and order of an autoregressive model. *Ann. Statist. 23*, 282-304.

Hansen, B.E. (1993) Tests for parameters instability in regression with $I(1)$ processes. *Journal of Business and economic Statistics* **10**, 321-335.

Hinkley, D.V. (1970) Inference about the cha -nge-point in a sequence of random variables. *Biometrika* **57**, 1-17.

Hinkley, D.V. and Hinkley, E.A. (1970) Inference about the change-point in a sequence of binomial variables. *Biometrika* **57**, 477-488.

Lai, T.L. (1995) Sequential changepoint detection in quality control and dynamical systems. With discussion and a reply by the author. *J. Roy. Statist. Soc. Ser.* **B 57**, 613-658.

Ling, S. (1999) On the probabilistic properties of a double threshold ARMA conditional heteroskedasticity model. *Journal of Applied Probability* **36**, 688-705.

Ling, S. (2002a) Exact testing for change-points in time series models. Working paper.

Ling, S. (2002b) Asymptotic theory for estimating change-points in time series models. Working paper, HKUST.

Ling, S. (2002c) Estimation for change-point in ARMA-GARCH models with a fixed drift. Working paper, HKUST.

Ling, S. and Li, W.K. (1997) On fractionally integrated autoregressive moving-average time series models with conditional heteroskedasticity. *J. Amer. Statist. Assoc.* **92**, 1184-1194.

Ling, S. and McAleer, M. (2002b) Necessary and sufficient moment conditions for the GARCH and Asymmetric Power GAR CH models. *Econometric Theory* **18**, 722-729.

Picard, D. (1985) Testing and estimating change-points in time series. *Adv. in Appl. Probab.* **17**, 841-867.

Quandt, R. E. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55**, 324-330.

Yao, Y.-C. (1987) Approximating the distribution of the maximum likelihood estimate of the change-point in a sequence of independent random variables. *Ann. Statist.* **15**, 1321-1328.

TABLE 1

Empirical Mean and Standard Deviation of MLE of λ_0 and λ_{0n}
$n = 250$ and 4000 Replications

		λ_0=(0.5, 1.0, 0.2, 0.7)			
τ_0		$\hat{\phi}_n$	$\hat{\alpha}_{0n}$	$\hat{\alpha}_n$	$\hat{\beta}_n$
.496	Mean	.4954	.9566	.1707	.7251
	SD	.0847	.1748	.1105	.0985
.500	Mean	.4952	.9580	.1706	.7251
	SD	.0845	.1692	.1104	.0980
.504	Mean	.4952	.9579	.1712	.7246
	SD	.0841	.1686	.1098	.0975

TABLE 2

Empirical Mean and Standard Deviation of MLE of λ_0 and λ_{0n}
$n = 400$ and 4000 Replications

		λ_{0n}=(-0.6, 0.5, 0.57, 0.02)			
τ_0		$\hat{\phi}_{1n}$	$\hat{\alpha}_{01n}$	$\hat{\alpha}_{1n}$	$\hat{\beta}_{1n}$
.496	Mean	$-.5968$	.4851	.5408	.0495
	SD	.0556	.0994	.1524	.0741
.500	Mean	$-.5931$	.4943	.5288	.0399
	SD	.0562	.0996	.1534	.0755
.504	Mean	$-.5932$	.4944	.5288	.0399
	SD	.0566	.0999	.1543	.0762

TABLE 3

MLE and Confidence Interval of the Change-point k_0
4000 Replications

n=250			
$\tau_0 = 0.500$	90% Range	EACI	ACI
Model 1	[120, 129]	[122, 128]	[122, 128]
Model 2	[115, 131]	[120, 130]	[120, 130]
Model 3a	[116, 135]	[119, 131]	[119, 131]
Model 3b	[115, 136]	[114, 136]	[114, 136]
n=400			
$\tau_0 = 0.500$	90% Range	EACI	ACI
Model 1	[195, 204]	[197, 203]	[197, 203]
Model 2	[191, 207]	[195, 205]	[195, 205]
Model 3a	[191, 210]	[194, 206]	[194, 206]
Model 3b	[190, 211]	[189, 211]	[189, 211]