GLOBAL JOURNAL OF MATHEMATICAL SCIENCES VOL 9, NO. 1, 2010: 45-51
COPYRIGHT© BACHUDO SCIENCE CO. LTD PRINTED IN NIGERIA. ISSN 1596-6208
www.globaljournalseries.com; Email: info@globaljournalseries.com

A MODEL FOR NONLINEAR INNOVATION IN TIME SERIES

D. W. U. EBONG

(Received 01, April 2010; Revision Accepted 4, May 2010)

ABSTRACT

This paper introduces a class of nonlinear innovation process that has similar properties as the white noise process. Consequently the process can be a replacement of the white noise process in cases where the latter is inadequate as residual process.

KEYWORDS: Asymptotic distribution of autocorrelation, nonlinear errors, nonlinear residuals, nonlinear time series

1. INTRODUCTION

There has been growing interest in non-linear time series of recent mainly as a result of the observation that many practical time series possess nonlinear components [Hinich & Patterson (1985), Granger & Andersen (1978), Baillie & Bollerslev (1989), Ding, Granger & Engle (1993), Tong (1990)]. Nonlinear AR models with conditionally heteroscedastic errors are common in financial and econometric time series. The conditional variance may be specified as nonlinear autoregressive conditional heteroscedasticity (ARCH) model [see e.g. Chen & Chen (2001)], or as linear generalized ARCH (GARCH) model of Bollerslev (1986), or even a GARCH model with complicated nonlinear structure. Models with threshold type nonlinearities characterized by discontinuous functions are also in use [see Tong (1990) Chen &Tsay (1993) for models for conditional mean; Glosten, Jaganathan & Runkle (1993), Rabemananjara & Zakoïan (1993) for GARCH models]. The functional AR model of Chen & Tsay(1993) encompasses various well-known nonlinear autoregressive models such as the smooth transition AR models [see Teräsvirta (1994), van Dijk, Terasvirta & Fances (2002)]. For nonlinear model involving moving average see, for example, Brännes et al (1998). For tests for nonlinearity under various assumptions on the type of nonlinearity see, for example, Tsay(1986), Lee(1991), Pera & Rodriguez (2005), Hinich, Mendes & Stone (2005), Subba Rao & Gabr (1980), Aparicio (1998), Barnett & Wolf (2005).
The current paper wishes to introduce a class of nonlinear processes that can be used to model certain nonlinearities in residual process. The nonlinear process will be introduced in the next section. Thereafter we will explore some properties of the process that are similar to those of the white noise process. The final section will give a generalization of the nonlinear process.

2. THE PROCESS

Let $\{\varepsilon_t\}$ be zero mean white noise (independent and identically distributed) process with finite fourth moment. In particular let $E(\varepsilon_t{}^2) = \sigma^2$, $E(\varepsilon_t^3) = \lambda\sigma^3$ and $E(\varepsilon_t^4) = \eta\sigma^4$.

Consider the process $\{W_t\}$ where

$$W_t = \varepsilon_t + \varepsilon_{t-1}\sum_{i=1}^{s}\beta_i\varepsilon_{t-i-1} \tag{1}$$

Clearly $E(W_t) = 0, \ \mathrm{var}(W_t) \equiv \gamma_0(W) = \sigma^2 + \sigma^4 \sum_{i=1}^{s}\beta_i^2$ (2)

and $\gamma_k(W) \equiv Cov(W_t, W_{t+k}) = E(W_tW_{t+k}) = 0$ for $k \neq 0$. Thus $\{W_t\}$ is uncorrelated $(0, \gamma_0(W))$ process just like the white noise (WN) process $\{\varepsilon_t\}$. However, unlike the latter, $\{W_t\}$ is not independent sequence. Nonetheless, it can serve as a non-linear innovation process in many modeling situations where the WN process is currently used, for example in a linear process, say ARMA process. To see this we first consider MA(q) process:

$$X_t = \varepsilon_t + \sum_{i=1}^{q}\theta_i\varepsilon_{t-i}$$

D. W. U. Ebong, Department of Mathematics and Statistics, University of Port, Harcourt, Port Harcourt, Nigeria.

For this process $E(X_t)=0,\ \gamma_k(X)\ =\ 0$ for all |*k*| > q and

$$\gamma_{-k}\ =\ \gamma_k(X)\ =\ \begin{cases}\sigma^2\ +\sigma^2\sum\theta_i^2 & \text{for}\ \ k\ =0\\ \sigma^2\ (\theta_k\ +\sum\theta_i\theta_{i+k}) & \text{for}\ \ k\ =1,2,...,\ \ q\end{cases}$$

