The Annals of Statistics
1994, Vol. 22, No. 1, 515–539

SEMIPARAMETRIC ANALYSIS OF LONG-MEMORY TIME SERIES[1]

BY P. M. ROBINSON

London School of Economics

We study problems of semiparametric statistical inference connected with long-memory covariance stationary time series, having spectrum which varies regularly at the origin: There is an unknown self-similarity parameter, but elsewhere the spectrum satisfies no parametric or smoothness conditions, it need not be in L_p, for any $p > 1$, and in some circumstances the slowly varying factor can be of unknown form. The basic statistic of interest is the discretely averaged periodogram, based on a degenerating band of frequencies around the origin. We establish some consistency properties under mild conditions. These are applied to show consistency of new estimates of the self-similarity parameter and scale factor. We also indicate applications of our results to standard errors of least squares estimates of polynomial regression with long-memory errors, to generalized least squares estimates of this model and to estimates of a "cointegrating" relationship between long-memory time series.

1. Introduction. There is now considerable theoretical and practical evidence on the performance of methods of analyzing long covariance stationary time series which possess a smooth, but nonparametric, spectrum [see, e.g., Brillinger (1975)]. Much less is known about how to deal with a series whose nonparametric spectrum is unbounded at some frequencies, especially when interest centres on behaviour around the singularity. In particular, consider the process $x_t, t = 0, \pm 1, \dots$, which has mean μ, lag-j-autocovariance γ_j and spectrum $f(\lambda) = (2\pi)^{-1}\sum_{j=-\infty}^{\infty}\gamma_j \cos j\lambda$ which satisfies the following condition.

CONDITION A. For some $H \in \left(\frac{1}{2}, 1\right)$,

$$f(\lambda) \sim L\left(\frac{1}{\lambda}\right)\lambda^{1-2H} \quad \text{as } \lambda \to 0+. \tag{1.1}$$

The symbol "$\sim$" indicates that the ratio of left- and right-hand sides tends to 1 and $L(\lambda)$ is a slowly varying function at infinity [see, e.g., Seneta (1976)], that is, a positive, measurable function satisfying

$$\frac{L(t\lambda)}{L(\lambda)} \to 1 \quad \text{as } \lambda \to \infty, \text{ for all } t > 0. \tag{1.2}$$

Received August 1991; revised March 1993.
[1]This article is based on research funded by the Economic and Social Research Council (ESRC) reference number R000233609.
AMS 1991 *subject classifications.* Primary 62M15; secondary 62G05, 60G18.
Key words and phrases. Long-memory time series, semiparametric inference, regular variation, autocorrelation-consistent standard errors, cointegration.

Thus Condition A asserts that $f(\lambda)$ is regularly varying at $\lambda = 0$ and is unbounded at $\lambda = 0$. (Our work can likely be extended to the case where the singularity is known to occur at some given nonzero frequency.) For some purposes L can be of unknown form; for others we require

$$L(\lambda) = GM(\lambda), \qquad G > 0, \quad M(\lambda) \text{ is a known function,} \tag{1.3}$$

where G is unknown. In either case, because (1.1) makes no parametric assumptions about f outside a neighbourhood of the origin, we can call (1.1) a semiparametric model. Indeed, apart from being integrable (due to covariance stationarity), f is not even required to satisfy any smoothness assumptions and need not be in L_p, for any $p > 1$ [it would be in $L_p, p < 1/(2H-1)$, if f were smooth away from $\lambda = 0$].

A time series exhibiting property (1.1) is often called long-memory, or long-range dependent. Two examples of such series are described by the "fractional noise" spectrum

$$\begin{aligned} f(\lambda) = {} & \frac{\sigma^2}{\pi}\sin(\pi H)\,\Gamma(2H+1)(1-\cos\lambda) \\ & \times \sum_{j=-\infty}^{\infty} |\lambda + 2\pi j|^{-2H-1}, \qquad -\pi < \lambda \leq \pi \end{aligned} \tag{1.4}$$

[see, e.g., Sinai (1976)] and the "fractional ARIMA" spectrum

$$f(\lambda) = \frac{\sigma^2}{2\pi}|1 - e^{i\lambda}|^{1-2H}\frac{|b(e^{i\lambda})|^2}{|a(e^{i\lambda})|^2}, \qquad -\pi < \lambda \leq \pi \tag{1.5}$$

[see, e.g., Adenstedt, (1974)], where $\sigma^2 > 0$ in (1.4) and (1.5), and a and b are polynomials of finite degree having no roots in or on the unit circle.

