The University of Chicago
Department of Statistics

TECHNICAL REPORT SERIES

Local asymptotic powers of nonparametric and semiparametric tests for fractional integration

Xiaofeng Shao and Wei Biao Wu

TECHNICAL REPORT NO. 562

January 23, 2006

The University of Chicago
Department of Statistics
5734 S. University Avenue
Chicago, IL 60637

Local asymptotic powers of nonparametric and semiparametric tests for fractional integration

By XIAOFENG SHAO AND WEI BIAO WU
January 23, 2006
Department of Statistics, University of Chicago,
5734 S. University Ave. Chicago, IL 60637, U.S.A.
shao,wbwu@galton.uchicago.edu

Summary

The paper concerns testing long memory for fractionally integrated nonlinear processes. We show that the exact local asymptotic power is of order $O[(\log n)^{-1}]$ for four popular nonparametric tests and is of order $O(m^{-1/2})$, where m is the bandwidth which is allowed to grow as fast as n^{κ}, $\kappa \in (0, 2/3)$, for the semiparametric Lagrange multiplier (LM) test proposed by Lobato and Robinson (1998). Our theory provides a theoretical justification for the empirical findings in finite sample simulations by Lobato and Robinson (1998) and Giraitis et al. (2003) that nonparametric tests have lower power than LM test in detecting long memory. Our result is applied to the daily indices data of CRSP (Center for Research in Security Prices).

Some key words: Fractional integration, KPSS test, Local Whittle estimation, Lagrange multiplier test, Long memory, R/S test.

1. Introduction

Consider the I(d) (fractionally integrated process of order $d \in \mathbb{R}$) model

$$(1 - B)^d (X_t - \mu) = u_t, \quad t \in \mathbb{Z}, \tag{1}$$

where B is the backward shift operator, μ is an unknown mean and $\{u_t\}_{t\in\mathbb{Z}}$ is a mean zero covariance stationary short-range dependent (or short memory) process. Loosely speaking, a stationary process is short-range dependent if its autocovariances are absolutely summable. The process $\{X_t\}$ is short memory if $d = 0$, long memory if $d \in (0, 1/2)$ and negative memory (antipersistence) if $d \in (-1/2, 0)$. A primary problem in studying such processes is to test the existence of long memory. We formulate it as the hypothesis testing problem: I(0) versus I(d), $d \in (0, 1/2)$, or more generally I(0) versus I(d), $d \in (-1/2, 0) \cup (0, 1/2)$.

In the literature various parametric, nonparametric and semiparametric tests have been proposed. Parametric tests include Lagrange multiplier (LM) test in the frequency domain [Robinson (1994)] and LM and Wald tests in the time domain [Tanaka (1999)]. Recently Nielsen (2004) incorporated likelihood ratio test and conducted a thorough study of parametric tests. All these tests assume certain parametric forms of the spectral density of $\{u_t\}$ and have a local power of order $O(n^{-1/2})$, where n is the sample size. Here we say that a test has a local power of order $O(a_n)$, ($a_n \to 0$ as $n \to \infty$) if it has nontrivial power against the local alternative $d = \delta a_n$, $\delta \neq 0$. To allow robustness to model misspecifications of the short memory component $\{u_t\}$, Lobato and Robinson (1998) proposed a frequency domain semiparametric LM test. As an important feature, their test does not impose parametric assumptions on the spectral density of $\{u_t\}$. It is expected that the local power of their test is of order $O(m^{-1/2})$, where m is the bandwidth which typically grows no faster than $n^{4/5}$. However Lobato and Robinson did not give theoretical justification of this order of local power. Another type of test is of nonparametric nature. Four popular nonparametric tests have been well studied both empirically and theoretically: modified R/S test [Lo (1991)], KPSS test [Kwiatkowski et al. (1992)], K/S test [Xiao (2001)] and V/S test [Giraitis et al. (2003)]. Wright (1999) showed that R/S and KPSS tests only have trivial powers against the local alternative $d = \delta n^{-1/2}$ in the sense that the asymptotic distribution under null and local alternatives are the same. This is not surprising since nonparametric tests are expected to be inferior to parametric ones in terms of the local power. On the other hand, however, for a given stretch of time series, the parametric form of the short memory component is often unknown. Both nonparametric and semiparametric tests are important alternatives if one wants to avoid possible model misspecifications of $\{u_t\}$ in conducting parametric tests.

In this paper, we show that all the four nonparametric tests mentioned above have a local power of order $O[(\log n)^{-1}]$, which refines and improves Wright's result. Further we prove that the local power of LM test is of order $O(m^{-1/2})$, where m can grow as fast as n^κ, $\kappa \in (0, 2/3)$. Therefore nonparametric tests are inferior to LM test so far as the local asymptotic power is concerned. Our theory confirms the findings in finite sample simulations by Lobato and Robinson (1998) and Giraitis et al. (2003) that semiparametric LM test is superior to nonparametric ones with respect to both size and power.

Now we introduce some notation. For a random variable ξ, write $\xi \in \mathcal{L}^p$ ($p > 0$) if $\|\xi\|_p := [\mathbb{E}(|\xi|^p)]^{1/p} < \infty$ and let $\|\cdot\| = \|\cdot\|_2$. For two sequences (a_n), (b_n), denote by $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$. The symbols "$\to_D$" and "$\to_{\mathbb{P}}$" stand for convergence in distribution and in probability, respectively. The symbols $O_{\mathbb{P}}(1)$ and $o_{\mathbb{P}}(1)$ signify being bounded in probability and convergence to zero in probability. Let $I\!\!B(\cdot)$ be the standard Brownian motion; let $\mathcal{D}[0,1]$ be the space of functions on $[0,1]$ which are right continuous and have left limits, endowed with the Skorohod topology (Billingsley, 1968). Denote weak convergence by "$\Rightarrow$". Let $N(\mu, \sigma^2)$ be a normal distribution with mean μ and variance σ^2.

The paper proceeds as follows. Section 2 introduces four nonparametric tests and states their asymptotic distributions under both null and local alternatives. Section 3 presents semiparametric LM test and its asymptotic distribution under local alternatives. In Section 4, we apply the LM test to the CRSP (Center for Research in Security Prices) equal-weighted daily indices data. Section 5 concludes and technical details are gathered in appendices A and B.

2. Nonparametric tests

Here we review four well-known nonparametric tests for long memory. Let $\bar{X}_n = n^{-1}\sum_{j=1}^n X_j$ be the sample mean and $\bar{S}_k = \sum_{j=1}^k (X_j - \bar{X}_n)$ the centered partial sum.

- Modified R/S statistic [Lo (1991)]:

$$Q_n = \frac{1}{w_{n,l}}\left\{\max_{1\le k\le n} \bar{S}_k - \min_{1\le k\le n} \bar{S}_k\right\},$$

  where $w_{n,l}^2$ is the long-run variance estimator. Following Lo (1991),

$$w_{n,l}^2 = \frac{1}{n}\sum_{j=1}^n (X_j - \bar{X}_n)^2 + 2\sum_{j=1}^l \left(1 - \frac{j}{l+1}\right)\hat{\gamma}_j, \tag{2}$$

  where $\hat{\gamma}_j = n^{-1}\sum_{i=1}^{n-j}(X_i - \bar{X}_n)(X_{i+j} - \bar{X}_n), 0 \le j < n$. Here the bandwidth satisfies $l = l(n) \to \infty$ and $l/n \to 0$ as $n \to \infty$.