See for example Fuller (1976) p. 20. Now consider the processes

$$Y_t\ =\ W_t\ +\ \sum_{i=1}^{q}\theta_i\varepsilon_{t-i} \tag{3}$$

and
$$Z_t\ =\ W_t\ +\ \sum_{i=1}^{q}\theta_i W_{t-i} \tag{4}$$

It is easily seen (see e.g. Fuller(1976) p. 20) that $E\left(Y_t\right)=0,\ \gamma_k(Y)=0$ for all |*k*| > q and

$$\gamma_{-k}(Y)=\gamma_k(Y)\ =\ \begin{cases}\gamma_0(W)+\sigma^2\sum\theta_i^2 & \text{for}\ k\ =\ 0\\ \sigma^2\left(\theta_k\ +\ \sum\theta_i\theta_{i+k}\right) & \text{for}\ k=1,2,...\ q\end{cases}$$

Also $E(Z_t)=0,\ \ \gamma_k(Z)=\ 0$ for all |*k*| > q and

$$\gamma_{-k}(Z)\ =\ \gamma_k(Z)\ =\begin{cases}\gamma_0(W)(1+\sum\theta_i^2) & \text{for}\ \ k\ =\ 0\\ \gamma_0(W)(\theta_k\ +\ \sum\theta_i\ \ \theta_{i+k}) & \text{for}\ \ k\ =1,2,..\ q\end{cases}$$

Consequently we can consider MA process in the non-linear process $\{W_t\}$ as an alternative to the one in WN process and still preserve the autocorrelation function structure. The process (3) seems to be of more practical importance in view of Volterra expansion (see Volterra, 1959) for nonlinear stationary series. However, notice that

process (4) is $$Z_t=\varepsilon_t+\varepsilon_{t-1}\sum_{i=1}^{q}\beta_i\varepsilon_{t-i-1}+\sum_{j=1}^{q}\theta_j\varepsilon_{t-j}+\sum_{j=1}^{q}\theta_j\varepsilon_{t-j-1}\sum_{i=1}^{s}\beta_i\varepsilon_{t-i-j-1}$$

$$=\sum_{j=0}^{q}\theta_j\varepsilon_{t-j}\ +\ \sum_{j=0}^{q}\theta_j\varepsilon_{t-j-1}\sum_{i=0}^{s}\beta_i\varepsilon_{t-i-j-1}$$

which is an extension of process (3) as MA(q) plus nonlinear terms.

If the WN process in a stationary AR (p) process is replaced by $\{W_t\}$ we have the process $Y_t\ =\ \sum_{i=1}^{p}\varphi_i\ Y_{t-i}\ +\ W_t$. We easily see (e.g. Fuller (1976) §2.3) that $E\left(Y_t\right)\ =\ 0,\ \gamma_k(Y)\ =\ \sum_{j=1}^{p}\varphi_j\gamma_{k-j}(Y),\ \ k\ >0$ and $\gamma_0(Y)\ =\sum_{j=1}^{p}\varphi_j\gamma_j(Y)+\gamma_0(W)$. This is again similar to what is obtained when WN process is used instead of $\{W_t\}$. A similar result holds for ARMA process generally. Terms similar to the right hand side of (1) can be included in ARCH and GARCH models to take care of certain nonlinearities. Consequently the $\{W_t\}$ process (1) can be used as innovation in situations where the WN process is inappropriately used to account for non-linear innovation. This leads to a new class of, for example, ARMA processes in non-linear innovation process $\{W_t\}$. The rest of the paper will be devoted to more properties of the process $\{W_t\}$ that are similar to those of the WN process.

