Working Paper 92-02
January 1992

División de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Madrid)

A GOODNESS-OF-FIT TEST BASED ON RANKS FOR ARMA MODELS

Nélida E. Ferretti, Diana M. Kelmansky and Victor J. Yohai*

Abstract

In this paper we introduce a goodness-of-fit test based on ranks for ARMA models.The classical portmanteau statistic is generalized to a class of estimators based on ranks. The asymptotic distributions of the proposed statistics are derived. Simulation results suggest that the proposed statistics have good robustness properties for an adequate choice of the score functions.

Key words: ARMA models; ranks; goodness-of-fit.

* Nélida E. Ferretti, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Universidad de la Plata, Argentina and CONICET, Argentina; Diana M. Kelmansky, Universidad de Buenos Aires, Argentina and Victor J. Yohai, Universidad de Buenos Aires, Argentina and CONICET, Argentina.

1. INTRODUCTION

Consider the observed series $(Z_1, \ldots, Z_T)$ of a stationary and invertible **ARMA**(p, q) model, i.e.,

$$\boldsymbol{\phi}_0(B)(Z_t - \mu_0) = \boldsymbol{\theta}_0(B)U_t \tag{1.1}$$

where U_t are independent identically distributed (i.i.d.) random variables with distribution F, μ_0 is the mean of Z_t, $\boldsymbol{\phi}_0(B)$ and $\boldsymbol{\theta}_0(B)$ are polynomials given by

$$\boldsymbol{\phi}_0(B) = 1 - \phi_{10}B \cdots - \phi_{p0}B^p$$

and

$$\boldsymbol{\theta}_0(B) = 1 - \theta_{10}B \cdots - \theta_{q0}B^q$$

and B is the backward shift operator defined by $BZ_t = Z_{t-1}$.

Usually, the parameters of a time series model are estimated by the maximum likelihood method assuming $\{U_t\}$ to be Gaussian. After a model of the form (1.1) has been fitted to a series $(Z_1, \ldots, Z_T)$, it is useful to study the adequacy of fit by examining the residuals. One of the most well-known statistics for testing the adequacy of a time series model is the Box-Pierce statistic (Box and Pierce (1970))

$$T \sum_{k=1}^{m} \rho_{1,k}^2,$$

where $\rho_{1,k}$ is the usual lag k residual autocorrelation. This statistic is asymptotically chi-squared distributed with degrees o f freedom $m - p - q$ for large T.

A modified test based on

$$Q_1 = T(T+2) \sum_{k=1}^{m} (T-k)^{-1} \rho_{1,k}^2,$$

was recommended by Ljung and Box (1978). It was shown that it provides a substantially improved chi-square approximation.

In general the usual maximum likelihood or least-squares (LS) procedures are not disturbed by innovation outliers. The LS estimator, however, is sensitive to additive outiers.

Li (1988) proposed to generalize the Q_1 statistic for a class of robust estimators based on residual autocovariances (RA-estimators; Bustos and Yohai (1986)).

In Section 2 a further modification of the Q_1 statistic is introduced and in Section 4 its asymptotic properties are derived. The basic idea is to replace the sample autocorrelation of the residuals by autocovariances based on ranks. Moreover, the robustness properties of the proposed statisics for the AR(1) and MA(1) models investigated in a Monte Carlo study are shown in Section 5.

2. STATISTICS BASED ON RANKS

Denote $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)$ and $\boldsymbol{\lambda} = (\boldsymbol{\phi}, \boldsymbol{\theta})$, and by $\boldsymbol{\phi}_0$, $\boldsymbol{\theta}_0$ and $\boldsymbol{\lambda}_0$ the corresponding true parameters. Also let $s_h(\boldsymbol{\phi})$, $t_h(\boldsymbol{\theta})$ and $g_h(\boldsymbol{\phi}, \boldsymbol{\theta})$ $(0 \le h < \infty)$ be the series expansion coefficients of the operators $\boldsymbol{\phi}^{-1}(B)$, $\boldsymbol{\theta}^{-1}(B)$ and $\boldsymbol{\theta}^{-1}(B)\boldsymbol{\phi}(B)$ respectively. For simplicity and without loss of generality it is assumed that $\mu_0 = 0$.

Let

$$U_t(\boldsymbol{\lambda}) = \sum_{i=0}^{t-1} g_i(\boldsymbol{\phi}, \boldsymbol{\theta}) Z_{t-i}\,, \quad p+1 \le t \le T \tag{2.1}$$

and

$$\mathbf{U}_T(\boldsymbol{\lambda}) = (U_{p+1}(\boldsymbol{\lambda}), \ldots, U_T(\boldsymbol{\lambda})).$$

Bustos and Yohai (1986) have shown that the LS equations for the autoregressive and moving average parameters are asymptotically equivalent to the following system of equations

$$\begin{cases} \displaystyle\sum_{h=0}^{T-j-p-1} s_h(\phi)\, \gamma_{1,h+j}(\mathbf{U}_T(\boldsymbol{\lambda})) = 0\,, & 1 \le j \le p\,, \\ \displaystyle\sum_{h=0}^{T-j-p-1} t_h(\boldsymbol{\theta})\, \gamma_{1,h+j}(\mathbf{U}_T(\boldsymbol{\lambda})) = 0\,, & 1 \le j \le q\,, \end{cases} \tag{2.2}$$

where

$$\gamma_{1,i}(\mathbf{U}_T(\boldsymbol{\lambda})) = \sum_{t=p+1+i}^{T} U_t(\boldsymbol{\lambda}) U_{t-i}(\boldsymbol{\lambda})\,, \quad 0 \le i \le T-p-1\,.$$

Bustos and Yohai (1986) have introduced the class of estimators based on the residual autocovariances (RA–estimators) which are defined by replacing in (2.2) the residual autocovariances $\gamma_{1,i}$'s by robust residual autocovariances of the form

$$\gamma_{2,i}(\mathbf{U}_T(\boldsymbol{\lambda})) = \sum_{t=p+1+i}^{T} \eta\left(\frac{U_t(\boldsymbol{\lambda})}{s}, \frac{U_{t-i}(\boldsymbol{\lambda})}{s}\right)\,, \quad 0 \le i \le T-p-1,$$

where $\eta(u,v)$ is a bounded function and s is an estimate of the innovation scale. Two canonical ways of taking η are: (i) the ***Hampel-Krasker*** type: $\eta(u,v) = \psi(u\,v)$ and (ii) the ***Mallows*** type: $\eta(u,v) = \psi(u)\psi(v)$, where ψ is an odd and bounded function.

Denote by $\widehat{U}_t$ the residuals obtained when λ is replaced by the corresponding RA–estimator. Also define

$$\widehat{\mathbf{U}}_T = (\widehat{U}_{p+1}, \ldots, \widehat{U}_T)$$

and

$$\rho_{2,k}(\widehat{\mathbf{U}}_T) = \gamma_{2,k}(\widehat{\mathbf{U}}_T)/\gamma_{2,0}(\widehat{\mathbf{U}}_T), \qquad 1 \leq k \leq m.$$

Li (1988) proposed the following robustified portmanteau statistic

$$Q_2(\widehat{\mathbf{U}}_T) = T^2 \sum_{k=1}^{m} (T-k)^{-1} \rho_{2,k}^2(\widehat{\mathbf{U}}_T).$$

Let us consider two score generating functions $J_i : [0,1] \rightarrow \Re$, $i = 1, 2$, satisfying $J_i(1-u) = -J_i(u)$. Also let $R_t(\boldsymbol{\lambda})$ be the rank of $U_t(\boldsymbol{\lambda})$ among $U_{p+1}(\boldsymbol{\lambda}), \ldots, U_T(\boldsymbol{\lambda})$. Define the lag i rank autocovariance of the residuals $\gamma_{3,i}$ by

$$\gamma_{3,i}(\mathbf{R}_T(\boldsymbol{\lambda})) = \sum_{t=p+1+i}^{T} J_1\left(\frac{R_t(\boldsymbol{\lambda})}{T-p+1}\right) J_2\left(\frac{R_{t-i}(\boldsymbol{\lambda})}{T-p+1}\right), \quad 0 \leq i \leq T-p-1 \tag{2.3}$$

where $\mathbf{R}_T(\lambda) = (R_{p+1}(\boldsymbol{\lambda}), \ldots, R_T(\boldsymbol{\lambda}))$.

Ferretti, Kelmansky and Yohai (1991) have introduced estimators based on ranks which are defined similarly to the RA–estimators but replacing in (2.2) the $\gamma_{2,i}$'s by the $\gamma_{3,i}$'s given by (2.3).

The following score generating functions J_1 and J_2 will be studied in this paper.

(i) $J_1 = J_2 = \Phi^{-1}$, where Φ is the standard normal distribution function. The RAR–estimators based on these functions give estimators which are optimal when F is normal.

(ii) $J_1(u) = J_2(u) = 2u - 1$.

To define the RAR–estimators more formally we need the following notation:

$$\mathbf{W}_T(\mathbf{R}_T(\lambda), \boldsymbol{\phi}, \boldsymbol{\theta}) = (W_{T,1}(\mathbf{R}_T(\lambda), \boldsymbol{\phi}, \boldsymbol{\theta}), \ldots, W_{T,p+q}(\mathbf{R}_T(\lambda), \boldsymbol{\phi}, \boldsymbol{\theta}))\,,$$

where

$$\begin{aligned} W_{T,j}(\mathbf{R}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) &= (T-j-p)^{-1} \sum_{h=0}^{T-j-p-1} \gamma_{3,h+j}(\mathbf{R}_T(\boldsymbol{\lambda})) s_h(\boldsymbol{\phi}), \quad 1 \leq j \leq p\,, \\ W_{T,p+j}(\mathbf{R}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) &= (T-j-p)^{-1} \sum_{h=0}^{T-j-p-1} \gamma_{3,h+j}(\mathbf{R}_T(\boldsymbol{\lambda})) t_h(\boldsymbol{\theta}), \quad 1 \leq j \leq q\,. \end{aligned} \tag{2.4}$$

Then the RAR–estimators of $\boldsymbol{\lambda}_0$, $\widehat{\boldsymbol{\lambda}}_T = (\widehat{\boldsymbol{\phi}}_T, \widehat{\boldsymbol{\theta}}_T)$, are defined as a sequence satisfying

$$T^{1/2} \mathbf{W}_T(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T), \widehat{\boldsymbol{\phi}}_T, \widehat{\boldsymbol{\theta}}_T) = \mathbf{0}. \tag{2.5}$$

In Ferretti, Kelmansky and Yohai (1991) it was shown that under suitable general assumptions of J the RAR–estimators are asymptotically normal:

$$T^{1/2}(\widehat{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_0) \xrightarrow{\mathcal{D}} N(0, \eta C^{-1}), \quad \text{as } T \to \infty ,$$

where C^{-1} is the covariance matrix of the LS–estimators and the scalar η depends on the J functions and the innovation's distribution F.