- KPSS statistic [Kwiatkowski et al. (1992)]: $K_n = (w_{n,l}^2 n^2)^{-1}\sum_{k=1}^n \bar{S}_k^2$.

- V/S statistic [Giraitis et al. (2003)]: $V_n = (w_{n,l}^2 n^2)^{-1}\{\sum_{k=1}^n \bar{S}_k^2 - n^{-1}\left(\sum_{k=1}^n \bar{S}_k\right)^2\}$.

- K/S statistic [Xiao (2001)]: $G_n = (n^{1/2} w_{n,l})^{-1}\max_{1\le k\le n}|\bar{S}_k|$.

Among the above four tests, KPSS and K/S tests were originally proposed to test trend stationarity versus unit root nonstationarity. They have been used by Lee and Schmidt (1996) and Lima and Xiao (2004) respectively for tests of fractional integration. Proposition 1 below describes their asymptotic null distributions. The proof is omitted since it easily follows from the continuous mapping theorem. Let $\sigma^2 = 2\pi f_u(0)$, where $f_u(\cdot)$ is the spectral density function of $\{u_t\}$. Throughout the paper, we assume without loss of generality $\mu = 0$ and $0 < f_u(0) < \infty$.

Proposition **1**. *Let $d = 0$. Assume that*

$$w_{n,l}^2 \to_{\mathbb{P}} \sigma^2 \quad \text{and} \quad (n\sigma^2)^{-1/2}\sum_{j=1}^{\lfloor nt\rfloor} X_j \Rightarrow I\!B(t) \quad \text{in } \mathcal{D}(0,1). \tag{3}$$

Let $\tilde{I\!B}(t) = I\!B(t) - tI\!B(1)$ be the Brownian bridge. Then we have
(a). $n^{-1/2}Q_n \to_D \sup_{0\le t\le 1}\tilde{I\!B}(t) - \inf_{0\le t\le 1}\tilde{I\!B}(t)$;
(b). $K_n \to_D \int_0^1 \tilde{I\!B}(t)^2 dt$;
(c). $V_n \to_D \int_0^1 \tilde{I\!B}(t)^2 dt - (\int_0^1 \tilde{I\!B}(t)dt)^2$;
(d). $G_n \to_D \sup_{0\le t\le 1}|\tilde{I\!B}(t)|$.

Giraitis et al. (2003) provided sufficient conditions for (3) and derived limiting distributions for R/S, KPSS and V/S statistics under both short memory null hypothesis and long memory alternatives. All of these tests are consistent against both long memory and antipersistent alternatives; see Shao and Wu (2005a) for the treatment of R/S and KPSS tests. The consistency was obtained for fixed d, $d \in (-1/2, 0) \cup (0, 1/2)$. Since we are interested in the local power, we allow d to be dependent on the sample size n.

Let $d_n = c/\log(n)$, where c is a fixed constant. Define

$$(1-B)^{d_n}(X_t - \mu) = u_t, \quad t = 1, \cdots, n. \tag{4}$$

Strictly speaking, the series $\{X_t\}_{t=1}^n$ form a triangular array of type $\{X_{tn} : t = 1, \cdots, n; n = 1, 2, \cdots\}$. For the convenience of presentation, we shall use $\{X_t\}_{t=1}^n$ and no confusion will arise. Throughout the paper, we assume

$$u_t = F(\ldots, \varepsilon_{t-1}, \varepsilon_t), \tag{5}$$

where ε_t are independent and identically distributed (iid) random variables and F is a measurable function such that u_t is well-defined. Then $\{u_t\}$ is a stationary causal ergodic process. The class (5) is large. It includes a number of widely used nonlinear time series models such as bilinear models, threshold models, GARCH and ARMA-GARCH models; see Wu and Min (2005) and Shao and Wu (2005b) for more details.

As in Wiener (1958), Priestley (1988) and Wu (2005b), (5) can be interpreted as a physical system with $\mathcal{F}_t = (\ldots, \varepsilon_{t-1}, \varepsilon_t)$ being the input, F being a filter and u_t being the output. Let $\{\varepsilon'_t\}_{t\in\mathbb{Z}}$ be iid copy of $\{\varepsilon_t\}_{t\in\mathbb{Z}}$. For $\xi \in \mathcal{L}^1$ define projection operators $\mathcal{P}_k\xi = \mathbb{E}(\xi|\mathcal{F}_k) - \mathbb{E}(\xi|\mathcal{F}_{k-1})$. Let $g_k(\mathcal{F}_0) = \mathbb{E}(u_k|\mathcal{F}_0)$ and $w_q(k) := \|g_k(\mathcal{F}_0) - g_k(\mathcal{F}_0^*)\|_q$, where $\mathcal{F}_0^* = (\mathcal{F}_{-1}, \varepsilon'_0)$; let $\delta_q(k) := \|u_k - u'_k\|_q$, where $u'_k = F(\mathcal{F}_0^*, \varepsilon_1, \cdots, \varepsilon_k)$. In Wu (2005b), $w_q(\cdot)$ and $\delta_q(\cdot)$ are called *predictive dependence measure* and *physical dependence measure* respectively. Intuitively, $w_q(k)$ measures the contribution of ε_0 in predicting u_k, while $\delta_q(k)$ quantifies the dependence of u_k on ε_0 by measuring the distance between u_k and its coupled version u'_k. In many applications physical and predictive dependence measures are easy to work with since they are directly related to data-generating mechanisms. For more details see Wu (2005b).

Theorem **1**. *Let $\{X_t\}_{t=1}^n$ be generated from (4). Assume $u_t \in \mathcal{L}^q$, $q > 2$ and*

$$\sum_{k=0}^{\infty} w_q(k) < \infty. \tag{6}$$

Then $n^{-1/2}\sum_{j=1}^{\lfloor nt \rfloor} X_j \Rightarrow \sigma e^c \mathbb{B}(t)$ in $\mathcal{D}(0,1)$.

Remark **1**. Since $\|\mathcal{P}_0 u_k\|_q \le w_q(k) \le 2\|\mathcal{P}_0 u_k\|_q$ (Wu, 2005b), (6) is equivalent to

$$\sum_{k=0}^{\infty} \|\mathcal{P}_0 u_k\|_q < \infty. \tag{7}$$

In the special case of the linear process $u_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$, $\mathcal{P}_0 u_k = a_k \varepsilon_0$, so (7) holds if $\sum_{k=0}^{\infty} |a_k| < \infty$ and $\varepsilon_1 \in \mathcal{L}^q$. See Remark 3 for nonlinear time series.

Theorem **2**. *Let $\{X_t\}_{t=1}^n$ be generated from (4). Assume $u_t \in \mathcal{L}^q$, $q > 2$, (6) and $l = \lfloor n^\alpha \rfloor$, $\alpha \in (0,1)$. Then $w_{n,l}^2 \to_{\mathbb{P}} \sigma^2 e^{2c\alpha}$. Consequently, we have*
(a). $n^{-1/2} Q_n \to_D e^{c(1-\alpha)} \{\sup_{0\le t\le 1} \tilde{I\!B}(t) - \inf_{0\le t\le 1} \tilde{I\!B}(t)\}$;
(b). $K_n \to_D e^{2c(1-\alpha)} \int_0^1 \tilde{I\!B}(t)^2 dt$;
(c). $V_n \to_D e^{2c(1-\alpha)} \{\int_0^1 \tilde{I\!B}(t)^2 dt - (\int_0^1 \tilde{I\!B}(t) dt)^2\}$;
(d). $G_n \to_D e^{c(1-\alpha)} \sup_{0\le t\le 1} |\tilde{I\!B}(t)|$.