3. SOME PROPERTIES OF W_T PROCESS

We have seen that $\{W_t\}$ is uncorrelated $\left(0,\gamma_0(W)\right)$ non-linear process. We will prove the following:

Theorem 1: $$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}W_t\ \xrightarrow{D}\ N(0,\gamma_0(W))\ as\ n\to\infty$$

This theorem shows that, just like WN process, the sample mean $\overline{W}$ process based on sample size n is asymptotically zero mean normal with variance $\gamma_0(W)/n$. Let the population correlation coefficient at lag k, ρ_k(W), be estimated by $\gamma_k(W) \equiv \gamma_{k,n}(W) = c_{k,n}(W)/c_{0,n}(W)$ where $c_k(W) = c_{k,n}(W) = \frac{1}{n-k}\sum_{t=1} W_t W_{t+k}$ is the sample autocovariance function at lag k based on sample size n. We also have the following results.

Theorem 2: $\sqrt{n}\,\gamma_{k,n}(W) \xrightarrow{D} N(0,1)$ *as* $n \to \infty$ *whenever* $k > s+1$

Theorem 3: $\sqrt{n}\,\gamma_{k,n}(W) \xrightarrow{D} N(0,\nu)$ *as* $n \to \infty$ *whenever* $1 \le k \le s+1$ *where* $\nu > 1$.

Theorem 2 indicates that for all lags k greater than s+1, the usual ±2/√n bounds for the (approximate) 95% confidence interval of WN $\rho_k(\varepsilon)$ ($k \neq 0$) also holds for $\{W_t\}$ process. Theorem 3 indicates that for positive lags not greater than s + 1, the above ±2/√n bounds are too conservative to be bounds for $\rho_k(W)$. These results incidentally also provide information on the parameter s of the process $\{W_t\}$ when it has to be estimated from data. It particular, if theorem 2 holds for all $k > 0$ we would be sure that the WN process should be used instead of the process $\{W_t\}$.

4. Proofs

The proofs will make use of central limit theorem for m-dependent random sequences given below in lemma 1. The result is due to Hoeffding and Robbins (1948) and is proved in Fuller (1976) p. 246. For the avoidance of doubt, a sequence of random variables $\{Z_t : t = 0, \pm 1, \pm 2, ..\}$ is said to be *m-dependent* (where m is a non-negative integer) if $\{\ldots, Z_{u-2}, Z_{u-1}, Z_u\}$ and $\{Z_\upsilon, Z_{\upsilon+1}, \ldots\}$ are independent sets of random variables whenever $\upsilon - u > m$.

Lemma 1: Let $\{Z_t\}$ be m-dependent sequence with E$\{Z_t\}$ = 0, var$\{Z_t\}$ = $\sigma_t^2 < \infty$ and $E(|Z_t|^3) \le \beta^3$. Let the limit $\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} A_{t+j} = A$ ($\neq 0$) be uniform for t = 1, 2,... where $A_t = E(Z_{t+m}^2) + 2\sum_{j=1}^{m} E(Z_{t+m-j} Z_{t+m})$. Then $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Z_t \xrightarrow{D} N(0, A)$ as $n \to \infty$.

The next lemma is proved, for example, in Rao (1973) p. 388.

Lemma 2: Let $\{V_n\}$ be a sequence of p-dimensional random vectors such that $\sqrt{n}\,(V_n - \mu) \xrightarrow{D} N_p(0,V)$ *as* $n \to \infty$. If $g : R^P \to R^q$ is a differentiable function in the neighbourhood of μ and J is q x p Jacobian matrix at μ, then $\sqrt{n}\,(g(V_n) \to g(\mu)) \xrightarrow{D} N_q(0, JVJ')$ *as* $n \to \infty$.

Proof of Theorem 1

W_t is a function of ε_{t-j} for j = 0, 1, …, s+1. Consequently $\{W_t\}$ is (s + 1) –dependent sequence. Since $E(W_{t+m-j}\, W_{t+m}) = \gamma_0(w)\delta_0^j$, the A_t in lemma 1 is $\gamma_0(W)$ and so A = $\gamma_0(W)$ and we have the result of the theorem.