In this paper we propose a robust portmanteau goodness–of–fit test based on the satistic

$$Q_3\ (\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)) = \xi^{-1}T(T+2)\sum_{k=1}^{m}(T-k)^{-1}\rho_{3,k}^2(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)),$$

where the constant ξ and the autocorrelation functions $\rho_{3,k}$ are given by

$$\xi = \frac{E(J_1^2(F(U_1)))E(J_2^2(F(U_1)))}{E^2(J_1(F(U_1))J_2(F(U_1)))} \tag{2.6}$$

and

$$\rho_{3,k}(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)) = \gamma_{3,k}(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T))/\gamma_{3,0}(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)), \qquad 1 \le k \le m$$

respectively.

It is shown, under general assumptions, that for m sufficiently large the assymptotic distribution of Q_3 may be approximated by a chi-square with degrees of freedom $m-p-q$.

3. Basic assumptions and notations

Assumption A.

(i) The $U_t's$ have finite moments up to the fourth order, with mean $E(U_t) = 0$ and variance $E(U_t^2) = \sigma^2$.

(ii) F is symmetric and continuous.

(iii) $F(x)$ has a uniformly continuous density $f(x)$ which is a non increasing function of $|x|$ and strictly decreasing for small x.

(iv) f has finite Fisher's information $I(f)$, i.e., f is absolutely continuous on finite intervals and $0 < I(f) = \int_{-\infty}^{\infty}(f'(x)/f(x))^2 f(x)dx < \infty$.

(v) Let $F^{-1}(u) = \inf\{x : F(x) \ge u\}$, $0 < u < 1$ and $\varphi(x) = -f'(x)/f(x)$, $x \in \Re$. Assume that $\varphi(x)$ is a.e. derivable and that its derivative $\varphi'(x)$ is a.e. Lipschitzian and square integrable: $|\varphi'(x) - \varphi'(y)| < K|x-y|$ and $\int_0^1 \varphi'^2(F^{-1}(u))du < \infty$.

Assumption B. The two score generating functions J_i, $i = 1, 2$, satisfy

(i) $\int_0^1 |J_i(u))|^4 du < \infty$.

(ii) $\int_0^1 J_1(u)J_2(u)\,du \neq 0$.

(iii) $\lim_{T\to\infty} E\left(\left(J_i(F(U_{p+1})) - J_i\left(\frac{R_{p+1}}{T-p+1}\right)\right)^4\right) = 0$ where R_{p+1}, is the rank of U_{p+1}, among $U_{p+1}, \ldots, U_T$.

(iv) $J_i(1-u) = -J_i(u)$.

(v) $J_i(F(v))$ are continuously differentiable and $|J_i(F(v))| \leq K|v|^m$ where m may be 0 or 1.

(vi) Let $J_i^{*\prime}(v) = dJ_i(F(v))/dv$, then $|J_i^{*\prime}(v)| \leq K|v|^m$ where m may be 0 or 1.

(vii) $|d^2 J_i(F(v))/dv^2| \leq K$

(viii) $E(J_1^{*\prime}(U_1)) \neq 0$ and $E(J_2(F(U_1))U_1) \neq 0$.

(ix) $|J_i(u) - J_i(v)| \leq K|u-v|$.

REMARK 3.1. Assumption B(iii) is verified for example if J_i 's satisfy

$$|d^j J_i(u)/du^j| \leq K(u(1-u))^{-j-1/4+\delta_l}, \quad j = 0, 1\ , 0 < u < 1$$

for some $\delta_l > 0$, $l = 1, 2$. This result can be obtained by similar arguments as those given in Theorem 3.6.6 of Puri and Sen (1971).

REMARK 3.2. Assumptions B(v) and B(vi) are satisfied, for example, if $J_i(u) = \Phi^{-1}(u)$ and F is normal, $J_i(u) = 2u - 1$ and F is normal or logistic or $J_i(u) = ln(u/(1-u))$ and F is logistic, $i = 1, 2$.

4. ASYMPTOTIC DISTRIBUTION

THEOREM 4.1. Assume that $(Z_1, \ldots, Z_T)$ is a stationary AR(p) process and that assumptions A and B hold. If, in addition, $\widehat{\boldsymbol{\phi}}_T$ is a sequence of estimators satisfying

$$T^{1/2}\mathbf{W}_T(\mathbf{R}_T(\widehat{\boldsymbol{\phi}}_T), \widehat{\boldsymbol{\phi}}_T) \xrightarrow{p} \mathbf{0}, \quad \text{as } T \to \infty$$

and such that $T^{1/2}(\widehat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0)$ is bounded in probability, then there exists a statistic Q_4 and a sequence of estimators $\widehat{\boldsymbol{\phi}}_T^m$ that satisfy

(i) $Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\phi}}_T^m))$ is asymptotically distributed as chi-squared with degrees of freedom $m-p$,

(ii) for exery $\epsilon > 0$ and $\delta > 0$ there exists $m_0 > 0$ and $T_0 > 0$ such that for $m \geq m_0$ and $T \geq T_0$

$$P(|Q_3\ (\mathbf{R}_T(\widehat{\boldsymbol{\phi}}_T)) - Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\phi}}_T^m))| \geq \epsilon) \leq \delta.$$

The proof of this theorem is given in the Appendix.

REMARK 4.1. The statistic Q_4 is defined as Q_3 but using a truncated version of the RA-estimators. A precise definition of Q_4 is given in (6.1.5) of the Appendix.

REMARK 4.2. We have only been able to prove Theorem 4.1 in the AR(p) case. However we conjecture that the result is also valid for the ARMA(p, q) model. The only part of the proof which is not valid for an arbitrary stationary ARMA(p,q) model is Proposition 6.2.3 and (6.2.46).

REMARK 4.3. Assumption B(ix) used in Theorem 4.1 is very restrictive. It is satisfied by the Wilcoxon scores generating function, $J_i(u) = 2u-1$, but not by the Normal scores generating function, $J_i(u) = \Phi^{-1}(u)$. The only part of the proof where this assumption is used is also Proposition 6.2.3 and (6.2.46). However, according to our Monte Carlo results we conjecture that Theorem 4.1 holds under weaker conditions which include $J_1(u) = J_2(u) = \Phi^{-1}(u)$.

5. THE MONTE CARLO STUDY

5.1. Description of the Study. The behaviour of the Q_1, Q_2 and Q_3 statistics has been studied for the AR(1) and MA(1) models without outliers (purely Gaussian) and with additive outliers. The AR(1) and MA(1) additive outlier models used in this Monte Carlo study assume that the observations $(Z_1, \dots, Z_T)$ satisfy

$$Z_t = W_t + V_t \quad 1 \le t \le T. \tag{5.1}$$

For the AR(1) model W_t in (5.1) are given by

$$W_t = \phi W_{t-1} + U_t \quad 1 \le t \le T,$$

and for the MA(1) model

$$W_t = -\theta U_{t-1} + U_t \quad 1 \le t \le T,$$

where the U_t are i.i.d. random variables with distribution N(0,1). The variables V_t, $1 \le t \le T$ are i.i.d. with distribution

$$H = (1-\varepsilon)\delta_0 + \varepsilon\, \mathrm{N}(0, \tau^2)$$

where δ_0 is the distribution which assigns probability 1 to the origen. Then a fraction $1-\varepsilon$ of the time Z_t coincides with the Gaussian model W_t and the rest of the time Z_t is equal to W_t plus some Gaussian noise V_t. The purely Gaussian case corresponds to $\varepsilon = 0$.

For each model three values of ε (0; 0.05; 0.10) and three values of τ (3; 10; 20) have been investigated. The Q_3 statistic considered is based on RAR–estimators with $J_1 = J_2$. This common function is called J. Two J–functions: (i)$J(u) = \Phi^{-1}(u)$, (ii) $J(u) = 2u - 1$, have been considered. These Q_3 statistics have been compared to Q_2 with Mallows type $\eta(u, v)$ and Huber ψ functions

$$\psi_{H,c}(u) = \operatorname{sign}(u) \min(|u|, c),$$

for 2 values of the tuning constants c: 1.65, 1.34. These values of c were chosen so that, under the purely Gaussian ARMA model, the corresponding RA–estimators have approximately the same efficiency as the selected RAR–estimators. The scale parameter was estimated by the median of $(|\widehat{U}_{p+1}|, \ldots, |\widehat{U}_T|)/0.6745$.

The proportion of Q_1, Q_2 and Q_3 values exceeding three nominal levels (0.01; 0.05; 0.1) of the χ^2_{m-1} distribution has been studied. Also the empirical mean and variance of each statistic have been obtained. The AR(1) cases with $\phi = 0.5$ and 0.8 and the MA(1) cases with $\theta = -0.5$ and -0.8 were investigated.

Moreover, the empirical power of Q_1, Q_2 and Q_3 was studied when the actual model was $AR(2)$ or $MA(2)$ but was identified as $AR(1)$ or $MA(1)$ respectively. Two second–order autoregressions were considered. The first one ($\phi_1 = 0.5$ and $\phi_2 = 0.28$) was chosen so that the empirical power of Q_1 was near 0.5. For the second one ($\phi_1 = 0.5$ and $\phi_2 = 0.38$) the empirical power was near 0.8. Also two second–order moving averages were studied. The first one ($\theta_1 = 0.5$ and $\theta_2 = 0.32$) was chosen so that the empirical power of Q_1 was near 0.5. For the second one ($\theta_1 = 0.5$ and $\theta_2 = 0.5$) the empirical power was near 0.8. Two nominal levels of significance (0.05; 0.1) were examined.

There were performed 500 replications, with sample size 100 and $m = 8$. Several routines given in Press, Flannery, Teukolsky and Vetterling (1986) were used: RAN1 (random number generator), GASDEV (Standard Normal generator), RANK (rearrengement of an array) and ZBRAK (bracketing of a root). The computer programmes where written in FORTRAN and performed in an IBM 3032 at the Centro de Estudios Superiores para el Procesamiento de la Información (CESPI), Universidad de La Plata.

5.2. ***Discussion of the results.*** For the AR model and $\phi = 0.5$ Table 1 shows that the significance levels of Q_1, Q_2 and Q_3 were not very much disturbed by additive outliers. However, for $\tau = 10$ and $\tau = 20$, the empirical variances were significatively different from its asymptotic value 14. For Q_1 this difference was larger. The results for the other additive outlier models are not reported here because they are qualitatively similar to those given.

If $\phi = 0.8$ Tables 2 and 3 show for the AR model that under the purely Gaussian model the distributions of Q_1, Q_2 and Q_3 were reasonably approximated by the asymptotic theory. However, if there were outliers the χ^2_7 is a poor approximation for the Q_1 statistic's distribution. On the other hand for $\tau = 3$ and $\varepsilon = 0.05$ the significance levels of Q_2 and Q_3

were similar to the nominal levels in all cases. Further for $\tau = 3$, $\varepsilon = 0.10$; $\tau = 10$, $\varepsilon = 0.05$ and $\tau = 20$, $\varepsilon = 0.05$ the significance levels of Q_2 ($c = 1.34$) and Q_3 ($\mathrm{J}(u) = 2u - 1$) were closer to the nominal levels than the significance levels of the other statistics considered.