See Appendix A for the proofs of Theorem 1 and 2.

Remark **2**. The long-run variance estimator (2) is equivalent to the nonparametric spectral density estimator evaluated at zero frequency with Bartlett window. The lack of power of nonparametric tests is not due to the choice of the window or the estimator of σ^2 itself. Even if we know σ^2 and replace $w_{n,l}^2$ by σ^2 in these test statistics, by Theorem 1 the exact local power is still of order $O[(\log n)^{-1}]$. Clearly our methods are applicable to other test statistics which can be expressed as continuous functionals of partial sum processes.

Theorem 2 shows that nonparametric tests have a nontrivial asymptotic power against the local alternative of the form $d_n = c/\log n$, $c \neq 0$. Their asymptotic distributions under the local alternative $d_n = o[(\log n)^{-1}]$ are the same as the asymptotic null distributions [cf. Remark 4]. In other words, nonparametric tests have no power to detect the local alternative, which is as close as $o[(\log n)^{-1}]$ to zero integration. Interestingly, the asymptotic distributions under null and local alternatives only differ by a multiplicative factor.

3. Semiparametric LM test

Lobato and Robinson (1998) introduced a frequency domain LM test of I(0) in a multivariate context. The idea is closely related to the local Whittle estimation of long memory parameter [Künsch (1987) and Robinson (1995b)]. For a process $\{X_t\}_{t\in\mathbb{Z}}$, denote the periodogram by

$$I_X(\lambda) = |w_X(\lambda)|^2, \quad \text{where} \quad w_X(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n X_t e^{it\lambda}.$$

The local Whittle estimator $\hat{d}$ is obtained by minimizing the local objective function

$$R(d) = \log\left(m^{-1} \sum_{j=1}^m \lambda_j^{2d} I_X(\lambda_j)\right) - \frac{2d}{m} \sum_{j=1}^m \log \lambda_j,$$

where $\lambda_j = 2\pi j/n$ and $m = m(n)$ is the bandwidth satisfying $m^{-1} + m/n \to 0$ as $n \to \infty$. Let $R'(d) = \partial R(d)/\partial d$ and $R''(d) = \partial^2 R(d)/\partial^2 d$. The standard LM test statistics can be written as $mR'(0)R''(0)^{-1}R'(0)$. Define

$$t_m = \frac{-m^{-1/2} \sum_{j=1}^m v_j I_X(\lambda_j)}{m^{-1} \sum_{j=1}^m I_X(\lambda_j)}, \tag{8}$$

where $v_j = \log j - m^{-1}\sum_{j=1}^m \log j$. Theorem 1 of Lobato and Robinson (1998) asserts that the standard LM test statistics is equal to $t_m^2 + o_{\mathbb{P}}(1)$. In fact, Lobato and Robinson's LM test statistics is defined to be t_m^2 in the univariate case. Note that the local Whittle estimator $\hat{d}$ is asymptotically equivalent to $t_m/2$ if $d = 0$ [cf. Shao and Wu (2005b)]. Under the null hypothesis, $t_m \to_D N(0,1)$. We reject the null in favor of long memory alternative if t_m falls into the upper tails of $N(0,1)$, or in favor of antipersistent alternative if t_m falls into the lower tails of $N(0,1)$. Theorem 2 of Lobato and Robinson (1998) shows that LM test is consistent against the fractional alternatives $d \in (-1/2, 0) \cup (0, 1/2)$. They mentioned that the LM test is expected to have good power against local alternatives of the form $d_m = \delta m^{-1/2}$, where δ is a fixed constant. But no proof was given in their paper. To investigate the local power, we define

$$(1-B)^{d_m}(X_t - \mu) = u_t,\ t = 1, \cdots, n, \tag{9}$$

where u_t is defined in (5) as before. Again, we avoid using the double array notation to ease the presentation.

Theorem **3**. *Let $\{X_t\}_{t=1}^n$ be generated from (9). Assume that*

$$\frac{(\log n)^3}{m} + \frac{m}{n^{2/3}} \to 0 \tag{10}$$

and $f_u(\lambda) = f_u(0)(1 + O(\lambda^2))$ as $\lambda \downarrow 0$. Further assume $u_t \in \mathcal{L}^q$, $q > 4$,

$$\sum_{k_1,k_2,k_3 \in \mathbb{Z}} |\mathrm{cum}(u_0, u_{k_1}, u_{k_2}, u_{k_3})| < \infty \tag{11}$$

and

$$\sum_{k=1}^{\infty} k\delta_q(k) < \infty. \tag{12}$$

Then $t_m \to_D N(2\delta, 1)$.

See Appendix B for the proof of Theorem 3.

Remark **3**. The conditions (10), (11) and (12) were originally imposed in Shao and Wu (2005b) to study asymptotic properties of local Whittle estimator for general fractionally integrated nonlinear processes. The fourth cumulants summability condition (11) is a common assumption adopted in spectral analysis [Brillinger (1975)]. For nonlinear processes (5), it is satisfied under a geometric moment contraction (GMC) condition with order 4 (Wu and Shao, 2004). The process $\{u_t\}$ is GMC(q), $q > 0$, if there exists a $C = C(q)$ and $\rho = \rho(q) \in (0, 1)$ such that

$$\mathbb{E}(|u_n^* - u_n|^q) \le C\rho^n,\ n \in \mathbb{N}, \tag{13}$$

where $u_n^* = F(\cdots, \varepsilon'_{-1}, \varepsilon'_0, \varepsilon_1, \cdots, \varepsilon_n)$ is a coupled version of u_n with all the past innovations $\{\varepsilon_t\}_{t \le 0}$ coupled. The property (13) indicates that the process $\{u_n\}$ forgets its past exponentially fast and it can be verified for many nonlinear time series models [Wu and Min (2005)]. Theorem 4.1 of Shao and Wu (2005b) provides another set of sufficient conditions for (11). Note that both (7) and (12) result from (13).

Theorem 3 provides a theoretical justification for the claim made in Lobato and Robinson (1998) in the univariate case for general nonlinear processes. It shows that the exact local power of LM test is of order $O(m^{-1/2})$, where m is allowed to grow as fast as n^{κ}, $\kappa \in (0, 2/3)$. For linear processes $u_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$, Theorem 3 holds if $m^5 (\log m)^2 = o(n^4)$, $\sum_{k=0}^{\infty} |k a_k| < \infty$ and $\varepsilon_1 \in \mathcal{L}^4$. In other words, κ is allowed to fall within $(0, 4/5)$.