Proof of Theorem 2

Let $k > 0$. Notice that $(n-k)c_{k,n}(W) = \sum_{t=1}^{n-k} W_t W_{t+k}$ and $\{W_t\, W_{t+h}\}$ *is* $(s+k+1)$-dependent sequence. By lemma 1,

$\sqrt{n-k}\; c_{k,n}(W) = \frac{1}{\sqrt{n-k}} \sum_{t=1}^{n-k} W_t W_{t+h} \xrightarrow{D} N(0,A)$ *as* $n \to \infty$ where *A* is as defined in lemma 1 with

$A_t = E(W_{t+m}^2 W_{t+k+m}^2) + 2\sum_{u=1}^{m} E\left(W_{t+m-u} W_{t+m+k-u} W_{t+m} W_{t+m+k}\right)$ and $m = s+k+1$. To compute A_t it suffices to compute

$$E\left(W_t^2 W_{t+k}^2\right) + 2\sum_{u=1}^{m} E\left(W_{t-u} W_{t+k-u} W_t W_{t+k}\right) \tag{5}$$

$$\begin{aligned} E(W_t^2 W_{t+k}^2) = \text{var}\left(W_t\, W_{t+k}\right) &= \text{var}\,(\varepsilon_{t+k} W_t + W_t \varepsilon_{t+k-1} \sum \beta_j\ \varepsilon_{t+k-j-1}) \\ &= \text{var}\left(\varepsilon_{t+k} W_t\right) + \text{var}\,(W_t \varepsilon_{t+k-1} \sum \beta_j \varepsilon_{t+k-j-1} \end{aligned}$$

since $Cov(\varepsilon_{t+k} W_t, W_t \varepsilon_{t+k-1} \sum \beta_j \varepsilon_{t+k-j-1}) = \sum \beta_j E(\varepsilon_{t+k} W_t^2\ \varepsilon_{t+k-1} \varepsilon_{t+k-j-1})$

$$= E(\varepsilon_{t+k}) \sum \beta_j E(W_t^2 \varepsilon_{t+k-1}\ \varepsilon_{t+k-j-1}) = 0$$

Now $\text{var}(\varepsilon_{t+k} W_t) = E(\varepsilon_{t+k}^2 W_t^2) = E(\varepsilon_{t+k}^2) E(W_t^2) = \sigma^2 \gamma_0(W) = \sigma^4 + \sigma^6 \sum_{i=1}^{s} \beta_i^2$ (6)

Also $\operatorname{var}(W_t \varepsilon_{t+k-1} \sum \beta_j \varepsilon_{t+k-j-1}) = E(W_t^2 \varepsilon_{t+k-1}^2 \sum_i \sum_j \beta_i \beta_j \, \varepsilon_{t+k-i-1} \varepsilon_{t+k-j-1})$ (7)

$$= E(\varepsilon_{t+k-1}^2) E(W_t^2 \sum_i \sum_j \beta_i \beta_j \varepsilon_{t+k-i-1} \, \varepsilon_{t+k-j-1}) \text{ for } k > s+1$$

$$= \sigma^2 E(W_t^2) \sum_i \sum_j \beta_i \beta_j \; E(\varepsilon_{t+k-i-1} \varepsilon_{t+k-j-1})$$

$$= \sigma^2 \gamma_0(W) \sum_i \beta_i^2 E(\varepsilon_{t+k-i-1}^2) = \sigma^4 \gamma_0(W) \Sigma \beta_i^2 = \gamma_0(W)(\gamma_0(W)) - \sigma^2)$$

Here we have used the fact that W_t and $\{\varepsilon_{t+k-i-1} \; : i = 1, 2, ..., s\}$ are uncorrelated if $k > s+1$.

Therefore, for $k > s+1$,

$$E\left(W_t^2 W_{t+k}^2\right) = \sigma^2 \gamma_0(W) + \gamma_0(W)(\gamma_0(W) - \sigma^2) = \gamma_0^2(W)$$

Next we compute the second term in (5).

$$E\left(W_{t-u} W_{t+k-u}, W_t W_{t+k}\right) = E\left(W_{t-u} W_{t+k-u} W_t \; \varepsilon_{t+k}\right) + \Sigma \beta_j E(W_{t-u} W_{t+k-u} W_t \varepsilon_{t+k-1} \varepsilon_{t+k-j-1}) \quad (8)$$

The first term on the right vanishes since ε_{t+k} is uncorrelated with $W_t W_{t+k-u} W_{t-u}$. If $k > 1$ and $u > 1$, the second term also vanishes since then ε_{t+k-1} and $W_t W_{t-u} W_{t+k-u} \varepsilon_{t+k-j-1}$ are uncorrelated. For $u = 1$ the second term vanishes whenever $k > s+1$ since then $W_t W_{t-1}$ and $\varepsilon_{t+k-1} \; W_{t+k-1} \varepsilon_{t+k-j-1}$ are uncorrelated.