The main theoretical contribution of this paper is its study of the behaviour of a statistic which is familiar from the smooth spectrum estimation literature, in the unfamiliar circumstances of (1.1). The statistic is the discretely averaged periodogram, where the averaging is done over a neighbourhood of the origin which slowly degenerates to zero as sample size n increases. A suitably normalized version of the averaged periodogram is well known to converge in probability to the spectrum at the origin under weak dependence conditions [see, e.g., Brillinger (1975), Chapter 6]. Under Condition A we show that the ratio of the averaged periodogram, discretely averaged over $(0, \lambda)$, to

$$F(\lambda) = \int_0^\lambda f(\theta)\, d\theta \tag{1.6}$$

converges in probability to 1, for a sequence λ tending to 0 slower than $1/n$ as $n \to \infty$. This result is described in detail in Section 2 and proved in Section 3. We have to impose additional regularity conditions but these seem rather mild and, in particular, entail no further restriction on $f(\lambda)$ beyond

(1.1). Under somewhat stronger conditions we establish also a rate of convergence. The proofs in Section 3 contain some results which may be useful in other problems concerning long-memory time series. Section 4 manipulates averaged periodograms to obtain an estimate of H which the results of Section 2 imply is consistent in the presence of nonparametric L, and also gives a consistent estimate of G under (1.3). The limiting distribution of the estimates remains to be derived; it is likely to be normal for $\frac{1}{2} < H < \frac{3}{4}$ and nonnormal for $\frac{3}{4} < H < 1$ [cf. Fox and Taqqu (1985)]. Section 5 demonstrates the usefulness of our consistency results in the analysis of models involving long memory and a semiparametric aspect.

2. The averaged periodogram. There is a substantial literature concerning the estimation of $F(\lambda)$ (1.6). Much of this [e.g., Ibragimov (1963)] has employed the continuously averaged periodogram

$$\widetilde{F}(\lambda) = \int_0^\lambda \widetilde{I}(\theta)\, d\theta, \qquad 0 < \lambda \le \pi,$$

where

$$\widetilde{I}(\lambda) = |\widetilde{w}(\lambda)|^2, \qquad \widetilde{w}(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} (x_t - \mu) e^{it\lambda}. \tag{2.1}$$

It is possible to write $\widetilde{F}(\lambda)$ as a linear combination of sample autocovariances of x_t about μ. Unless μ is known, $\widetilde{F}(\lambda)$ is infeasible, and an alternative estimate which is invariant to μ and also lends itself to somewhat more direct computation is

$$\widehat{F}(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{[n\lambda/2\pi]} I(\lambda_j),$$

where $[\cdot]$ here denotes integer part and $\lambda_j = 2\pi j/n$,

$$I(\lambda) = |w(\lambda)|^2, \qquad w(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t e^{it\lambda}.$$

Like $\widetilde{F}, \widehat{F}$ is centered at μ because $\widetilde{w}(\lambda_j) = w(\lambda_j)$, for all integers $j \neq 0, \bmod(n)$.

There are various available asymptotic statistical results relevant to $\widetilde{F}(\lambda)$ and $\widehat{F}(\lambda)$, including pointwise ones for fixed λ, and functional limit results on $[0, \pi]$. The bulk of these assume x_t is weakly dependent (e.g., having a spectrum which is at least bounded), although there are some results [e.g., Ibragimov (1963)] which cover long-memory processes, although with λ not depending on n. A very important thread to the time series literature has been the use of $\widetilde{F}$ and $\widehat{F}$ in weakly dependent series, in particular, $\widetilde{F}(\lambda_m)/\lambda_m$

and $\widehat{F}(\lambda_m)/\lambda_m$ can be consistent estimates of a finite $f(0)$ when the following assumption holds:

CONDITION B. $1/m + m/n \to 0$ as $n \to \infty$.

[See, e.g., Brillinger (1975), Chapter 7.] $\widetilde{F}(\lambda_m)/\lambda_m$ and $\widehat{F}(\lambda_m)/\lambda_m$ are special cases of the general classes of weighted autocovariance and weighted periodogram spectrum estimates, respectively.

We wish to show that

$$\frac{\widehat{F}(\lambda_m)}{F(\lambda_m)} \to_p 1 \quad \text{as } n \to \infty, \tag{2.2}$$

under Conditions A and B and mild additional restrictions. Lemma 3(ii) in Section 3 and (1.1) imply

$$F(\lambda) \sim L\left(\frac{1}{\lambda}\right)\frac{\lambda^{2(1-H)}}{2(1-H)} \quad \text{as } \lambda \to 0+. \tag{2.3}$$

Thus we can substitute $L(\lambda_m^{-1})\lambda_m^{2(1-H)}/2(1-H)$ for the denominator in (2.2). In the sequel we refer to F only with a vanishing argument, so that we can treat (2.3) as an equality and substitute the right-hand side for $F(\lambda)$ whenever this occurs. Condition B is a minimal restriction on m, because on the one hand consistency requires the accumulation of information entailed in $m \to \infty$, whereas on the other the semiparametric model (1.1) describes f over only a negligible interval, so that λ_m must tend to 0 as $n \to \infty$. An optimality theory for m is described by Robinson (1991). Because $\lambda^{2(1-H)}$ (unlike λ^{1-2H}) is continuous at $\lambda = 0$ for $\frac{1}{2} < H < 1$, $\widehat{F}(\lambda_m)$ seems a promising statistic for use in statistical inference on long-memory time series.