4. An Empirical study

In this section, we revisit the CRSP (Center for Research in Security Prices) data which contains equal-weighted daily indices of stock returns from July 3 1962 to Dec 31 1987 with a sample size of $n = 6409$; see Willinger et al. (1999). For this series there is still no consensus about the existence of long memory. Lo (1991) applied modified R/S test statistics and found no evidence for long memory. Later Willinger et al. (1999) questioned Lo's conclusion based on the argument that modified R/S test shows a strong preference for accepting the null hypothesis of no long memory, irrespective of the existence of long memory. The latter authors fitted a FARIMA$(1, d, 1)$ model using the full Whittle likelihood, i.e.

$$(1 - B)^d (1 - \phi B)(X_t - \mu) = (1 + \psi B)\varepsilon_t.$$

The resulting estimates are $\hat{d} = 0.13$, $\hat{\phi} = -0.16$ and $\hat{\psi} = -0.40$ respectively. Note that μ is estimated by the sample mean 0.000568. To verify the results of Willinger et al., we fit a FARIMA$(1, d, 1)$ model to the demeaned data using time domain maximum likelihood. The implementation was done in R using the function `fracdiff`. The estimates are comparable to those obtained by Willinger et al.:

	$\hat{d}$	$\hat{\phi}$	$\hat{\psi}$
Estimate	0.162	-0.159	-0.398
Standard Error	0.0107	0.00622	0.00600

Willinger et al. also tried other methods to isolate the "pure" long memory by removing the "extra" short-range dependence.

Figure 1 shows the autocorrelation of the original indices, absolute indices and squared indices. The absolute indices show obvious long memory feature while the squared indices appear to possess significant autocorrelations at lags 1 through 10, which indicates the existence of conditional heteroskedasticity. To see if a FARIMA$(1, d, 1)$ model can capture these empirical features, we simulate a Gaussian FARIMA$(1, d, 1)$ with $\phi = -0.159, \psi = -0.398, d = 0.162, \mu = 0.000568$ and iid standard normal innovations. Figure 2 shows the autocorrelation of one realization, the absolute value of this realization and the square of the realization. It can be seen that the autocorrelation of the original indices is well mimicked by the simulation, but the absolute value and square of the realization do not possess the strong autocorrelation as we see in Figures 1b and 1c. We also tried other type of innovations, for example, t distribution with various degrees of freedom. However, none of them can mimick the strong correlations of absolute and squared indices. This seems to suggest that FARIMA$(1, d, 1)$ is not an appropriate model for the equal-weighted indices. Note that we are mainly interested in testing whether the equal-weighted indices themselves have long memory or not. To avoid the impact of model misspecification, we shall use nonparametric or semiparametric LM tests. Recently, Lima and Xiao (2004) applied modified R/S, KPSS, V/S, K/S tests with a partially adaptive bandwidth selection that is robust against the presence of long memory. No long memory for the original indices was detected in their paper.

Like nonparametric tests, semiparametric LM test also involves the bandwidth selection issue. Since LM test has a close connection to local Whittle estimation, Lobato and Robinson (1998) used the optimal bandwidth for local Whittle estimator, which was first derived by Henry and Robinson (1996); also see Henry (2001). It takes the form

$$m_{opt} = \left(\frac{3n}{4\pi}\right)^{4/5} |E_2|^{-2/5},$$

where E_2 is related to the smoothness of the spectral density of $\{u_t\}$ in a vicinity of zero frequency and it has to be estimated in practice. However, as argued in Geweke's comment to Lobato and Savin (1998), this bandwidth does not satisfy the restriction on m for the asymptotic null distribution of LM tests; see Assumption 3 in Lobato and Robinson (1998). The reason is that m_{opt} is proposed to minimize the mean square error of local Whittle estimator, i.e. the bias and variance of local Whittle estimator are of the same magnitude if we use m_{opt}. While in the LM test the variance has to dominate the bias in order to ensure the asymptotic normality. So the ideal bandwidth of LM test should be smaller than the optimal bandwidth for the local Whittle estimation. Although m_{opt} is not directly useful, it provides an upper bound for the bandwidth we choose. To find m_{opt}, we tried the following "direct" iterated procedure proposed in Henry and Robinson (1996). Suppose at step k, we have $(m^{(k)}, G^{(k)}, d^{(k)})$. Then at step $k+1$,

- Perform the local Whittle estimation with bandwidth $m^{(k)}$ and get the estimator $d^{(k+1)}$.
- Update $G^{(k)}$ by $G^{(k+1)} = [m^{(k)}]^{-1} \sum_{j=1}^{m^{(k)}} I_X(\lambda_j)\lambda_j^{2d^{(k+1)}}$.
- Fit the following linear regression model,

$$\frac{I_X(\lambda_j)\lambda_j^{2d^{(k+1)}}}{G^{(k+1)}} = a^{(k+1)} + b^{(k+1)}\lambda_j^2, \quad j = 1, \cdots, m^{(k)}.$$

  Denote the least square estimates by $\hat{a}^{(k+1)}$ and $\hat{b}^{(k+1)}$.
- Update the bandwidth by $m^{(k+1)} = \left(\frac{3n}{4\pi}\right)^{4/5} |\hat{b}^{(k+1)}|^{-2/5}$. Here we take the integer part of $m^{(k+1)}$ for the use of next iteration.

- Start the new round and iterate until the convergence.

To start the above iteration, one has to specify an initial bandwidth $m^{(0)}$. Table 1 shows $m^{(k)}, k = 0, 1, \cdots, 18$ for a range of starting values. The convergence is not guaranteed. Nevertheless, it seems to suggest that m_{opt} is less than 125.

As suggested by Lobato and Savin (1998), it is helpful to plot the LM test statistics on a grid of bandwidths. We calculate t_m for each $m \in [51, 192]$. Let $Z \sim N(0, 1)$. Figure 3 shows $P(Z > t_m)$ versus m. The p-value is sensitive to the choice of the bandwidth. It is less than 0.05 if $m \geq 135$ and less than 0.01 if $m \geq 154$. Since the optimal bandwidth for LM test is less than m_{opt}, it is safe to say that there is no significant evidence (at 5% level) to support the existence of long memory in the original indices. However, the issue of choosing the optimal bandwidth for both local Whittle estimation and LM test remains open. We conjecture that resampling based bandwidth selection rule might work better.

5. Conclusions

In this paper, we consider local asymptotic powers of four nonparametric tests for fractional integration. The result seems surprising in that these tests only have nontrivial power against local fractional alternative such as $d_n = c(\log n)^{-1}$, where c is a fixed constant. In contrast, Lobato and Robinson's semiparametric LM test is shown to have a local power of order $O(m^{-1/2})$, where m can grow as fast as n^{κ}, $\kappa \in (0, 2/3)$ for nonlinear processes. We conclude that, to avoid possible model misspecifications in conducting parametric tests, semiparametric LM test is preferable to all the nonparametric tests discussed in this paper, at least in the large sample case. On the other hand, a reliable bandwidth selection for LM test is not yet available. This limits the use of LM test in practice and motivates us to pursue further study along this direction.

We further remark that our theoretical investigation is limited to the framework (1). Recently, there have been a substantial amount of work on testing and estimation of long memory in the volatility of financial time series [cf. Hurvich and Soulier (2002), Hurvich et al. (2005) and the references therein]. It is not clear whether in that setting one can still obtain similar results for nonparametric and semiparametric LM tests.

Acknowledgements

The authors would like to thank Sheng Guo for providing us with the CRSP data used in the paper.