For the MA model and $\theta = -0.5$ the conclusions related to the significance levels are similar to those of the AR model and $\phi = 0.5$. Some of these results are shown in Table 4. Further for the MA model and $\theta = -0.8$ Table 5 shows the unstable behaviour of the statistics considered .

In Tables 6 and 7 it can be seen that for the AR model under the purely Gaussian model the powers of Q_2 ($c = 1.65$, $c = 1.34$) and Q_3 ($\mathrm{J}(u) = \Phi^{-1}(u)$, $\mathrm{J}(u) = 2u - 1$) were similar to that of Q_1. For additive outliers model with $\tau = 10$ and $\tau = 20$ the power of Q_1 is significantly lower than Q_2 ($c = 1.65$, $c = 1.34$) and Q_3 ($\mathrm{J}(u) = \Phi^{-1}(u)$, $\mathrm{J}(u) = 2u - 1$). Further the powers of Q_2 and Q_3 were insensitive to departures from normality of the $U_t's$.

Finally Table 8 shows that for the MA model the power is more sensitive to additive outliers than for the AR model in all cases. However the powers of Q_2 ($c = 1.65$, $c = 1.34$) and Q_3 ($\mathrm{J}(u) = \Phi^{-1}(u)$, $\mathrm{J}(u) = 2u - 1$) are significatively higher than those of Q_1 mainly for $\tau = 10$.

REMARK 5.1. The stability of the significance levels of Q_1 for $\phi = 0.5$ and $\theta = -0.5$ is in accordance with Anderson and Walker (1964) who have shown that the asymptotic normality of the residual autocorrelations does not require normality of the $U_t's$.

REMARK 5.2. The nonstandard contamination appearing in Li (1988) leads to different conclusions than ours.

TABLE 1

Empirical means, variances and significance levels of Q_1, Q_2 and Q_3 for AR(1) model and $\phi = 0.5$

	$\varepsilon = 0$				$\varepsilon = 0.05$							
					$\tau = 10$				$\tau = 20$			
			Nominal level				Nominal level				Nominal level	
Test statistics	**Mean**	**Var**	0.05	0.10	**Mean**	**Var**	0.05	0.10	**Mean**	**Var**	0.05	0.10
Q_1	6.75	12.67	0.03	0.07	5.62	22.73	0.06	0.09	4.65	30.01	0.06	0.09
Q_2 (c=1.65)	6.62	12.10	0.03	0.07	7.02	14.72	0.05	0.10	7.12	15.58	0.06	0.11
Q_3 (J(u)=$\Phi^{-1}(u)$)	6.80	12.20	0.03	0.09	7.24	16.05	0.06	0.11	7.34	17.58	0.07	0.11
Q_2 (c=1.34)	6.63	12.05	0.03	0.07	6.99	14.28	0.05	0.10	7.04	14.71	0.06	0.10
Q_3 (J(u)=$2u-1$)	6.80	12.11	0.03	0.08	7.15	14.66	0.03	0.10	7.19	15.14	0.05	0.11

TABLE 2
Empirical means, variances and significance levels of Q_1, Q_2 and Q_3 for AR(1) model and $\phi = 0.8$

					$\tau = 3$							
	$\varepsilon = 0$				$\varepsilon = 0.05$				$\varepsilon = 0.10$			
			Nominal level				Nominal level				Nominal level	
Test statistics	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10
Q_1	6.86	13.65	0.04	0.08	8.30	21.37	0.09	0.17	9.33	28.25	0.16	0.25
Q_2 (c=1.65)	6.73	13.25	0.04	0.07	7.24	15.27	0.06	0.10	7.78	18.02	0.07	0.14
Q_3 ($J(u)=\Phi^{-1}(u)$)	6.89	13.35	0.04	0.08	7.50	17.17	0.07	0.13	8.11	19.20	0.08	0.16
Q_2 (c=1.34)	6.77	13.62	0.04	0.07	7.11	14.50	0.05	0.09	7.51	17.28	0.06	0.13
Q_3 ($J(u)=2u-1$)	6.93	14.15	0.04	0.09	7.22	14.95	0.06	0.10	7.49	15.49	0.06	0.12

TABLE 3

Empirical means, variances and significance levels of Q_1, Q_2 and Q_3 for AR(1) model and $\phi = 0.8$

	$\tau = 10$								$\tau = 20$							
	$\epsilon = 0.05$				$\epsilon = 0.10$				$\epsilon = 0.05$				$\epsilon = 0.10$			
			Nominal level				Nominal level				Nominal level				Nominal level	
Test statistics	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10
Q_1	8.66	40.48	0.16	0.21	8.08	29.83	0.14	0.20	6.49	41.67	0.10	0.15	6.05	21.75	0.08	0.10
Q_2 (c=1.65)	7.76	18.14	0.09	0.14	9.19	24.21	0.13	0.23	7.95	19.23	0.09	0.15	9.76	25.85	0.17	0.27
Q_3 (J(u)=$\Phi^{-1}(u)$)	8.26	21.71	0.10	0.16	9.63	26.04	0.16	0.26	8.48	23.54	0.11	0.17	10.11	28.26	0.19	0.29
Q_2 (c=1.34)	7.47	16.54	0.06	0.13	8.54	20.28	0.09	0.19	7.60	17.17	0.07	0.13	8.94	22.87	0.13	0.21
Q_3 (J(u)=$2u-1$)	7.56	17.32	0.07	0.13	8.42	18.45	0.09	0.17	7.66	17.89	0.08	0.13	8.77	22.05	0.11	0.20

TABLE 4
Empirical means, variances and significance levels of Q_1, Q_2 and Q_3 for MA(1) model and $\theta = -0.5$

					$\varepsilon = 0.05$							
	$\varepsilon = 0$				$\tau = 3$				$\tau = 10$			
			Nominal level				Nominal level				Nominal level	
Test statistics	**Mean**	**Var**	0.05	0.10	**Mean**	**Var**	0.05	0.10	**Mean**	**Var**	0.05	0.10
Q_1	7.02	14.86	0.06	0.12	6.68	12.81	0.03	0.10	5.63	21.40	0.07	0.09
Q_2 (c=1.65)	6.94	13.35	0.05	0.10	6.71	12.80	0.04	0.08	6.83	13.87	0.06	0.09
Q_3 (J(u)=$\Phi^{-1}(u)$)	7.15	14.46	0.04	0.11	6.83	12.85	0.04	0.09	6.90	13.91	0.05	0.09
Q_2 (c=1.34)	6.84	14.28	0.04	0.10	6.78	13.51	0.04	0.09	6.94	14.85	0.06	0.11
Q_3 (J(u)=$2u-1$)	7.14	16.08	0.07	0.12	7.09	14.20	0.05	0.10	7.18	15.07	0.05	0.12

TABLE 5

Empirical means, variances and significance levels of Q_1, Q_2 and Q_3 for MA(1) model and $\theta = -0.8$

					$\varepsilon = 0.05$							
	$\varepsilon = 0$				$\tau = 3$				$\tau = 10$			
			Nominal level				Nominal level				Nominal level	
Test statistics	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10	Mean	Var	0.05	0.10
Q_1	7.64	17.52	0.08	0.12	6.60	12.81	0.04	0.08	5.40	18.49	0.05	0.07
Q_2 (c=1.65)	7.71	17.76	0.09	0.15	6.89	14.29	0.06	0.09	7.13	15.21	0.07	0.11
Q_3 ($J(u)=\Phi^{-1}(u)$)	8.06	19.37	0.10	0.16	7.05	14.31	0.05	0.11	7.02	15.04	0.06	0.10
Q_2 (c=1.34)	7.71	17.50	0.09	0.13	6.91	14.14	0.06	0.10	7.10	16.57	0.06	0.09
Q_3 ($J(u)=2u-1$)	8.12	19.56	0.10	0.18	7.13	14.02	0.05	0.11	7.23	17.65	0.06	0.10

TABLE 6

Empirical power of Q_1, Q_2 and Q_3 for $\phi_1 = 0.5$ and $\phi_2 = 0.28$

	Nominal level=0.05							Nominal level=0.10						
		$\tau = 3$		$\tau = 10$		$\tau = 20$			$\tau = 3$		$\tau = 10$		$\tau = 20$	
Test statistics	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$
Q_1	0.43	0.44	0.43	0.22	0.13	0.13	0.06	0.57	0.57	0.53	0.28	0.20	0.16	0.11
Q_2 (c=1.65)	0.42	0.44	0.42	0.45	0.44	0.45	0.45	0.52	0.56	0.54	0.56	0.57	0.57	0.57
Q_3 (J(u)=$\Phi^{-1}(u)$)	0.38	0.39	0.40	0.41	0.39	0.43	0.41	0.49	0.54	0.50	0.55	0.50	0.56	0.50
Q_2 (c=1.34)	0.39	0.41	0.41	0.41	0.40	0.41	0.38	0.51	0.53	0.52	0.54	0.53	0.54	0.55
Q_3 (J(u)=$2u-1$)	0.35	0.37	0.38	0.37	0.36	0.36	0.34	0.46	0.50	0.48	0.49	0.48	0.50	0.46

TABLE 7

Empirical power of Q_1, Q_2 and Q_3 for $\phi_1 = 0.5$ and $\phi_2 = 0.38$

	Nominal level=0.05							Nominal level=0.10						
		$\tau = 3$		$\tau = 10$		$\tau = 20$			$\tau = 3$		$\tau = 10$		$\tau = 20$	
Test statistics	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$	$\epsilon = 0.05$	$\epsilon = 0.10$
Q_1	0.74	0.73	0.70	0.49	0.32	0.23	0.12	0.84	0.82	0.80	0.57	0.42	0.30	0.19
Q_2 (c=1.65)	0.72	0.71	0.69	0.69	0.68	0.68	0.68	0.81	0.83	0.79	0.80	0.78	0.79	0.78
Q_3 (J(u)=$\Phi^{-1}(u)$)	0.69	0.68	0.66	0.66	0.63	0.67	0.61	0.79	0.79	0.76	0.77	0.73	0.77	0.73
Q_2 (c=1.34)	0.71	0.71	0.68	0.68	0.64	0.67	0.67	0.79	0.82	0.77	0.79	0.77	0.78	0.76
Q_3 (J(u)=$2u - 1$)	0.67	0.65	0.63	0.64	0.59	0.64	0.58	0.76	0.78	0.73	0.76	0.72	0.75	0.70

TABLE 8

Empirical power of Q_1, Q_2 and Q_3

	$\theta_1 = 0.5$ $\theta_2 = 0.32$						$\theta_1 = 0.5$ $\theta_2 = 0.5$					
	Nominal level=0.05			Nominal level=0.10			Nominal level=0.05			Nominal level=0.10		
		$\epsilon = 0.05$			$\epsilon = 0.05$			$\epsilon = 0.05$			$\epsilon = 0.05$	
Test statistics	$\epsilon = 0$	$\tau = 3$	$\tau = 10$	$\epsilon = 0$	$\tau = 3$	$\tau = 10$	$\epsilon = 0$	$\tau = 3$	$\tau = 10$	$\epsilon = 0$	$\tau = 3$	$\tau = 10$
Q_1	0.41	0.21	0.10	0.58	0.33	0.14	0.73	0.41	0.15	0.86	0.59	0.20
Q_2 (c=1.65)	0.36	0.24	0.17	0.54	0.38	0.27	0.71	0.53	0.40	0.83	0.69	0.55
Q_3 ($J(u)=\Phi^{-1}(u)$)	0.41	0.27	0.18	0.56	0.40	0.28	0.74	0.54	0.39	0.87	0.68	0.56
Q_2 (c=1.34)	0.34	0.23	0.17	0.48	0.39	0.28	0.66	0.52	0.40	0.79	0.69	0.57
Q_3 ($J(u)=2u-1$)	0.38	0.27	0.18	0.52	0.38	0.29	0.67	0.54	0.37	0.80	0.71	0.57

6. Appendix

6.1. Notation and Definitions. Given $\boldsymbol{\beta} = (\beta_1, \dots, \beta_h) \in \Re^h$, $\boldsymbol{\beta}(B)$ denotes the polynomial operator $\boldsymbol{\beta}(B) = 1 - \beta_1 B - \cdots - \beta_h B^h$, where 1 is the identity operator and B the backward shift operator.