Thus $E\left(W_{t-u} W_{t+k-u} W_t W_{t+k}\right) = 0 \text{ for } k > s+1$.

Consequently $A_t = \gamma_0^2(W)$. This also equals A since A_t is a constant. Hence

$\sqrt{n-k}\, c_{k,n}(W) \xrightarrow{D} N(0, \gamma_0^2(W))$ as $n \to \infty$ whenever $k > s+1$.

We now compute the asymptotic distribution of $c_{0,n}(W)$. Let $V_t = W_t^2 - \gamma_0(W)$. Then $\{V_t\}$ is (s+1)-dependent sequence. By lemma 1, $\sqrt{n}\,(c_{0,n}(W) - \gamma_0(W)) \equiv \dfrac{1}{\sqrt{n}} \sum_{y=1}^{m} V_t \xrightarrow{D} N\,(0, A)$ as $n \to \infty$ where A is as defined in lemma 1 with

$$A_t = E(V_{t+m}^2) + 2 \sum_{u=1}^{m} E(V_{t+m-u} V_{t+m}) \text{ and } m = s+1$$

Now $E(V_{t+m}^2) = \gamma_0(W^2)$ and $E(V_{t+m-j} V_{t+m}) = Cov(W_{t+m-u}^2, W_{t+m}^2)$. To compute the latter it suffices to compute $Cov(W_{t-u}^2, W_t^2)$. Straight forward computation using the properties of WN process gives

$$Cov(W_{t-j}^2, W_t^2) = Cov(W_{t-u}^2, W_t \varepsilon_t + W_t \varepsilon_{t-1} \Sigma \beta_i \varepsilon_{t-i-1})$$

$$= \Sigma \beta_i \; Cov(W_{t-u}^2, W_t \; \varepsilon_{t-1} \varepsilon_{t-i-1})$$

$$= \sum_i \sum_j \beta_i \beta_j Cov(W_{t-u}^2, \varepsilon_{t-1}^2 \varepsilon_{t-i-1} \varepsilon_{t-j-1})$$

$$= \{(\eta - 1)\sigma^6 \Sigma \beta_i^2 + 4\lambda \sigma^7 \beta_i \Sigma \beta_i \beta_{i+1}\} \delta_1^u + (\eta - 1)\sigma^6 \beta_{u-1}^2$$

$$+ (\eta - 1)\sigma^8 \beta_u^2 \sum \beta_i^2 + (\eta - 3)\sigma^8 \Sigma \beta_i^2 \beta_{i+u}^2 + 2\sigma^8 (\sum \beta_i \beta_{i+u})^2$$

Since A_t is independent of t we see that the asymptotic variance $A \equiv A(c_0)$ is given by

$$A(c_0) = \gamma_0(W^2) + 2(\eta - 1)\gamma_0^2(W) - 2(\eta - 1)\sigma^4 + 8\lambda \sigma^7 \beta_1 \sum_{i-1}^{s-1} \beta_{i+1}$$

$$+ 2(\eta - 3)\sigma^8 \sum_{u=1}^{s-1} \sum_{i=1}^{s-u} \beta_i^2 \beta_{i+u}^2 + 2\sigma^8 \sum_{u=1}^{s-1} \sum_{i=1}^{s-u} (\beta_i \beta_{i+u})^2$$

Next we show that

$$Cov(c_{k,m,}(W), c_{0,k}(W)) = \frac{1}{n(n-k)} \sum_{u=1}^{n} \sum_{t=1}^{n-k} E(W_t W_{t+k} W_u^2) = 0 \text{ for } k > s+1$$

To see this, noting that $\{W_t\}$ is (s+1) – dependent, we have for all k > s + 1:

a) $u < t+k \Rightarrow E(W_t W_{t+k} W_u^2) = E(W_{t+k}) E(W_t W_u^2) = 0$

b) $u = t+k \Rightarrow E(W_t W_{t+k} W_u^2) = E(W_t W_{t+k}^3) = E(W) E(W_{t+k}^3) = 0$

c) $u > t+k \Rightarrow E(W_t W_{t+k} W_u^2) = E(W_t)\, E(W_{t+k} W_u) = 0$

That is, $Cov(c_{k,n}(W), c_{0,k}(W)) = 0$ for *k* > *s* + 1.