Appendix A

Lemma **1**. *Let* $f(d, j) := [\Gamma(d+j)/\Gamma(j+1) - j^{d-1}]/j^{d-2}$, $j \in \mathbb{N}$. *Then for any fixed* $M > 0$, *there exists a finite constant* $C = C(M)$ *such that*

$$\sup_{d \in [-M, M]} \sup_{j \in \mathbb{N}} |f(d, j)| < C.$$

Proof of Lemma 1: The proof follows from Stirling's formula $\Gamma(z) = e^{-z}z^{z-1/2}\sqrt{2\pi}(1 + z/12 + O(z^{-2}))$ and some elementary algebras. $\diamondsuit$

By the binomial expansion, we write $X_t = \sum_{j=0}^{\infty} \phi_j(d_n)u_{t-j}$, where $\phi_j(d_n) = \Gamma(d_n + j)/\{\Gamma(j+1)\Gamma(d_n)\}$. Note that $\phi_0(d) = 1$ for any $d \in \mathbb{R}$. For $1 \leq m \leq n$, write $S_m = \sum_{t=1}^{m} X_t = \sum_{k=0}^{\infty} A_{k,m}u_{m-k}$, where $A_{k,m} = \psi_k - \psi_{k-m}$, $\psi_k = \sum_{j=0}^{k} \phi_j(d_n)$ if $k \geq 0$ and 0 if $k < 0$. Let $\xi_t = \sum_{k=t}^{\infty} \mathcal{P}_t u_k$, $T_n = \sum_{t=1}^{n} u_t$ and $M_n = \sum_{t=1}^{n} \xi_t$. We approximate S_m by $\tilde{S}_m = \sum_{k=0}^{\infty} A_{k,m}\xi_{m-k}$. Let $R_m = S_m - \tilde{S}_m$ be the remainder term.

LEMMA **2**. *Suppose X_t is generated from (4) and $u_t \in \mathcal{L}^q$, $q \geq 2$. Assume (7), then for any $l = l(n) = \lfloor n^\alpha \rfloor$, $\alpha \in (0,1]$, $l^{-1}\mathbb{E}S_l^2 \to \sigma^2 e^{2c\alpha}$.*

Proof of Lemma 2: By Theorem 1 of Wu (2005a), (7) implies $\|T_n - M_n\|_q = o(\sqrt{n})$ for $q \geq 2$. By Karamata's theorem and summation by parts, for $q \geq 2$,

$$\begin{aligned}
\|R_l\|_q &\leq \sum_{k=0}^{2l} |A_{k,l} - A_{k-1,l}|\|T_k - M_k\|_q + \sum_{k=2l+1}^{\infty} |A_{k,l} - A_{k-1,l}|\|T_k - M_k\|_q \\
&= \sum_{k=0}^{2l} k^{d_n-1}o(\sqrt{k}) + \sum_{k=2l+1}^{\infty} k^{d_n-1}o(\sqrt{k}) = o(\sqrt{l}). \qquad (14)
\end{aligned}$$

Since $\{\xi_t\}_{t\in\mathbb{Z}}$ are stationary martingale differences, we have $\|\sum_{k=0}^{\infty} A_{k,l}\xi_{l-k}\|^2 = \sum_{k=0}^{\infty} A_{k,l}^2\|\xi_1\|^2$. Then our conclusion follows from the fact that $\|\xi_1\|^2 = \sigma^2$ [cf. Wu (2005a)], $\sum_{k=0}^{l-1}\psi_k^2 \sim \sum_{k=0}^{l-1} k^{2d_n}\Gamma(d_n+1)^{-2} \sim e^{2c\alpha}l$ and $\sum_{k=l}^{\infty} A_{k,l}^2 \sim \sum_{k=l}^{\infty}(k^{d_n} - (k-l)^{d_n})^2 \leq l^{2d_n+1}\sum_{h=1}^{\infty}[h^{d_n} - (h-1)^{d_n}]^2 = o(l)$, where we have applied Lemma 1. $\diamondsuit$

REMARK **4**. If d_n decays to zero faster than $(\log n)^{-1}$, then it is easy to see that $l^{-1}\mathbb{E}S_l^2 \to \sigma^2$.

Proof of Theorem 1: Set $C_q = 18q^{3/2}(q-1)^{-1/2}$. By (14) and Proposition 1 in Wu (2005a),

$$\left\|\max_{m\leq 2^k} |R_m|\right\|_q \leq \sum_{s=0}^{k} 2^{(k-s)/q}\|R_{2^s}\|_q = \sum_{s=0}^{k} 2^{(k-s)/q}o(2^{s/2}) = o(2^{k/2}).$$

Thus it suffices to show $n^{-1/2}\tilde{S}_{\lfloor nt\rfloor} \Rightarrow \sigma e^c I\!B(t)$ in $\mathcal{D}(0,1)$. The finite dimensional convergence follows from the argument in the proof of Lemma 4.1 of Shao and Wu (2005a). By Lemma 3 in Wu and Min (2005), $\|\tilde{S}_n\|_q^2 \leq C_q^2\|\xi_1\|^2\sum_{k=0}^{\infty} A_{k,n}^2 \leq Cn$, where C is a generic constant. Then for any $0 \leq t_1 \leq t \leq t_2 \leq 1$, we have

$$\begin{aligned}
\mathbb{E}(|\tilde{S}_{\lfloor nt_2\rfloor} - \tilde{S}_{\lfloor nt\rfloor}|^{q/2}|\tilde{S}_{\lfloor nt\rfloor} - \tilde{S}_{\lfloor nt_1\rfloor}|^{q/2}) &\leq \|\tilde{S}_{\lfloor nt_2\rfloor} - \tilde{S}_{\lfloor nt\rfloor}\|_q^{q/2}\|\tilde{S}_{\lfloor nt\rfloor} - \tilde{S}_{\lfloor nt_1\rfloor}\|_q^{q/2} \\
&\leq C(\lfloor nt_2\rfloor - \lfloor nt\rfloor)^{q/4}(\lfloor nt\rfloor - \lfloor nt_1\rfloor)^{q/4} \\
&\leq C(nt_2 - nt_1)^{q/2}.
\end{aligned}$$

Therefore the tightness follows from Theorem 15.6 of Billingsley (1968) and the proof is completed. $\diamondsuit$

Proof of Theorem 2: The convergence in (a)-(d) is a direct consequence of the assertion $w_{n,l}^2 \to_{\mathbb{P}} \sigma^2 e^{2c\alpha}$ and Theorem 1 by the continuous mapping theorem. By the same argument as in the proof of Theorem 3.1 of Shao and Wu (2005a), where the case of fixed d is treated, $w_{n,l}^2 \to_{\mathbb{P}} \sigma^2 e^{2c\alpha}$ holds in view of Lemma 2. We omit the details. $\diamondsuit$

Appendix B

For notational convenience, write $I_{Xj} = I_X(\lambda_j)$, $I_{uj} = I_u(\lambda_j)$, $f_{uj} = f_u(\lambda_j)$, $w_{Xj} = w_X(\lambda_j)$ and $w_{uj} = w_u(\lambda_j)$. Denote by $\alpha_n(\lambda) = (1 - e^{i\lambda})^{-d_m}$ and $\alpha_{nj} = \alpha_n(\lambda_j)$. Then it is easily seen that $\alpha_n(\lambda)$ is differentiable in a neighborhood of the origin $(0, \epsilon)$ and $\alpha_n'(\lambda) = O(|\alpha_n(\lambda)|\lambda^{-1})$ as $\lambda \downarrow 0$. Denote by $D(w) = D_n(w) = \sum_{t=1}^n e^{itw}$. Let $K(w) = (2\pi n)^{-1}|D(w)|^2$ be the Fejér's kernel.