Define

$$R^{*h} = \{\boldsymbol{\beta} \in \Re^h : \boldsymbol{\beta}(B) \text{ has all the roots with absolute value} > 1\} .$$

Since Z_t is stationary $\boldsymbol{\phi}_0 \in R^{*p}$ and since it is invertible $\boldsymbol{\theta}_0 \in R^{*q}$.

Given $\boldsymbol{\phi} \in R^{*p}$, $\boldsymbol{\theta} \in R^{*q}$ let $g_i(\boldsymbol{\phi}, \boldsymbol{\theta})$ be defined as in the beginning of Section 2, i.e., by

$$\boldsymbol{\theta}^{-1}(B)\boldsymbol{\phi}(B) = \sum_{i=0}^{\infty} g_i(\boldsymbol{\phi}, \boldsymbol{\theta}) B^i .$$

It is easy to prove that the functions g_i are continuously differentiable for $\boldsymbol{\phi} \in R^{*p}$ and $\boldsymbol{\theta} \in R^{*q}$. Moreover, given $C_1 \subset R^{*p}$ and $C_2 \subset R^{*q}$, compact sets, there exist, $A^* > 0$, $0 < b < 1$ such that

$$\sup\left\{|g_i(\boldsymbol{\phi}, \boldsymbol{\theta})| : \boldsymbol{\phi} \in C_1,\ \boldsymbol{\theta} \in C_2\right\} \le A^* b^i \tag{6.1.1}$$

$$\sup\left\{\left|\frac{\partial g_i(\boldsymbol{\phi}, \boldsymbol{\theta})}{\partial \phi_l}\right| : \boldsymbol{\phi} \in C_1,\ \boldsymbol{\theta} \in C_2\right\} \le A^* b^i, \quad 1 \le l \le p \tag{6.1.2}$$

$$\sup\left\{\left|\frac{\partial g_i(\boldsymbol{\phi}, \boldsymbol{\theta})}{\partial \theta_l}\right| : \boldsymbol{\phi} \in C_1,\ \boldsymbol{\theta} \in C_2\right\} \le A^* b^i, \quad 1 \le l \le q. \tag{6.1.3}$$

Given $\boldsymbol{\lambda} = (\boldsymbol{\phi}, \boldsymbol{\theta})$, define the residuals of order k by

$$U_t^{(k)}(\boldsymbol{\lambda}) = \sum_{i=0}^{k} g_i(\boldsymbol{\phi}, \boldsymbol{\theta}) Z_{t-i}, \quad 1 \le k \le \infty.$$

Note that

$$U_t(\boldsymbol{\lambda}) = U_t^{(t-1)}(\boldsymbol{\lambda}), \quad p + 1 \le t \le T.$$

Let us now define

$$\gamma_{4,i}(\mathbf{U}_T(\boldsymbol{\lambda})) = \sum_{t=p+1+i}^{T} J_1(F(U_t(\boldsymbol{\lambda})))J_2(F(U_{t-i}(\boldsymbol{\lambda})), \quad 0 \le i \le T - p - 1$$

and

$$W_{T,j}^*(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (T - j - p)^{-1} \sum_{h=0}^{T-j-p-1} \gamma_{4,h+j}(\mathbf{U}_T(\boldsymbol{\lambda})) s_h(\boldsymbol{\phi}), \quad 1 \le j \le p,$$

$$W_{T,p+j}^*(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (T - j - p)^{-1} \sum_{h=0}^{T-j-p-1} \gamma_{4,h+j}(\mathbf{U}_T(\boldsymbol{\lambda})) t_h(\boldsymbol{\theta}), \quad 1 \le j \le q.$$

Also put

$$\mathbf{W}_T^*(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (W_{T,1}^*(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}), \ldots, W_{T,p+q}^*(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta})).$$

Observe that

$$\mathbf{W}_T^*(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*), \widehat{\boldsymbol{\phi}}_T^*, \widehat{\boldsymbol{\theta}}_T^*) = \mathbf{0}$$

are the equations of the RA–estimators $\widehat{\boldsymbol{\lambda}}_T^* = (\widehat{\boldsymbol{\phi}}_T^*, \widehat{\boldsymbol{\theta}}_T^*)$ with $\eta(u, v) = J_1(F(u))J_2(F(v))$.

Let

$$\rho_{4,k}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*) = \gamma_{4,k}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))/\gamma_{4,0}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*)), \qquad 1 \le k \le m. \tag{6.1.4}$$

Then define the portmanteau statistic based on the RA–estimators $\widehat{\boldsymbol{\lambda}}_T^*$ by

$$Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*) = \xi^{-1}T(T+2)\sum_{k=1}^{m}(T-k)^{-1}\rho_{4,k}^2(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*)). \tag{6.1.5}$$

Let us also define

$$W_{T,j}^{*,m}(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (m-j+1)^{-1}\sum_{h=0}^{m-j}\gamma_{4,h+j}(\mathbf{U}_T(\boldsymbol{\lambda}))s_h(\boldsymbol{\phi}), \quad 1 \le j \le p,$$

$$W_{T,p+j}^{*,m}(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (m-j+1)^{-1}\sum_{h=0}^{m-j}\gamma_{4,h+j}(\mathbf{U}_T(\boldsymbol{\lambda}))t_h(\boldsymbol{\theta}), \quad 1 \le j \le q$$

and

$$\mathbf{W}_T^{*,m}(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}) = (W_{T,1}^{*,m}(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta}), \ldots, W_{T,p+q}^{*,m}(\mathbf{U}_T(\boldsymbol{\lambda}), \boldsymbol{\phi}, \boldsymbol{\theta})).$$

Then $\widehat{\boldsymbol{\lambda}}_T^m = (\widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m)$ is defined as a sequence satisfying

$$\mathbf{W}_T^{*,m}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m), \widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m) = \mathbf{0}$$

and the corresponding portmanteau statistic is obtained replacing $\widehat{\boldsymbol{\lambda}}_T^*$ by $\widehat{\boldsymbol{\lambda}}_T^m$ in (6.1.4) and (6.1.5).

Let, for $1 \le k \le T-1$,

$$c_k = ((T+2)/(T-k))^{1/2},$$

$$\tilde{\rho}_{3,k}(\mathbf{R}_T(\boldsymbol{\lambda}) = c_k\gamma_{3,k}(\mathbf{R}_T(\boldsymbol{\lambda}))/\gamma_{3,0}(\mathbf{R}_T(\boldsymbol{\lambda})),$$

and

$$\tilde{\rho}_{4,k}(\mathbf{U}_T(\boldsymbol{\lambda}) = c_k\gamma_{4,k}(\mathbf{U}_T(\boldsymbol{\lambda}))/\gamma_{4,0}(\mathbf{U}_T(\boldsymbol{\lambda})).$$

Now denote for $s \geq 1$

$$\tilde{\gamma}_3^s(\mathbf{R}_T(\boldsymbol{\lambda})) = (c_1\gamma_{3,1}(\mathbf{R}_T(\boldsymbol{\lambda}))/T, \ldots, c_s\gamma_{3,s}(\mathbf{R}_T(\boldsymbol{\lambda}))/T),$$

$$\tilde{\rho}_3^s(\mathbf{R}_T(\boldsymbol{\lambda})) = (\tilde{\rho}_{3,1}(\mathbf{R}_T(\boldsymbol{\lambda})), \ldots, \tilde{\rho}_{3,s}(\mathbf{R}_T(\boldsymbol{\lambda}))).$$

and

$$\tilde{\gamma}_4^s(\mathbf{U}_T(\boldsymbol{\lambda})) = (c_1\gamma_{4,1}(\mathbf{U}_T(\boldsymbol{\lambda}))/T, \ldots, c_s\gamma_{4,s}(\mathbf{U}_T(\boldsymbol{\lambda}))/T),$$

$$\tilde{\rho}_4^s(\mathbf{U}_T(\boldsymbol{\lambda})) = (\tilde{\rho}_{4,1}(\mathbf{U}_T(\boldsymbol{\lambda})), \ldots, \tilde{\rho}_{4,s}(\mathbf{U}_T(\boldsymbol{\lambda}))).$$

Finally, let us denote $\mathbf{R}_T = (R_{p+1}, \ldots, R_T)$ and $\mathbf{U}_T = (U_{p+1}, \ldots, U_T)$ where R_j, $p+1 \leq j \leq T$, is the rank of U_j, $p+1 \leq j \leq T$, among $U_{p+1}, \ldots, U_T$.

6.2. Asymptotic Distribution of Q_3 . In this section we derive the asymptotic distribution of Q_3 through the asymptotic distribution of Q_4.