To round up the proof we use lemma 2 with the function g defined by $g(u,\upsilon) = \frac{u}{\upsilon},\ \upsilon \neq 0,$

$V_{n,k} = (c_{k,n}(W),\ c_{0,n}(W))',\ \mu = E\left(V_{n,k}\right) = \left(0,\ \gamma_0(W)\right)',\ g\left(\mu\right) = 0$

and $g\left(V_{n,k}\right) = \gamma_{k,n}(W)$. The Jacobian of the transformation at μ is $J = (\frac{\partial g}{\partial u}(\mu),\ \frac{\partial g}{\partial u}(\mu)) = (\gamma_0(W)^{-1}\ ,\ 0)$. From the above discussion it is evident that

$\sqrt{n}\ (V_{n,k} - \mu) \xrightarrow{D} N_2(0,\ diag\{\gamma_0^2(W),\ A(c_0)\})$ as $n \to \infty$ whenever *k* > *s* + 1

By lemma 2 we have

$\sqrt{n}(g(V_{n,k}) - g\left(\mu\right)) \xrightarrow{D} N_2(0,\ Jdiag\{\gamma_0^2(W),\ A(c_0)\}J')$ as $n \to \infty$ whenever *k* > *s* + 1

That is, $\sqrt{n}\ \gamma_{k,n}(W) \xrightarrow{D} N(0,1)$ as $n \to \infty$ for all *k* > *s* + 1. The proof of theorem 2 is now complete.

Proof of Theorem 3

Using the same transformation g and same variable $V_{n,k}$ as in the proof of theorem 2, we see that if $\sqrt{n}\ ((V_{n,k} - \mu) \xrightarrow{D} N_2(0,V)$ as $n \to \infty$ then by lemma 2,

$\sqrt{n}\ (g(V_{n,k}) - g(\mu)) \equiv \sqrt{n}\ \gamma_{k,n}(W) \xrightarrow{D} N(0,JVJ')$ as $n \to \infty$ where *JVJ* = $V_{11} / \gamma_0^2(W)$ and V_{11} is the asymptotic variance, $A(c_k)$, of $\sqrt{n-k}\ c_{k,n}(W)$ for 1 ≤ k ≤ *s* + 1. Consequently, to prove the theorem it suffices to compute $A(c_k)$ and show that it is greater than $\gamma_0^2(W)$, the asymptotic variance of $\sqrt{n-k}c_{k,n}(W)$ when k > s + 1. To compute A_t we proceed as in the proof of theorem 2 and use equations (5) to (8). Thus for $1 \le k \le s+1$, $E(W_t^2 W_{t+k}^2) = \ \sigma^2\gamma_0(W) + \sum_i\sum_j \beta_i\beta_j E(W_t^2 \varepsilon_{t+k-1}^2 \varepsilon_{t+k-i-1}\varepsilon_{t+k-j-1})$ and

$\sum_{u=1}^{m} E\left(W_{t-u}W_{t+k-u}W_t W_{t+k}\right) = \sum_{u=1}^{m}\sum_{j=1}^{s} \beta_j E\left(W_{t-u}W_{t+k-u}W_t \varepsilon_{t+k-1}\varepsilon_{t+k-j-1}\right).$

For *k* = 1, straight forward computation gives

$$E\left(W_t^2 W_{t+k}^2\right) = \sigma^2\gamma_0(W) + \sum_i\sum_j \beta_i\beta_j E\left(W_t^2\varepsilon_t^2\varepsilon_{t-i}\varepsilon_{t-j}\right)$$

$$= \gamma_0^2(W) + (\eta-1)\sigma^6\sum_{i=1}^{s}\beta_i^2 + (\eta-1)\sigma^8\beta_1^2\sum_{i=1}^{S}\beta_i^2 + (\eta-3)\sigma^8\sum_{i=1}^{s-1}\beta_i^2\beta_{i+1}^2$$

$$+\ 2\sigma^8(\sum_{i=1}^{s-1}\beta_i\beta_{i+1})^2 \ +\ 4\lambda\sigma^7\beta_1\sum_{i=1}^{s-1}\beta_i\beta_{i+1}$$