The following lemmas correspond to Lemmas 6.1-6.3 in Shao and Wu (2005b), where fixed d is treated. Let $g_j = w_{Xj}/(|\alpha_{nj}|\sqrt{f_{uj}})$ and $h_j = w_{uj}/\sqrt{f_{uj}}$.

Lemma **3**. *Suppose $\{X_t\}_{t=1}^n$ is generated from (9). Under (12), the following relations hold uniformly over $1 \le k \le j \le m = o(n)$:*

$$|\mathbb{E}\{g_j \bar{g}_j\} - 1| + |\mathbb{E}\{h_j \bar{h}_j\} - 1| + |\mathbb{E}\{g_j \bar{h}_j\} - \alpha_n(-\lambda_j)/|\alpha_{nj}|| = O(\log j/j); \tag{15}$$

$\mathbb{E}\{g_j g_j\} = O(\log j/j)$, $\mathbb{E}\{g_j g_k\} = O(\log j/k)$, $\mathbb{E}\{g_j \bar{g}_k\} = O(\log j/k)$;
$\mathbb{E}\{h_j h_j\} = O(\log j/j)$, $\mathbb{E}\{h_j h_k\} = O(\log j/k)$, $\mathbb{E}\{h_j \bar{h}_k\} = O(\log j/k)$;
$\mathbb{E}\{g_j h_j\} = O(\log j/j)$, $\mathbb{E}\{g_j h_k\} = O(\log j/k)$, $\mathbb{E}\{g_j \bar{h}_k\} = O(\log j/k)$.

Proof of Lemma 3: Since $\mathbb{E}[(u_k - u_k')|\mathcal{F}_0] = \mathcal{P}_0 u_k$ and $\|\mathcal{P}_0 u_k\| \le \|\mathcal{P}_0 u_k\|_q \le \delta_q(k)$ by Jensen's inequality, (12) implies $\sum_{k=0}^\infty k\|\mathcal{P}_0 u_k\| < \infty$. Thus we have $\sum_{k=0}^\infty |k\gamma(k)| < \infty$, where $\gamma(\cdot)$ is the autocovariance function of $\{u_t\}$; see the Appendix of Shao and Wu (2005b) for a rigorous proof. Then we have $\alpha_n'(\lambda) = O(\alpha_n(\lambda)\lambda^{-1})$ and $f_u'(\lambda) = O(\lambda^{-1})$ as $\lambda \downarrow 0$. The rest of the proof follows from the argument in the proof of Theorem 2 of Robinson (1995a), where d is fixed. We omit the details. $\diamondsuit$

Lemma **4**. *Suppose $\{X_t\}_{t=1}^n$ is generated from (9). Under (12), we have*

$$\mathbb{E}||g_j|^2 - |h_j|^2| = O(j^{-1/2}) \ \ \textit{uniformly over } j = 1, \cdots, m. \tag{16}$$

Further, if (11) holds, then we have

$$\mathbb{E}\left|\sum_{j=1}^r (|g_j|^2 - |h_j|^2)\right| \le C(r^{1/4}(1 + \log r)^{1/2} + r^{1/2}n^{-1/4}),\ r \le m = o(n), \tag{17}$$

where C is a generic constant independent of r, m and n.

Proof of Lemma 4: For (16), the argument of Lemma 2 of Shao and Wu (2005b) applies. The key step is to show that

$$\int_{-\pi}^{\pi} K(\lambda-\lambda_j)\left|\frac{\alpha_n(\lambda)}{\alpha_{nj}}-1\right|^2 = O(1/j) \text{ uniformly over } j=1,\cdots,m. \tag{18}$$

Using the same argument as Lemma 3 of Robinson (1995b), by the properties of $\alpha_n(\lambda)$, (18) holds. The assertion (17) corresponds to Lemma 6.3 of Shao and Wu (2005b), where the argument uses (11) and (18). The conclusion follows. $\diamondsuit$

Proof of Theorem 3: Let $b_{nj} = v_j|\alpha_{nj}|^2$. We first list some useful facts.
Fact 1. $|\alpha_{nj}|^2 = 1 + O((\log n)/\sqrt{m})$ uniformly over $j = 1,\cdots,m$.
Fact 2. $|b_{nj} - b_{n(j+1)}| \le Cj^{-1}$ uniformly over $j = 1,\cdots,m$. Here C is a generic constant.
Fact 3. Under (12), $\text{cov}(w_{uj}, \bar{w}_{uk}) = f_{uj}\mathbf{1}(j=k) + O(1/n)$ and $\text{cov}(w_{uj}, w_{uk}) = O(1/n)$ uniformly over $j,k = 1,\cdots,m$.
Fact 4. Under (11) and (12), we have $\text{cov}(I_{uj}, I_{uk}) = f_{uj}^2\mathbf{1}(j=k) + O(1/n)$ uniformly over $j,k = 1,\cdots,m$.
Facts 1&2 follow from elementary calculations. By (12), $\sum_{k=0}^{\infty}|k\gamma(k)| < \infty$, which yields

$$\text{cov}(w_{uj}, \bar{w}_{uk}) = \frac{1}{2\pi n}\sum_{t,s=1}^{n}\gamma(t-s)e^{it\lambda_j - is\lambda_k} = f_{uj}\mathbf{1}(j=k) + O(1/n).$$

So the first assertion of Fact 3 holds and the second assertion similarly follows. Regarding Fact 4, we have

$$\begin{aligned}\text{cov}(I_{uj}, I_{uk}) &= \text{cum}(w_{uj}, \bar{w}_{uj}, w_{uk}, \bar{w}_{uk}) + \text{cov}(w_{uj}, w_{uk})\text{cov}(\bar{w}_{uj}, \bar{w}_{uk}) \\ &\quad + \text{cov}(w_{uj}, \bar{w}_{uk})\text{cov}(\bar{w}_{uj}, w_{uk}) = O(1/n) + f_{uj}^2\mathbf{1}(j=k),\end{aligned}$$

which follows from Fact 3 and (11).

We first deal with the denominator of (8). Fact 1 and (16) imply that

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m} I_{Xj} &= \frac{1}{m}\sum_{j=1}^{m}|g_j|^2|\alpha_{nj}|^2 f_{uj} = \frac{1}{m}\sum_{j=1}^{m}|h_j|^2|\alpha_{nj}|^2 f_{uj} + o_{\mathbb{P}}(1) \\ &= \frac{1}{m}\sum_{j=1}^{m} I_{uj} + o_{\mathbb{P}}(1) = f_u(0) + o_{\mathbb{P}}(1).\end{aligned} \tag{19}$$

The last equality above is due to Facts 3&4. Now we shall treat the numerator of (8). By Lemma 4 and Fact 2,

$$\begin{aligned}\mathbb{E}\left|\sum_{j=1}^{m} b_{nj}(|g_j|^2 - |h_j|^2)\right| &\le \sum_{r=1}^{m-1}|b_{nr} - b_{n(r+1)}|\mathbb{E}\left|\sum_{j=1}^{r}(|g_j|^2 - |h_j|^2)\right| \\ &\quad + |b_{nm}|\mathbb{E}\left|\sum_{j=1}^{m}(|g_j|^2 - |h_j|^2)\right| = o(\sqrt{m}).\end{aligned}$$