PROPOSITION 6.2.1. *Assume that assumptions B(i), B(ii) B(iii) and B(iv) hold. Then*

$$T^{1/2}(\tilde{\rho}_3^m(\mathbf{R}_T) - \tilde{\rho}_4^m(\mathbf{U}_T)) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty .$$

PROOF: To prove the proposition it suffices to show that for $1 \leq j \leq m$

(6.2.1) $$T^{1/2}(\tilde{\rho}_{3,j}(\mathbf{R}_T) - \tilde{\rho}_{4,j}(\mathbf{U}_T)) \xrightarrow{p} 0, \qquad \text{as } T \to \infty .$$

We will first prove that

(6.2.2) $$T^{-1/2}(\gamma_{3,j}(\mathbf{R}_T) - \gamma_{4,j}(\mathbf{U}_T)) \xrightarrow{p} 0, \qquad \text{as } T \to \infty .$$

We have

$$T^{-1/2}(\gamma_{3,j}(\mathbf{R}_T) - \gamma_{4,j}(\mathbf{U}_T)) = T^{-1/2} \sum_{t=p+1+j}^{T} [J_1(\frac{R_t}{T-p+1})J_2(\frac{R_{t-j}}{T-p+1})$$
$$- J_1(F(U_t))J_2(F(U_{t-j})) - \overline{J}_T + \overrightarrow{J}_T^*(\mathbf{U}_T)] + T^{-1/2}(T-p-j)[\overline{J}_T - \overrightarrow{J}_T^*(\mathbf{U}_T)],$$

where

$$\overline{J}_T = [(T-p)(T-p-1)]^{-1} \sum_{i=1}^{T-p} \sum_{\substack{j=1 \\ i \neq j}}^{T-p} J_1(\frac{i}{T-p+1})J_2(\frac{j}{T-p+1})$$

and

$$\overline{J}_T^*(\mathbf{U}_T) = [(T-p)(T-p-1)]^{-1} \sum_{t_1=p+1}^{T} \sum_{\substack{t_2=p+1 \\ t_1 \neq t_2}}^{T} J_1(F(U_{t_1}))J_2(F(U_{t_2})).$$

From the weak law of large numbers and the central limit theorem we obtain

$$T^{-1/2}(T-p-j)[\overline{J}_T - \overline{J}_T^*(\mathbf{U}_T)] \xrightarrow{p} 0,$$

as $T \to \infty$.

Hence, it sufices to show that

$$\lim_{T\to\infty} E\,[\,\Delta_{T,j}^2(\mathbf{U}_T)] = 0\,,$$

where

$$\Delta_{T,j}(\mathbf{U}_T) = T^{-1/2} \sum_{t=p+1+j}^{T} [J_1(\frac{R_t}{T-p+1})J_2(\frac{R_{t-j}}{T-p+1})$$

$$-J_1(F(U_t))J_2(F(U_{t-j})) - \overline{J}_T + \overline{J}_T^*(\mathbf{U}_T)].$$

Let $V_s = F(U_s), p+1 \leq s \leq T$ and let $\mathbf{V}_{(\cdot)} = (V_{(1)}, \ldots, V_{(T-p)})$ where $\mathbf{V}_{(i)}, 1 \leq i \leq T-p$, is the ith order statistic.

Define

$$\alpha(R_t, R_{t-j}, V_{(R_t)}, V_{(R_{t-j})}) = J_1(\frac{R_t}{T-p+1})J_2(\frac{R_{t-j}}{T-p+1}) - J_1(V_{(R_t)})J_2(V_{(R_{t-j})}).$$

Hence we have

$$\Delta_{T,j}(\mathbf{U}_T) = T^{-1/2}[\sum_{t=p+1+j}^{T} \alpha(R_t, R_{t-j}, V_{(R_t)}, V_{(R_{t-j})}) - (T-p-j)(\overline{J}_T - \overline{J}_T^*(\mathbf{U}_T)).$$

Then,

$$E[\Delta_{T,j}^2(\mathbf{U}_T)] = T^{-1}E[E[(S_T^j(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) - (T-p-j)(\overline{J}_T - \overline{J}_T^*(\mathbf{U}_T)))^2 | \mathbf{V}_{(\cdot)}]]$$

where

$$S_T^j(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) = \sum_{t=p+1+j}^{T} \alpha(R_t, R_{t-j}, V_{(R_t)}, V_{(R_{t-j})})$$

and

$$E(S_T^j(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) \,|\, \mathbf{V}_{(\cdot)}) = (T-p-j)(\overline{J}_T - \overline{J}_T^*(\mathbf{U}_T)).$$

Then

$$E[\Delta^2_{T,j}(\mathbf{U}_T)] = T^{-1}E[\mathrm{Var}(S^j_T(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) \mid \mathbf{V}_{(\cdot)})]$$

where $\mathrm{Var}(S^j_T(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) \mid \mathbf{V}_{(\cdot)})$ denotes the conditional variance of $S^j_T(\mathbf{U}_T, \mathbf{V}_{(\cdot)})$.

From Lemma 6.2.1 of Ferretti, Kelmansky and Yohai (1991) it follows that

$$\mathrm{Var}(S^j_T(\mathbf{U}_T, \mathbf{V}_{(\cdot)}) \mid \mathbf{V}_{(\cdot)}) \leq T(3 + TK_T)E[(\alpha(R_{p+1+j}, R_{p+1}, V_{(R_{p+1+j})}, V_{(R_{p+1})}))^2 \mid \mathbf{V}_{(\cdot)}].$$

Hence

$$E[\Delta^2_{T,j}(\mathbf{U}_T)] \leq (3 + TK_T)E[(\alpha(R_{p+1+j}, R_{p+1}, V_{(R_{p+1+j})}, V_{(R_{p+1})}))^2].$$

Moreover, from assumptions B(i) and B(iii) and the Cauchy-Schwartz inequality we obtain that

$$\lim_{T\to\infty} E\{[J_1(F(U_{p+1}))J_2(F(U_{p+2})) - J_1(\frac{R_{p+1}}{T-p+1})J_2(\frac{R_{p+2}}{T-p+1})]^2\} = 0.$$

Hence (6.2.2) holds.

Therefore from the weak law of large numbers and Theorem 7.7.5 of Anderson (1971) one now concludes (6.2.1). ∎

PROPOSITION 6.2.2. *Assume that assumptions A(i), B(i), B(ii), B(iii)and B(iv) hold. Let* $\mathbf{A} \in \Re^{p+q}$ *and put* $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0 + T^{-1/2}\mathbf{A}$ *then*

$$T^{1/2}(\tilde{\rho}^m_3(\mathbf{R}_T(\boldsymbol{\lambda})) - \tilde{\rho}^m_4(\mathbf{U}_T(\boldsymbol{\lambda})) \xrightarrow{P} 0\,, \quad \text{as } T \to \infty\,.$$

PROOF: The proof of this proposition is an immediate consequence of the Proposition 6.2.1, the Proposition 6.2.2 of Ferretti Kelmansky and Yohai (1991), and the definition of contiguity. ∎

Let us denote the usual Euclidean norm by $\|\cdot\|_2$.

PROPOSITION 6.2.3. *Assume that* $(Z_1, \ldots, Z_T)$ *is a stationary AR(p) process and that assumptions A and B hold. Let* $\mathbf{A} \in \Re^p$ *and* $A_0 > 0$ *and put* $\boldsymbol{\phi} = \boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}$. *Then*

$$\sup_{\|\mathbf{A}\|_2 \leq A_0} \|T^{1/2}(\tilde{\rho}^m_3(\mathbf{R}_T(\boldsymbol{\phi})) - \tilde{\rho}^m_4(\mathbf{U}_T(\boldsymbol{\phi}))\|_2 \xrightarrow{P} 0, \quad \text{as } T \to \infty. \tag{6.2.3}$$

PROOF:

Due to Proposition 6.2.2 in order to prove (6.2.3) it suffices to show that for all $A_0 > 0$,

$$\sup_{\|\mathbf{A}\|_2 \leq A_0, \|\boldsymbol{\epsilon}\|_2 \leq \epsilon_0} \|T^{1/2}(\tilde{\rho}^m_3(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon})) - \tilde{\rho}^m_3(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}))\|_2 \xrightarrow{P} 0 \tag{6.2.4}$$

and

$$\text{(6.2.5)} \qquad \sup_{\|\mathbf{A}\|_2 \le A_0, \|\boldsymbol{\epsilon}\|_2 \le \epsilon_0} \|T^{1/2}(\tilde{\boldsymbol{\rho}}_4^m(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon})) - \tilde{\boldsymbol{\rho}}_4^m(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}))\|_2 \xrightarrow{p} 0$$

as $T \to \infty$ and $\epsilon_0 \to 0$.

We will first show (6.2.4).

Let

$$\overline{\gamma}_T = T^{-1} \sum_{j=1}^{T-p} J_1\left(\frac{j}{T-p+1}\right) J_2\left(\frac{j}{T-p+1}\right),$$

$$S_{1,j}(T,\mathbf{A},\boldsymbol{\epsilon}) = T^{-1/2} \sum_{t=p+1+j}^{T} \left|J_2\left(\frac{R_{t-j}(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon}))}{T-p+1}\right)\right| \left|J_1\left(\frac{R_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon}))}{T-p+1}\right) - J_1\left(\frac{R_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})}{T-p+1}\right)\right|$$

and

$$S_{2,j}(T,\mathbf{A},\boldsymbol{\epsilon}) = T^{-1/2} \sum_{t=p+1+j}^{T} \left|J_1\left(\frac{R_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})}{T-p+1}\right)\right| \left|J_2\left(\frac{R_{t-j}(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon}))}{T-p+1}\right) - J_2\left(\frac{R_{t-j}(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})}{T-p+1}\right)\right|.$$

For $1 \le j \le m$ we have

$$\begin{aligned} &|T^{1/2}(\tilde{\rho}_{3,j}(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon})) - \tilde{\rho}_{3,j}(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}))| \\ \text{(6.2.6)} \qquad &\le |1/\overline{\gamma}_T|(S_{1,j}(T,\mathbf{A},\boldsymbol{\epsilon}) + S_{2,j}(T,\mathbf{A},\boldsymbol{\epsilon})). \end{aligned}$$

Given $\mathbf{X} \in \Re^h$ and $v \in \Re$ define $F_h(\mathbf{X}, v)$ as the empirical distribution determined by $\mathbf{X}$. Therefore

$$F_h(\mathbf{X}, v) = \frac{\sum_{i=1}^h I(X_i \le v)}{h}$$

where $I(B)$ denotes the indicator of the event B. Let $\boldsymbol{\Delta} \in \Re^p$ then we have

$$\text{(6.2.7)} \qquad R_t(\boldsymbol{\phi}_0 + T^{-1/2}\boldsymbol{\Delta}) = (T-p)F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0 + T^{-1/2}\boldsymbol{\Delta}), U_t(\boldsymbol{\phi}_0 + T^{-1/2}\boldsymbol{\Delta})).$$

From assumption B(ix), (6.2.7) and the Cauchy-Schwartz's inequality we obtain

$$\begin{aligned} S_{1,j}(T,\mathbf{A},\boldsymbol{\epsilon}) \le{}& T^{-1/2}(T-p+1)^{-1}(T-p)K\Big[\sum_{j=1}^{T-p}[J_2(\frac{j}{T-p+1})]^2/(T-p)\Big]^{1/2} \\ &\Big[\sum_{t=p+1+j}^{T} [(T-p)^{1/2}|F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon})), U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon}))) \\ &- F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}), U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}))|]^2\Big]^{1/2}. \end{aligned}$$

Let $F_{T-p}(u) = F_{T-p}(\mathbf{U}_T, u)$.