And $$E\left(W_{t-u}W_{t+k-u}W_t W_{t+k}\right) = \begin{cases} \lambda\sigma^5\beta_1 + 2\sigma^6\sum_{i=1}^{s-1}\beta_i\beta_{i+1} & \text{for } k=1,\ u=1 \\ \sigma^6\sum_{i=1}^{s-u}\beta_i\beta_{i+u} & \text{for } k=1,\ u>1 \end{cases}$$

For $1 < k \le s+1$,

$$E(W_t^2 W_{t+k}^2) = \sigma^2\gamma_0(W) + E(\varepsilon_{t+k-1}^2)\sum_i\sum_j \beta_i\beta_j E(W_t^2\varepsilon_{t+k-i-1}\varepsilon_{t+k-j-1})$$

$$= \gamma_0^2(W) + (\eta-1)\sigma^6\beta_{k-1}^2 + (\eta-1)\sigma^8\beta_k^2\sum_{i=1}^{s}\beta_i^2 + (\eta-3)\sigma^8\sum_{i=1}^{s-k}\beta_i^2\beta_{i+k}^2 + 2\sigma^8(\sum_{i=1}^{s-k}\beta_i\beta_{i+k})^2$$

$$E\left(W_{t-u}W_{t+k-u}W_t W_{t+k}\right) = \begin{cases} \sigma^6\sum_{i=1}^{s-k}\beta_i\beta_{i+k} & \text{for } 1<k<s, u=1 \\ 0 & \text{otherwise} \end{cases}$$

Consequently for $1 \le k \le s+k$ the asymptotic variance, $A(c_k)$, of $\sqrt{n-k}\ c_{k,n}(w)$ is given by

$$A\left(c_1\right) = \gamma_0^2(W) + (\eta-1)\sigma^6\sum_{i=1}^{s}\beta_i^2 + (\eta-1)\sigma^8\beta_1^2\Sigma\beta_i^2 + (\eta-3)\sigma^8\Sigma\beta_i^2\beta_{i+1}^2 + 2\sigma^8(\sum_{i=1}^{s-k}\beta_i\beta_{i+1})^2$$

$$+\ 4\sigma^6 \sum_{i=1}^{s-1} \beta_i \beta_{i+1} + 2\sigma^6 \sum_{u=2}^{s-1} \sum_{i=1}^{s-u} \beta_i \beta_{i+u} + 2\lambda\sigma^5 \beta_1 + 4\lambda\sigma^7 \beta_1 \sum_{i=1}^{s-1} \beta_i \beta_{i+1} \text{ for k = 1}$$

For 1 < $k \leq s$ + 1 we have

$$A(c_k) = \gamma_0^2(W) + (\eta - 1)\sigma^6 \beta_{k-1}^2 + (\eta - 1)\sigma^8 \beta_k^2 \sum_{i=1}^{s} \beta_i^2 + (\eta - 3)\sigma^8 \sum_{i=1}^{s-k} \beta_t^2 \beta_{i+k}^2$$

$$+ 2\sigma^8 (\sum_{i=1}^{s-k} \beta_i \beta_{i+k})^2 \ + 2\sigma^6 \sum_{i=1}^{s-k} \beta_i \beta_{i+k}$$

These values are evidently greater than $\gamma_0^2(W)$. The proof of theorem 3 is now complete.

Discussion

a) As a passing remark we notice that since

$$A(c_s) = \gamma_0^2(W) + (\eta - 1)\sigma^6 \beta_{s-1}^2 + (\eta - 1)\sigma^8 \beta_s^2 \sum_{i=1}^{s} \beta_i^2,$$

$A(c_{s+1}) = \gamma_0^2(W) + (\eta - 1)\sigma^6 \beta_s^2$ and $(\eta - 1)\sigma^6 \beta_s^2$ can be much smaller than $\gamma_0^2(W) = \sigma^4 + 2\sigma^6 \Sigma\beta_i^2 + \sigma^8 (\Sigma\beta_i^2)^2$. Consequently the asymptotic variance of $\sqrt{n}\, r_{s+1,n}$ may be quite close to 1 especially if *s* is large.

b) By allowing the parameters β_i in (1) to be functions of t, heteroscedasticity in $\{W_t\}$ can be introduced. In that case var (Y_t) in the MA model given by equation (3) will be heteroscedastic but the autocovariances at non-zero lags will not be affected. In comparison, in the MA model given by (4) both var (Z_t) and the autocovariances would be affected.