Then we have

$$\begin{aligned}\frac{1}{\sqrt{m}}\sum_{j=1}^{m} v_j I_{Xj} &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m} b_{nj}|g_j|^2 f_{uj} = \frac{f_u(0)}{\sqrt{m}}\sum_{j=1}^{m} b_{nj}|g_j|^2 + o_{\mathbb{P}}(1)\\ &= \frac{f_u(0)}{\sqrt{m}}\sum_{j=1}^{m} b_{nj}|h_j|^2 + o_{\mathbb{P}}(1) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} b_{nj} I_{uj} + o_{\mathbb{P}}(1). \qquad (20)\end{aligned}$$

By Fact 3, we get

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m} b_{nj}\mathbb{E}I_{uj} = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} b_{nj} f_{uj} + o(1) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} b_{nj} f_u(0) + o(1) = -2\delta f_u(0) + o(1),$$

where the last equality above follows from elementary calculations. By Theorem 2 of Wu and Shao (2005), $m^{-1/2}f_u(0)^{-1}\sum_{j=1}^m b_{nj}(I_{uj}-\mathbb{E}I_{uj}) \to_D N(0,1)$. Therefore $t_m \to_D N(2\delta, 1)$ in view of (19) and (20). $\diamondsuit$

Remark **5**. Condition (10) is needed to prove the asymptotic normality for $\sum_{j=1}^m b_{nj}(I_{uj} - \mathbb{E}I_{uj})$. By Fact 4,

$$\begin{aligned}\operatorname{var}\left(m^{-1/2}\sum_{j=1}^{m} b_{nj}(I_{uj}-\mathbb{E}I_{uj})\right) &= m^{-1}\sum_{j=1}^{m} b_{nj}^2 \operatorname{var}(I_{uj}) + O(m\log m/n)\\ &= m^{-1}\sum_{j=1}^{m} b_{nj}^2 f_{uj}^2 + o(1) = [f_u(0)]^2 + o(1).\end{aligned}$$

In other words, $t_m = O_{\mathbb{P}}(1)$ under the less restrictive assumption $m^5(\log m)^2 = o(n^4)$.

References

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

Giraitis, L., Kokoszka, P., Leipus, R. & Teyssiére, G. (2003). Rescaled variance and related tests for long memory in volatility and levels. *J. Econometrics* **112**, 265-294.

Henry, M. (2001). Robust automatic bandwidth for long memory. *J. Time Ser Anal.* **22**, 293-316.

Henry, M. & Robinson, P. M. (1996). Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In *Athens Conference on Applied Probability and Time Series Analysis*, Vol. II: *Time Series Analysis, In Memory of E. J. Hannan* (eds P. M. Robinson and M. Rosenblatt). New York: Springer, pp. 220-232.

Hurvich, C. M. & Soulier, P. (2002). Testing for long memory in volatility. *Econom. Theory* **18**, 1291-1308.

Hurvich, C. M., Moulines, E. & Soulier, P. (2005). Estimating long memory in volatility. *Econometrica* **73**, 1283-1328.

Künsch, H. (1987). Statistical aspects of self-similar processes, in Yu. A. Prohorov and V. V. Sazonov (eds.), *Proceedings of the 1st World Congress of the Bernoulli Society* Vol.1, Science Press, Utrecht, 67-74.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *J. Econometrics* **54**, 159-178.

Lee, D. & Schmidt, P. (1996). On the power of the KPSS test of stationarity against fractionally-integrated alternatives. *J. Econometrics* **73**, 285-302.

Lima, L. R. & Xiao, Z. (2004). Robustness of stationarity tests under long-memory alternatives. Preprint.

Lo, A. (1991). Long-term memory in stock market prices. *Econometrica* **59**, 1279-1313.

Lobato, I. & Robinson, P. M. (1998). A nonparametric test for I(0). *Rev. Econ. Stud.* **68**, 475-495.

Lobato, I. & Savin, N. E. (1998). Real and spurious long-memory properties of stock-market data (with comments). *J. Bus. Econ. Stat.* **16**, 261-283.

Nielsen, M. O. (2004). Efficient likelihood inference in nonstationary univariate models. *Econom. Theory* **20**, 116-146.

Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. *J. Am. Statist. Assoc.* **89**, 1420-1437.

Robinson, P. M. (1995a). Log-periodogram regression of time series with long range dependence. *Ann. Statist.* **23**, 1048-1072.

Robinson, P. M. (1995b). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23**, 1630-1661.

Shao, X. & Wu, W. B. (2005a). Invariance principles for fractionally integrated nonlinear processes. Technical Report 561, Department of Statistics, University of Chicago.

Shao, X. & Wu, W. B. (2005b). Local Whittle estimation of fractional integration for nonlinear processes. Technical Report 560, Department of Statistics, University of Chicago.

Tanaka, K. (1999). The nonstationary fractional unit root. *Econom. Theory* **15**, 549-582.

Willinger, W., Taqqu, M. S. & Teverovsky, V. (1999). Stock market prices and long-range dependence. *Finance Stochast.* **3**, 1-13.

Wright, J. H. (1999). The local asymptotic power of certain tests for fractional integration. *Econom. Theory* **15**, 704-709.

Wu, W. B. (2005a). A strong convergence theory for stationary processes. Preprint.

Wu, W. B. (2005b). Nonlinear system theory: another look at dependence. *P. Natl. Acad. Sci. USA.* **102**, 14150-14154.

Wu, W. B. & Min, M. (2005). On linear processes with dependent innovations. *Stochast. Process. Appl.* **115**, 939-958.

Wu, W. B. & Shao, X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41**, 425-436.

Wu, W. B. & Shao, X. (2005). A limit theorem for quadratic forms and its applications. Preprint.

Xiao, Z. (2001). Test the null hypothesis of stationarity against an autoregressive unit root alternative. *J. Time Ser. Anal.* **22**, 87-105.

Figure 1: (a). The autocorrelation plot of the original indices. (b). The autocorrelation plot of the absolute indices. (c). The autocorrelation plot of the squared indices.