From Theorem 2 of Section 1 in Koul (1990), we have

$$\sup_{\substack{u \in \Re \\ \|\mathbf{A}\| \leq A_0, \|\boldsymbol{\epsilon}\| \leq \epsilon_0}} (T-p)^{1/2} |F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon})), u) - F_{T-p}(u)| = o_p(1) \tag{6.2.8}$$

$$\sup_{\substack{u \in \Re \\ \|\mathbf{A}\| \leq A_0}} (T-p)^{1/2} |F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}), u) - F_{T-p}(u)| = o_p(1) \tag{6.2.9}$$

From (6.2.8), 6.2.9) and the fact that $\mathbf{U}_T(\boldsymbol{\phi}_0) = \mathbf{U}_T$, it readily follows that

$$\begin{aligned} &(T-p)^{1/2} |F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon})), U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon}))) \\ &\qquad - F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon})))| \\ &+ (T-p)^{1/2} |F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}), U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})) \\ &\qquad - F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A}))| = \bar{o}_p(1) \end{aligned} \tag{6.2.10}$$

where $\bar{o}_p(1)$ is a sequence of stochastic processes converging to zero uniformly in probability over the set $\{\|\mathbf{A}\|_2 \leq A_0, \ \|\boldsymbol{\epsilon}\|_2 \leq \epsilon_0\}$.

From equation (5) in Theorem 1 of Section 3.2 in Koul (1990) we immediately obtain

$$\begin{aligned} \sup_{\substack{p+1 \leq t \leq T \\ \|\mathbf{A}\| \leq A_0, \|\boldsymbol{\epsilon}\| \leq \epsilon_0}} &(T-p)^{1/2} |F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon}))) - F_{T-p}(U_t) \\ &+ (T-p)^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon})'\mathbf{Z}_t f(U_t)| = o_p(1) \end{aligned}$$

and

$$\sup_{\substack{p+1 \leq t \leq T \\ \|\mathbf{A}\| \leq A_0}} (T-p)^{1/2} |F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})) - F_{T-p}(U_t) + (T-p)^{-1/2}\mathbf{A}'\mathbf{Z}_t f(U_t)| = o_p(1),$$

where $\mathbf{Z}_t = (Z_{t-1}, \ldots, Z_{t-p})$.

Then,

$$\begin{aligned} \sup_{\substack{p+1 \leq t \leq T \\ \|\mathbf{A}\| \leq A_0, \|\boldsymbol{\epsilon}\| \leq \epsilon_0}} &(T-p)^{1/2} |F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}(\mathbf{A} + \boldsymbol{\epsilon}))) - F_{T-p}(U_t(\boldsymbol{\phi}_0 + T^{-1/2}\mathbf{A})) \\ &+ (T-p)^{-1/2}\boldsymbol{\epsilon}'\mathbf{Z}_t f(U_t)| = o_p(1) \end{aligned} \tag{6.2.11}$$

From (6.2.10) and (6.2.11) it easily follows that

$$\sup_{\substack{p+1\le t\le T\\ \|\mathbf{A}\|\le A_0,\|\boldsymbol{\epsilon}\|\le\epsilon_0}} (T-p)^{1/2}|F_{T-p}(\mathbf{U}_T(\boldsymbol{\phi}_0+T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon})), U_t(\boldsymbol{\phi}_0+T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon})))$$

(6.2.12)

$$-F_{T-p}(U_t(\boldsymbol{\phi}_0+T^{-1/2}\mathbf{A}))+(T-p)^{-1/2}\boldsymbol{\epsilon}'\mathbf{Z}_t f(U_t)| = \bar{o}_p(1)+o_p(1)$$

From (6.2.12) and using some algebra one obtains

$$\sup_{\|\mathbf{A}\|\le A_0,\|\boldsymbol{\epsilon}\|\le\epsilon_0} S_{1,j}(T,\mathbf{A},\boldsymbol{\epsilon}) \xrightarrow{p} 0, \quad \text{as } T\to\infty,\ \varepsilon_0\to 0\,. \tag{6.2.13}$$

Similar arguments can be used to show that

$$\sup_{\|\mathbf{A}\|\le A_0,\|\boldsymbol{\epsilon}\|\le\epsilon_0} S_{2,j}(T,\mathbf{A},\boldsymbol{\epsilon}) \xrightarrow{p} 0, \quad \text{as } T\to\infty,\ \varepsilon_0\to 0\,. \tag{6.2.14}$$

Then, from the fact that

$$\bar{\gamma}_T \to \int_0^1 J_1(u)J_2(u)\,du \qquad \text{as } T\to\infty,$$

(6.2.6), (6.2.13) and (6.2.14) we obtain

$$\sup_{\|\mathbf{A}\|\le A_0,\|\boldsymbol{\epsilon}\|\le\epsilon_0} |T^{1/2}(\tilde{\rho}_{3,j}(\boldsymbol{\phi}_0+T^{-1/2}(\mathbf{A}+\boldsymbol{\epsilon}))-\tilde{\rho}_{3,j}(\boldsymbol{\phi}_0+T^{-1/2}\mathbf{A}))| \xrightarrow{p} 0,$$

as $T\to\infty$, $\varepsilon_0\to 0$. Therefore (6.2.4) follows.

Further, from the Mean Value Theorem, assumptions B(v) and B(vi), and the fact that

$$T^{-1}\sum_{t=p+1}^{T} J_1(F(U_t))J_2(F(U_t)) \xrightarrow{p} \int_0^1 J_1(u)J_2(u)\,du, \qquad \text{as } T\to\infty$$

we obtain (6.2.5). This completes the proof of the proposition. ∎

PROPOSITION 6.2.4. *Assume that assumptions A(i), B(ii) and B(v) hold. Then*

$$T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T) \xrightarrow{\mathcal{D}} N(\mathbf{0},\xi I_m), \qquad T\to\infty \tag{6.2.15}$$

where I_m is the $m\times m$ identity matrix and ξ is defined by (2.6).

PROOF: We will first prove that

$$T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T) \xrightarrow{\mathcal{D}} N(\mathbf{0},\omega I_m), \qquad T\to\infty, \tag{6.2.16}$$

where

$$\omega = E((J_1^2(F(U_1)))E((J_2^2(F(U_2))).$$

Let

$$\delta_{1,j}(\mathbf{U}_T) = T^{-1/2} \sum_{t=p+1+m}^{T} J_1(F(U_t))J_2(F(U_{t-j})), \quad 1 \le j \le m,$$

$$\delta_{2,j}(\mathbf{U}_T) = \begin{cases} T^{-1/2}\sum_{t=p+j+1}^{p+m} J_1(F(U_t))J_2(F(U_{t-j})), & 1 \le j \le m-1, \\ 0 & j = m, \end{cases}$$

$$\boldsymbol{\delta}_1(\mathbf{U}_T) = (\delta_{1,1}(\mathbf{U}_T), \ldots, \delta_{1,m}(\mathbf{U}_T) \text{ and } \boldsymbol{\delta}_2(\mathbf{U}_T) = (\delta_{2,1}(\mathbf{U}_T), \ldots, \delta_{2,m}(\mathbf{U}_T).$$

From Theorem 7.7.6 of Anderson (1971) it follows that $\boldsymbol{\delta}_1(\mathbf{U}_T)$ is asymptotically normally distributed with mean $\mathbf{0}$ and covariance matrix ωI_m. Moreover, we have

$$\boldsymbol{\delta}_2(\mathbf{U}_T) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty.$$

On the other hand, from Theorem 7.7.5 of Anderson (1971), we have, for $1 \le j \le m$,

$$c_j\; T^{-1/2}\gamma_{4,j}(\mathbf{U}_T) - (\delta_{1,j}(\mathbf{U}_T) + \delta_{2,j}(\mathbf{U}_T)) \xrightarrow{p} 0, \qquad \text{as } T \to \infty.$$

Therefore

$$T^{1/2}\tilde{\gamma}_4^m(\mathbf{U}_T) - (\boldsymbol{\delta}_1(\mathbf{U}_T) + \boldsymbol{\delta}_2(\mathbf{U}_T)) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty.$$

Hence (6.2.16) holds.

Further, from the weak law of large numbers we obtain

$$\gamma_{4,0}(\mathbf{U}_T) \xrightarrow{p} \int_0^1 J_1(u)J_2(u)\,du, \qquad \text{as } T \to \infty. \tag{6.2.17}$$

Then from (6.2.16) and (6.2.17) one now concludes (6.2.15). ∎

Let

$$\nu = E(J_1^{*\prime}(U_1))E(J_2(U_1))U_1) \tag{6.2.18}$$

where $J_1^{*\prime}(v)$ is defined in B(viii) and X^m is a $m \times (p+q)$ matrix given by

$$X_{i,j}^m = \begin{cases} s_{i-j}(\boldsymbol{\phi}_0) & \text{if } j \le i \le m \text{ and } 1 \le j \le p \\ -t_{i-j+p}(\boldsymbol{\theta}_0) & \text{if } j-p \le\ i \le m \text{ and } p+1 \le j \le p+q \\ 0 & \text{otherwise.} \end{cases} \tag{6.2.19}$$

PROPOSITION 6.2.5. *Assume that assumptions A(i), B(v), B(vi) and B(vii) hold. If, in addition,* $\widehat{\boldsymbol{\lambda}}_T^m$ *is a sequence of estimators satisfying*

$$T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m), \widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty$$

and such that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0)$ *is bounded in probability, then*

$$T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) = T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T) - \nu X^m T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) + o_p(1). \tag{6.2.20}$$

PROOF: We have

$$\begin{aligned} T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) =& T^{1/2}(\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)) \\ &+ T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)). \end{aligned} \tag{6.2.21}$$

We will first prove that

$$T^{1/2}(\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)) \xrightarrow{p} \mathbf{0}, \quad \text{as } T \to \infty. \tag{6.2.22}$$

Given $C \subset R^{*p} \times R^{*q}$, compact set, by (6.1.1), there exist, $\tilde{A} > 0$, $0 < \tilde{b} < 1$ such that

$$sup\{|U_t^{(k)}(\boldsymbol{\lambda}) - U_t^{(\infty)}(\boldsymbol{\lambda})| : \boldsymbol{\lambda} \in C, 1 \le k \le \infty\} \le \tilde{b}^{k+1}U_0^* \quad \text{a.s.} \tag{6.2.23}$$

and

$$sup\{|U_t^{(k)}(\boldsymbol{\lambda})| : \boldsymbol{\lambda} \in C, 1 \le k \le \infty\} \le U_t^* \quad \text{a.s.} \tag{6.2.24}$$

where

$$U_t^* = \tilde{A}\sum_{j=0}^{\infty}(j+1)\tilde{b}^j|U_{t-j}|.$$

Using the fact that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0)$ is bounded in probability, the Mean Value Theorem, B(v), B(vi), (6.2.23) and (6.2.24) we have

$$\begin{aligned} |J_1(F(U_t(\widehat{\boldsymbol{\lambda}}_T^m)))J_2(F(U_{t-j}(\widehat{\boldsymbol{\lambda}}_T^m))) - J_1(F(U_t^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)))J_2(F(U_{t-j}^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)))| \\ \le KU_0^*(\tilde{b}^tU_t^* + \tilde{b}^{t-j}U_{t-j}^*). \end{aligned}$$

Therefore,

$$T^{-1/2}c_j|\gamma_{4,j}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \gamma_{4,j}(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m))| \le T^{-1/2}c_jKU_0^*\sum_{t=p+1+j}^{T}(\tilde{b}^tU_t^* + \tilde{b}^{t-j}U_{t-j}^*).$$

Since $E(U_t^2) < \infty$, we have $\sum_{t=p+1+j}^{\infty} (\tilde{b}^t U_t^* + \tilde{b}^{t-j} U_{t-j}^*) < \infty$ a.s.. Then (6.2.22) follows.