c) The non-linear process (1) is a member of a class of uncorrelated non-linear processes that have properties similar to the WN process. indeed for any integer a ≥ 0, the process $U_t \equiv U_{t,a} = \varepsilon_t + \varepsilon_{t-a} \sum_{i=1}^{s} \beta_i \varepsilon_{t-i-a}$ is uncorrelated $(0, \gamma_0(U))$ process where $\gamma_0(U)$ is as given in (2). Moreover $\{U_{t,a}\}$ is (*s* + *a*)-dependent sequence and theorems 1-3 given above (with *s* + 1 replaced by *s* + a) also hold for each member of this class of processes. This provides some flexibility in the choice of non-linear innovation process. The problem of identifying suitable values of the parameters *a* and *s* for a given problem will be taken up elsewhere.

REFERENCES

Aparicio, F. M., 1998. Testing linearity with information theoretic statistics and the bootstrap, In Nonlinear modeling of high frequency financial time series, C. Duns and B. Zoo, Eds. Chichester, UK: John Wiley, pp 51 – 67

Baillie, R. and T. Bollerslev., 1989. Common stochastic trends in a system of exchange rates. Journal of finance, 44, 167 – 181.

Barnett, A. G. and Wolff, R. C., 2005. A time series domain test for some types of nonlinearity. IEEE transactions on signal processing, 53 (1), 26 – 33.

Bollerslov, T., 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics, 31, 307 – 327.

Brännes, K.,Jan G. deGooijer and Timo Teräsvirta 1998. Testing linearity against nonlinear moving average models. Communications in Statistics, theory and methods, 27 (8), 2025 – 2035.

Chen, G. and Chen, M. 2001. On a class of nonlinear AR(p) models with nonlinear ARCH errors. Australian and New Zealand Journal of Statistics, 43, 445 – 454.

Chen, R. and Tsay, R. S., 1993. Functional-coefficient autoregressive models. Journal of the American Statistical Association, 88, 298 – 308.

Ding, Z., Granger, C. W. J. and Engle, R. F., 1993. A long memory property of stock market returns and a new model. Journal of empirical finance, 1, 83 – 106.

Fuller, W. A., 1976. Introduction to statistical time series. New York: John Wiley.

Glosten, L. W., Jaganathan R. and Runkle, D. E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of finance, 48, 1779 – 1801.

Granger, C. W. J. and Andersen, A. P., 1978. Introduction to bilinear time series models. Gottingen: Vandenhoeck and Ruprecht.

Hinich, M., Mendes, E. and Stone, L., 2005. Detecting nonlinearity in time series:surrogate and bootstrap approaches. Studies in nonlinear dynamics and econometrics, 9 (4), 1268 – 1268.

Hinich, M. J. and Patterson D. M., 1985. Evidence of nonlinearity in daily stock returns. J. Bus. Econ. Statist. 3, 69 – 77.

Hoeffding, W. and Robbins, H. 1948. The central limit theorem for dependent random variables. Duke Math. J. 15, 773 – 780.

Lee, J. H. H., 1991. A Lagrange multiplier test for GARCH models. Economics letters, 37 (3), 265 – 271.

Pera, D. and Rodriguez, J., 2005. Detecting nonlinearity in time series by model selection criteria. International journal of forecasting, 21 (4), 731 – 748.

Rabemananjara, R. and Zakoïan J. M., 1993. Threshold ARCH models and asymmetrics in volatility. Journal of applied econometrics, 8, 31 – 49.

Rao, C. R., 1973. Linear statistical inference and its applications, 2^{nd} edition. New york: John Wiley.

Subba Rao, T. and Gabr, M. M., 1980. A test for linearity of stationary time series. Journal of time series analysis, 1, 145 – 158.

Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. J. Amer. Statist. Assoc. 89, 208 – 218.

Tong, H., 1990. Nonlinear time series: a dynamical system approach. London: Oxford University Press.

Tsay, R. S., 1986. Nonlinearity tests for time series. Biometrika, 73 (2), 461 – 466.

Van Dijk, D., Teräsvirta T. and Frances P. H., 2002. Smooth transition autoregressive models – a survey of recent developments. Economic reviews, 21, 1 – 47.

Volterra, V., 1959. Theory of functionals and of integrals and integro-differential equations. New York: Dover.