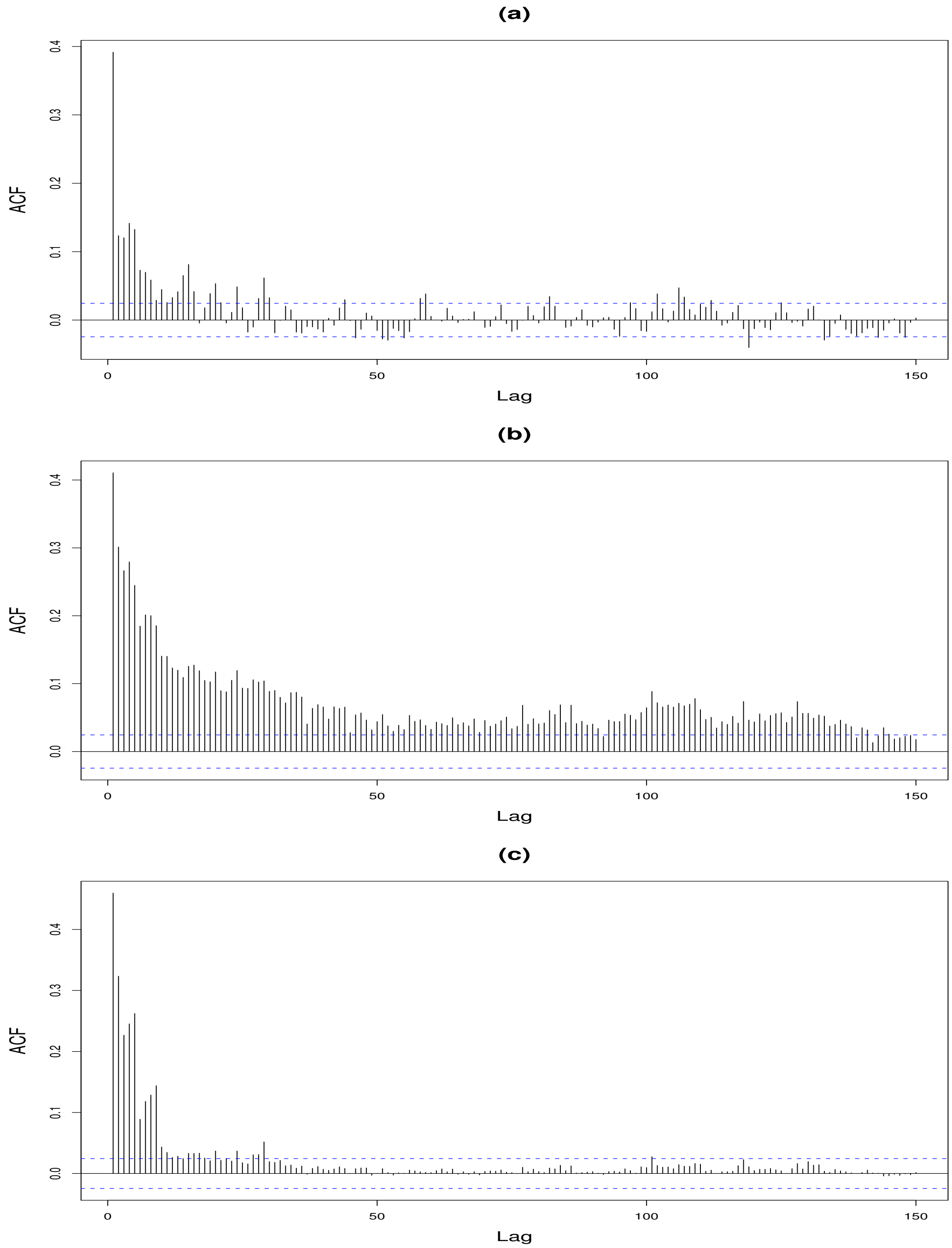

Figure 2: (a). The autocorrelation plot of a simulated Gaussian FARIMA(1,d,1). (b). The autocorrelation plot of the absolute value of the simulation. (c). The autocorrelation plot of the square of the simulation.

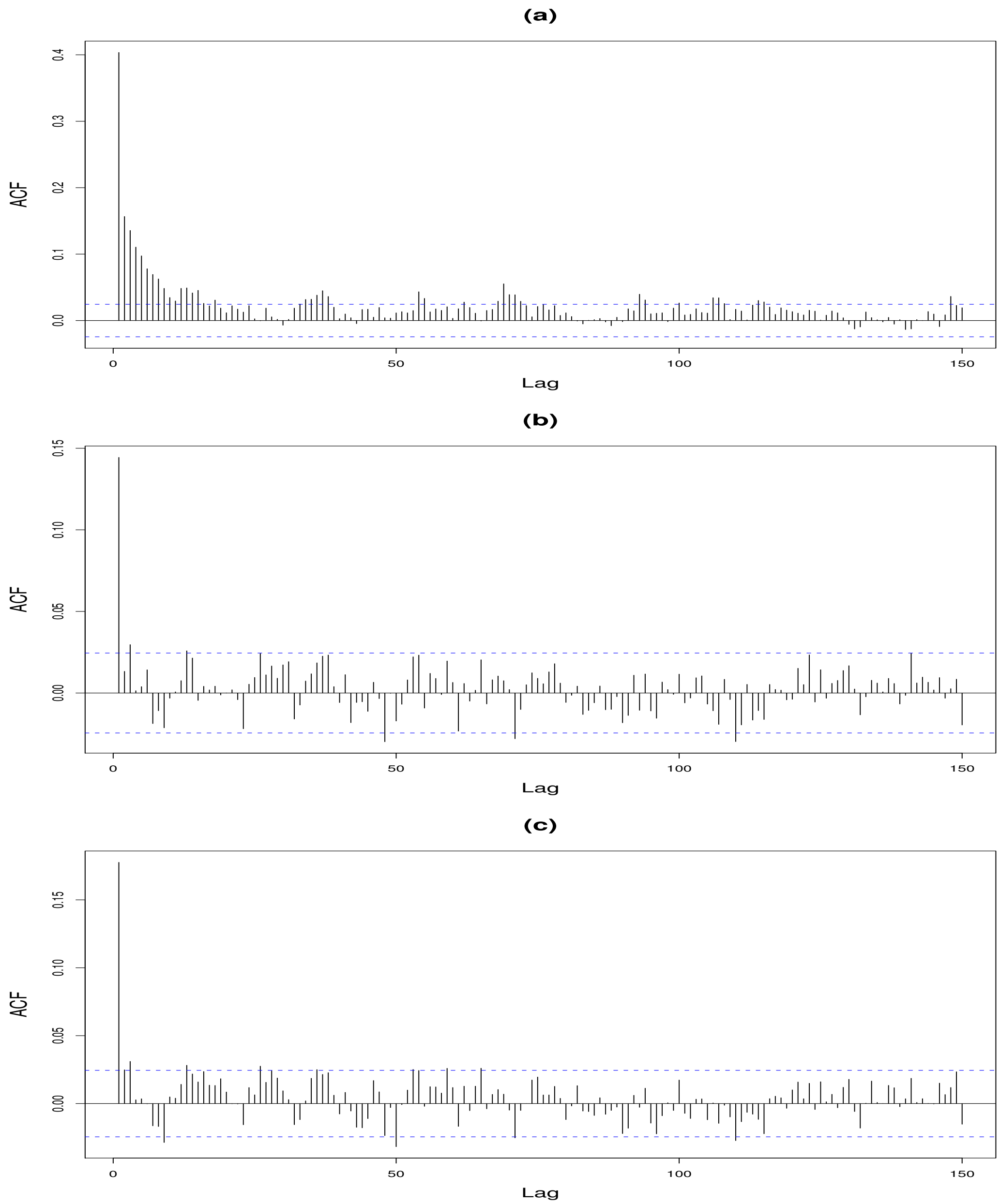

Table 1: The bandwidth $m^{(k)}$, $k = 0, 1, \cdots, 18$ for different starting values.

k	$m^{(0)} = 40$	$m^{(0)} = 60$	$m^{(0)} = 80$	$m^{(0)} = 120$	$m^{(0)} = 180$	$m^{(0)} = 200$	$m^{(0)} = 300$
0	40	60	80	120	180	200	300
1	35	51	151	123	145	175	248
2	33	54	120	121	120	140	230
3	65	64	123	122	123	114	190
4	63	57	121	123	121	98	161
5	59	55	122	121	122	91	123
6	95	62	123	122	123	91	121
7	90	55	121	123	121	91	122
8	96	62	122	121	122	91	123
9	87	55	123	122	123	91	121
10	110	62	121	123	121	91	122
11	100	55	122	121	122	91	123
12	102	62	123	122	123	91	121
13	99	55	121	123	121	91	122
14	97	62	122	121	122	91	123
15	88	55	123	122	123	91	121
16	102	62	121	123	121	91	122
17	99	55	122	121	122	91	123
18	97	62	123	122	123	91	121

Figure 3: The plot of p-value versus $m \in [51, 192]$.