Now we will prove that

$$T^{1/2}\tilde{\gamma}_4^m(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)) = T^{1/2}\tilde{\gamma}_4^m(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0)) - \nu X^m T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) + o_p(1). \tag{6.2.25}$$

By the Mean Value Theorem we have

$$T^{1/2}\tilde{\gamma}_4^m(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m)) = T^{1/2}\tilde{\gamma}_4^m(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0)) + T^{1/2}D\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T))(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) \tag{6.2.26}$$

where $D\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}))$ is the differential matrix of $\tilde{\boldsymbol{\gamma}}_4^m$ with respect to $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}_T$ satisfies $\|\tilde{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_0\|_2 \le \|\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0\|_2$.

We have, for $1 \le i \le m$,

$$\begin{aligned} c_i T^{-1}\frac{\partial \gamma_{4,i}(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}))}{\partial \lambda_j} =& c_i T^{-1}\sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t^{(\infty)}(\boldsymbol{\lambda}))J_2(F(U_{t-i}^{(\infty)}(\boldsymbol{\lambda})))\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j} \\ &+ c_i T^{-1}\sum_{t=p+1+i}^{T} J_1(F(U_t^{(\infty)}(\boldsymbol{\lambda})))J_2^{*\prime}(U_{t-i}^{(\infty)}(\boldsymbol{\lambda}))\frac{\partial U_{t-i}^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j} \end{aligned} \tag{6.2.27}$$

where $J_i^{*\prime}(v)$, $i = 1, 2$, are defined in assumption B(vi).

We will show that

$$\begin{aligned} & c_i T^{-1}\sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T))J_2(F(U_{t-i}^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T)))\Big[\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j}\Big]_{\boldsymbol{\lambda}=\tilde{\boldsymbol{\lambda}}_T} \\ & \qquad \xrightarrow{p} -\nu s_{i-j}(\boldsymbol{\phi}_0), \qquad 1 \le j \le p, \\ & c_i T^{-1}\sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T))J_2(F(U_{t-i}^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T)))\Big[\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j}\Big]_{\boldsymbol{\lambda}=\tilde{\boldsymbol{\lambda}}_T} \\ & \qquad \xrightarrow{p} \nu t_{i-j}(\boldsymbol{\theta}_0), \qquad p+1 \le j \le p+q, \end{aligned} \tag{6.2.28}$$

as $T \to \infty$.

It is easy to show that

$$\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j} = -\boldsymbol{\phi}^{-1}(B)U_{t-j}^{(\infty)}(\boldsymbol{\lambda}), \qquad 1 \le j \le p$$

and

$$\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j} = \boldsymbol{\theta}^{-1}(B)U_{t-j+p}^{(\infty)}(\boldsymbol{\lambda}), \qquad p+1 \le j \le p+q.$$

From the Mean Value Theorem, (6.2.23), (6.2.24), A(i), B(v), B(vi) and B(vii) we have, for $1 \leq j \leq p+q$,

$$c_i T^{-1} \sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T)) J_2(F(U_{t-i}^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T)))\Big[\frac{\partial U_t^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j}\Big]_{\boldsymbol{\lambda}=\tilde{\boldsymbol{\lambda}}_T}$$
$$= M_{i,j,T}(\mathbf{U}_T) + o_p(1)$$

where

$$M_{i,j,T}(\mathbf{U}_T) = \begin{cases} -c_i T^{-1} \sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t) J_2(F(U_{t-i}))\boldsymbol{\phi}_0^{-1}(B) U_{t-j} & if\ 1 \leq j \leq p \\ c_i T^{-1} \sum_{t=p+1+i}^{T} J_1^{*\prime}(U_t) J_2(F(U_{t-i}))\boldsymbol{\theta}_0^{-1}(B) U_{t-j+p} & if\ p+1 \leq j \leq p+q. \end{cases}$$

If $j \leq i$, by the ergodic theorem and the fact that $J_2(1-u) = -J_2(u)$ we have

$$M_{i,j,T}(\mathbf{U}_T) \xrightarrow{p} -\nu s_{i-j}(\boldsymbol{\phi}_0), \qquad 1 \leq j \leq p$$

and

$$M_{i,j,T}(\mathbf{U}_T) \xrightarrow{p} \nu t_{i-j}(\boldsymbol{\theta}_0), \qquad p+1 \leq j \leq p+q$$

as $T \to \infty$. Hence (6.2.28) holds.

By steps similar to those in the proof of (6.2.28) and the fact that $J_1(1-u) = -J_1(u)$ we obtain

$$c_i T^{-1} \sum_{t=p+1+i}^{T} J_1(F(U_t^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T))) J_2^{*\prime}(U_{t-i}^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T))\Big[\frac{\partial U_{t-i}^{(\infty)}(\boldsymbol{\lambda})}{\partial \lambda_j}\Big]_{\boldsymbol{\lambda}=\tilde{\boldsymbol{\lambda}}_T}$$
$$\xrightarrow{p} 0, \qquad ,1 \leq j \leq p+q,$$

as $T \to \infty$.

Therefore, from (6.2.27) we have

$$D\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T)) \xrightarrow{p} -\nu X^m \qquad \text{as } T \to \infty. \tag{6.2.29}$$

Hence (6.2.25) follows from (6.2.26), (6.2.29) and the fact that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0)$ is bounded in probability.

Moreover from Lemma 3.2 of Bustos, Fraiman and Yohai (1984) we have $U_t^{(\infty)}(\boldsymbol{\lambda}_0) = U_t$ a.s. and hence (6.2.20) follows from (6.2.21), (6.2.22) and (6.2.25). ∎

REMARK 6.1. Proposition 6.2.5 is based on generalized RA estimators. Li (1988) in Lemma 2 shows a similar result for RA estimators, but the details of the proof and the assumptions under which the lemma holds are omitted.

PROPOSITION 6.2.6. Assume that assumptions A(i),B(ii), B(v), B(vi) and B(vii) hold. If, in addition, $\widehat{\boldsymbol{\lambda}}_T^m$ is a sequence of estimators satisfying

$$T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m), \widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty \tag{6.2.30}$$

and such that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0)$ is bounded in probability, then

$$T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) \xrightarrow{\mathcal{D}} N(\mathbf{0}, G), \qquad \text{as } T \to \infty \tag{6.2.31}$$

where $G = \xi(I_m - X^m[(X^m)'X^m]^{-1}(X^m)')$, ξ is defined by (2.6) and X^m is defined by (6.2.7).

PROOF: By the Mean Value Theorem, we have

$$\begin{aligned} &T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\widehat{\boldsymbol{\lambda}}_T^m), \widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m) = \\ &T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0), \boldsymbol{\phi}_0, \boldsymbol{\theta}_0) + D\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T), \tilde{\boldsymbol{\phi}}_T, \tilde{\boldsymbol{\theta}}_T)T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0), \end{aligned} \tag{6.2.32}$$

where $\tilde{\boldsymbol{\lambda}}_T$ satisfies $\|\tilde{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_0\|_2 \leq \|\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0\|_2$.

As in Lemma 3.5 of Bustos, Fraiman and Yohai (1984), and since $T^{1/2}(\tilde{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_0)$ is bounded in probability we can prove that

$$D\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\tilde{\boldsymbol{\lambda}}_T), \tilde{\boldsymbol{\phi}}_T, \tilde{\boldsymbol{\theta}}_T) \xrightarrow{a.s.} \nu(X^m)'X^m \qquad \text{as } T \to \infty. \tag{6.2.33}$$

Therefore from (6.2.32), (6.2.33) and the fact that $\widehat{\boldsymbol{\lambda}}_T^m$ is a sequence of estimators satisfying (6.2.30) we have

$$T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) = -\nu^{-1}((X^m)'X^m)^{-1}T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0), \boldsymbol{\phi}_0, \boldsymbol{\theta}_0) + o_p(1).$$

It is easy to show that

$$T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0), \boldsymbol{\phi}_0, \boldsymbol{\theta}_0) = (X^m)'T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0)) + o_p(1).$$

Hence

$$T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) = -\nu^{-1}((X^m)'X^m)^{-1}(X^m)'T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T^{(\infty)}(\boldsymbol{\lambda}_0)) + o_p(1).$$

Moreover from Lemma 3.2 of Bustos, Fraiman and Yohai (1984) we have $U_t^{(\infty)}(\boldsymbol{\lambda}_0) = U_t$ a.s. and therefore

$$T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0) = -\nu^{-1}((X^m)'X^m)^{-1}T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T) + o_p(1). \tag{6.2.34}$$

Then from Proposition 6.2.5 and (6.2.34) we obtain

$$T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) = (I_m - X^m[(X^m)'X^m]^{-1}(X^m)')T^{1/2}\tilde{\boldsymbol{\gamma}}_4^m(\mathbf{U}_T) + o_p(1). \tag{6.2.35}$$

Therefore from (6.2.17) we have

$$T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) = (I_m - X^m[(X^m)'X^m]^{-1}(X^m)')T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T) + o_p(1).$$

Hence, using Proposition 6.2.4 it follows (6.2.31). ∎

In Proposition 6.2.7 we will use the Frobenius matrix norm given by

$$\|B\|_F = \left[\sum_{i=1}^{r}\sum_{j=1}^{s}|b_{ij}|^2\right]^{1/2}$$

where $B = (b_{ij})$ is a $r \times s$ matrix.

PROPOSITION 6.2.7. Assume that assumptions A(i),B(ii), B(v), B(vi) and B(vii) hold. If, in addition, $\widehat{\boldsymbol{\lambda}}_T^m$ is a sequence of estimators satisfying

$$T^{1/2}\mathbf{W}_T^{*,m}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m), \widehat{\boldsymbol{\phi}}_T^m, \widehat{\boldsymbol{\theta}}_T^m) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty$$

and such that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^m - \boldsymbol{\lambda}_0)$ is bounded in probability, then

(i) $Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m))$ is asymptotically distributed as chi-squared with degrees of freedom $m-p-q$.

(ii) If $\widehat{\boldsymbol{\lambda}}_T^*$ is a sequence of estimators satisfying

$$T^{1/2}\mathbf{W}_T^{*}(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*), \widehat{\boldsymbol{\phi}}_T^*, \widehat{\boldsymbol{\theta}}_T^*) \xrightarrow{p} \mathbf{0}, \qquad \text{as } T \to \infty$$

and such that $T^{1/2}(\widehat{\boldsymbol{\lambda}}_T^* - \boldsymbol{\lambda}_0)$ is bounded in probability. Then for every $\epsilon > 0$ and $\delta > 0$ there exist $m_0 > 0$ and $T_0 > 0$ such that for $m \geq m_0$ and $T \geq T_0$

$$P(|Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))| \geq \epsilon) \leq \delta.$$

PROOF: Since $(\xi^{-1}T)^{1/2}\tilde{\boldsymbol{\rho}}_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m))$ has an asymptotic covariance matrix that is idempotent of rank $m-p-q$ we obtain (i) from Proposition 6.2.6.

Now we will prove (ii). From the Cauchy-Schwartz inequality we have

$$
\begin{aligned}
|Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))| \leq{}& \xi^{-1}T\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))\|_2^2 \\
&+ 2\xi^{-1}T\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))\|_2\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m))\|_2.
\end{aligned} \tag{6.2.36}
$$

We will show that for every $\epsilon > 0$ and $\delta > 0$ there exist $m_1 > 0$ and $T_1 > 0$ such that for $m \geq m_1$ and $T \geq T_1$

$$
P(\xi^{-1}T\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))\|_2^2 \geq \epsilon) \leq \delta. \tag{6.2.37}
$$

From (6.2.17) and (6.2.35) we obtain

$$
\begin{aligned}
T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) ={}&[I_m - X^m((X^m)'X^m)^{-1}(X^m)'] \\
&\times T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T) + o_p(1).
\end{aligned} \tag{6.2.38}
$$

Let C be the $(p+q) \times (p+q)$ symmetric matrix given by

$$
\begin{cases}
C_{i,j} = \sum_{k=0}^{\infty} s_k(\boldsymbol{\phi}_0)s_{k+j-i}(\boldsymbol{\theta}_0), & i \leq j \leq p \\
C_{i,p+j} = -\sum_{k=0}^{\infty} t_k(\boldsymbol{\theta}_0)s_{k+j-i}(\boldsymbol{\phi}_0), & i \leq p,\ j \leq q,\ i \leq j \\
C_{i,p+j} = -\sum_{k=0}^{\infty} s_k(\boldsymbol{\phi}_0)t_{k+j-i}(\boldsymbol{\theta}_0), & i \leq p,\ j \leq q,\ j \leq i \\
C_{p+i,p+j} = \sum_{k=0}^{\infty} t_k(\boldsymbol{\theta}_0)t_{k+j-i}(\boldsymbol{\phi}_0), & i \leq j \leq q
\end{cases}
$$

By steps similar to those in the proof of (6.2.35) and from (6.2.17) we have

$$
\begin{aligned}
T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*)) ={}&T^{1/2}\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T) \\
&- X^m C^{-1}(X^{T-p-1})'T^{1/2}\tilde{\boldsymbol{\rho}}_4^{T-p-1}(\mathbf{U}_T) + o_p(1).
\end{aligned} \tag{6.2.39}
$$

where X^{T-p-1} is obtained replacing m by $T-p-1$ in (6.2.19). Then, from (6.2.38) and (6.2.39) it follows

$$
\begin{aligned}
&T\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^m)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T^*))\|_2^2 \\
&\leq T\|X^m C^{-1}(X^{T-p-1})'\tilde{\boldsymbol{\rho}}_4^{T-p-1}(\mathbf{U}_T) - X^m[(X^m)'X^m]^{-1}(X^m)'\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T)\|_2^2 + o_p(1).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&T\|\tilde{\rho}_4^m(\mathbf{U}_T(\widehat{\lambda}_T^m)) - \tilde{\rho}_4^m(\mathbf{U}_T(\widehat{\lambda}_T^*))\|_2^2 \\
&\qquad \leq T\|X^m\|_F^2\|C^{-1}\|_F^2\|(X^{T-p-1})'\tilde{\rho}_4^{T-p-1}(\mathbf{U}_T) - (X^m)'\tilde{\rho}_4^m(\mathbf{U}_T)\|_2^2 \\
&\qquad + T\|X^m\|_F^2\|C^{-1} - [(X^m)'X^m]^{-1}\|_F^2\|(X^m)'\tilde{\rho}_4^m(\mathbf{U}_T)\|_2^2 + o_p(1).
\end{aligned} \tag{6.2.40}
$$

From the Chebyschev inequality and the fact that $\sum_{j=0}^{\infty}|s_j(\boldsymbol{\phi}_0)| < \infty$ and $\sum_{j=0}^{\infty}|t_j(\boldsymbol{\theta}_0)| < \infty$ it follows that for every $\epsilon > 0$ and $\delta > 0$ there exist $m_2 > 0$ and $T_2 > 0$ such that for $m_2 \leq m \leq T-p-2$ and $T \geq T_2$

$$
P(T\|(X^{T-p-1})'\tilde{\rho}_4^{T-p-1}(\mathbf{U}_T) - (X^m)'\tilde{\rho}_4^m(\mathbf{U}_T)\|_2^2 \geq \epsilon) \leq \delta. \tag{6.2.41}
$$

Moreover,

$$
\begin{aligned}
&\|C^{-1} - [(X^m)'X^m]^{-1}\|_F^2 \\
&\leq \|[(X^m)'X^m]^{-1}\|_F^2\|(X^m)'X^m - C\|_F^2\|C^{-1}\|_F^2.
\end{aligned} \tag{6.2.42}
$$

Then from the fact that $\sum_{j=0}^{\infty}|s_j(\boldsymbol{\phi}_0)| < \infty$, $\sum_{j=0}^{\infty}|s_j(\boldsymbol{\phi}_0)|^2 < \infty$, $\sum_{j=0}^{\infty}|t_j(\boldsymbol{\theta}_0)| < \infty$ and $\sum_{j=0}^{\infty}|t_j(\boldsymbol{\theta}_0)|^2 < \infty$ we obtain, for every T

$$
\lim_{m\to\infty}\|(X^m)'X^m - C\|_F^2 = 0. \tag{6.2.43}
$$

Hence from (6.2.41), (6.2.42) and (6.2.43) it follows (6.2.37). Then (ii) is an immediate consequence of (6.2.36), (6.2.37) and (i). ∎

PROOF OF THEOREM 4.1: From Proposition 6.2.3 (i) of Ferretti, Kelmansky and Yohai (1991) we obtain

$$
T^{1/2}\mathbf{W}_T^*(\mathbf{U}_T(\widehat{\lambda}_T), \widehat{\boldsymbol{\phi}}_T, \widehat{\boldsymbol{\theta}}_T) \xrightarrow{p} 0, \quad \text{as } T \to \infty. \tag{6.2.44}
$$

Then, from Proposition 6.2.7 (ii), it immediately follows that for every $\epsilon > 0$ and $\delta > 0$ there exist $m_0 > 0$ and $T_0 > 0$ such that for $m \geq m_0$ and $T \geq T_0$

$$
P(|Q_4(\mathbf{U}_T(\widehat{\lambda}_T^m)) - Q_4(\mathbf{U}_T(\widehat{\lambda}_T))| \geq \epsilon) \leq \delta. \tag{6.2.45}
$$

Also, from the Cauchy-Schwartz inequality we have

$$
\begin{aligned}
|Q_3(\mathbf{R}_T(\widehat{\lambda}_T)) - Q_4(\mathbf{U}_T(\widehat{\lambda}_T))| &\leq \xi^{-1}T\|\tilde{\rho}_3^m(\mathbf{R}_T(\widehat{\lambda}_T)) - \tilde{\rho}_4^m(\mathbf{U}_T(\widehat{\lambda}_T))\|_2^2 \\
&+ 2\xi^{-1}T\|\tilde{\rho}_3^m(\mathbf{R}_T(\widehat{\lambda}_T)) - \tilde{\rho}_4^m(\mathbf{U}_T(\widehat{\lambda}_T))\|_2\|\tilde{\rho}_4^m(\mathbf{U}_T(\widehat{\lambda}_T))\|_2.
\end{aligned} \tag{6.2.46}
$$

From Proposition 6.2.3 it follows that

$$T\|\tilde{\boldsymbol{\rho}}_3^m(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T))\|_2 \xrightarrow{p} 0 \quad \text{as } T \to \infty. \tag{6.2.47}$$

Then from (6.2.44) and (6.2.47) we have for every $\epsilon > 0$ and $\delta > 0$ there exist $m_1 > 0$ and $T_1 > 0$ such that for $m \geq m_1$ and $T \geq T_1$

$$P(2\xi^{-1}T\|\tilde{\boldsymbol{\rho}}_3^m(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)) - \tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T))\|_2\|\tilde{\boldsymbol{\rho}}_4^m(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T))\|_2 \geq \epsilon) \leq \delta. \tag{6.2.48}$$

Hence from (6.2.46), (6.2.47) and (6.2.48) we obtain that for every $\epsilon > 0$ and $\delta > 0$ there exist $m_2 > 0$ and $T_2 > 0$ such that for $m \geq m_2$ and $T \geq T_2$

$$P(|Q_3(\mathbf{R}_T(\widehat{\boldsymbol{\lambda}}_T)) - Q_4(\mathbf{U}_T(\widehat{\boldsymbol{\lambda}}_T))| \geq \epsilon) \leq \delta. \tag{6.2.49}$$

Therefore, from Proposition 6.2.7 (i), (6.2.45) and (6.2.49) Theorem 4.1 follows. ∎

REFERENCES

Anderson, T.W. (1971). *The Statistical Analysis of Times Series.* Wiley, New York.

Anderson, T.W. and Walker, A.M. (1964). On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *Ann. Math. Statis.* **35** 1296-303.

Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *J. Amer. Statist. Assoc.* **65** 1509-26.

Bustos, O., Fraiman, R. and Yohai, V. (1984). Asymptotic behaviour of the estimates based on residual autocovariance for ARMA models. In *Robust and Nonlinear Time Series.* (J. Franke, W. Härdle and D. Martin, eds.) 26-49. Springer-Verlag, New York.

Bustos, O. and Yohai, V. (1986). Robust Estimates for ARMA models. *J. Amer. Statist. Assoc.* **81** 155-168.

Ferretti, N.E., Kelmansky D.M. and Yohai V.J. . Estimators based on ranks for ARMA models. *Communications in Statistics-Theory and Methods* Vol 20, **12**, 1991 .

Koul, H. (1990). A weak convergence result useful in robust autoregression. To appear in *Journal of Statistical Planning and Linear Inference.*

Li, W.K. (1988). A goodness of fit test in robust time series modelling. ***Biometrika***. **75**, 2, 355-61.

Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. ***Biometrika*** **65** 297-303.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press.

Puri, M.L. and Sen, P. K. (1971). *Nonparametric Methods on Multivariate Analysis*. Wiley, New York.