It is possible, indeed slightly easier, to study similarly $\widetilde{F}(\lambda_m)$ or a feasible version of it in which μ in (2.1) is replaced by the sample mean $\bar{x}$ of $x_1, \dots, x_n$. Due to the influence of $\bar{x}$ [which is only $(n)^{1-H}/L(n)^{1/2}$-consistent under (1.1)] the latter often provides a poor estimate in modest-sized samples, markedly inferior to $\widehat{F}(\lambda_m)$ (as Monte Carlo simulation evidence confirms). We thus focus on $\widehat{F}$.

To establish (2.2), Condition A can be strengthened in either of two leading directions. One is to impose fourth moment conditions and prove mean square convergence. However, the main practical interest of the paper is not in $\widehat{F}$ itself but in certain nonlinear functions of $\widehat{F}$ and of analogous statistics; since only convergence in probability of these can be inferred from mean square convergence of $\widehat{F}$ we prefer to invest in conditions capable of leading only to convergence in probability of $\widehat{F}$, or little more than that. Instead of strengthening the necessary second moment assumption on x_t, we instead restrict the dependence and heterogeneity of the white noise innovations e_t in the Wold representation

$$x_t = \mu + \sum_{j=0}^{\infty} \alpha_j e_{t-j}, \qquad \sum_{j=0}^{\infty} \alpha_j^2 < \infty. \tag{2.4}$$

CONDITION C. x_t satisfies (2.4), where the following hold:

(i) $E(e_t e_u) = 0, \quad t > u.$

(ii) $E(e_r e_s e_t e_u) = \begin{cases} \sigma^4, & \text{if } r = s > t = u, \\ 0, & \text{if } r = s > t > u, \text{ or } r > s = t > u, \text{ or } r > s > t \geq u. \end{cases}$

(iii) There exists a nonnegative random variable e such that, for all $\eta > 0$ and some $K < 1$,

$$E(e^2) < \infty, \qquad P(|e_t| > \eta) \leq KP(e > \eta).$$

(iv) $$\frac{1}{n}\sum_{t=1}^{n} E\left(e_t^2 \mid e_s^2, s < t\right) \to_p \sigma^2 \quad \text{as } n \to \infty.$$

Condition C is satisfied if the e_t are an independent identically distributed (iid) sequence with finite variance, or, more generally, if the e_t and $e_t^2 - \sigma^2$ form integrable martingale difference sequences. To check the latter claim, denote by $\mathcal{F}_t$ the σ-field of events generated by $e_t, e_{t-1}, \ldots$. Then we have the following:

$$\begin{aligned} E(e_t e_u) &= E(e_u E(e_t \mid \mathcal{F}_{t-1})) = 0, \qquad \text{for } t > u; \\ E(e_t^2 e_r^2) &= E(e_t^2 E(e_r^2 \mid \mathcal{F}_{r-1})) = \sigma^2 E(e_t^2) = \sigma^4, \qquad \text{for } r > t; \\ E(e_r^2 e_t e_u) &= E(e_t e_u E(e_r \mid \mathcal{F}_{r-1})) = 0, \qquad \text{for } r > t > u; \\ E(e_r e_s e_t e_u) &= E(e_t e_u E(e_s E(e_r \mid \mathcal{F}_{r-1}) \mid \mathcal{F}_{s-1})) = 0, \qquad \text{for } r > s > t > u; \end{aligned}$$

and

$$n^{-1}\sum_{1}^{n} \left\{E(e_t^2 \mid e_s^2,\, s < t) - \sigma^2\right\} = n^{-1}\sum_{1}^{n} E\left\{E(e_t^2 - \sigma^2 \mid \mathcal{F}_{t-1}) \mid e_s^2,\, s < t) = 0\right\}.$$

Thus finiteness of only second moments is assumed, condition (iii) being a homogeneity restriction. The level of technical achievement in our proof of (2.2) under Condition C with only finiteness of second moments assumed seems substantially greater than a proof of mean square convergence under fourth moment conditions.

We do not need to strengthen the assumptions on the Wold representation weights α_j which are implicit in Condition A, but we note the implications of Condition A for their Fourier transform

$$\alpha(\lambda) = \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda}.$$

We have $f(\lambda) = (\sigma^2/2\pi)\{\alpha_1^2(\lambda) + \alpha_2^2(\lambda)\}$, where $\alpha_1(\lambda) = \mathrm{Re}\{\alpha(\lambda)\}$, $\alpha_2(\lambda) = \mathrm{Im}\,\{\alpha(\lambda)\}$. Thus, Condition (A) implies that either

$$\alpha_i(\lambda) \sim \alpha_i L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H} \quad \text{as } \lambda \to 0+, \quad i = 1, 2, \tag{2.5}$$

where $\alpha_1^2 + \alpha_2^2 = 2\pi/\sigma^2$, or

$$\alpha_i(\lambda) \sim \pm\left(\frac{2\pi}{\sigma^2}\right)^{1/2} L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H} \quad \text{as } \lambda \to 0+, \quad i = 1 \text{ or } 2, \tag{2.6}$$

and

$$\alpha_j(\lambda) = o\left(L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H}\right) \quad \text{as } \lambda \to 0+, \quad j \neq i. \tag{2.7}$$

Yong [(1974), pages 53, 64, 71 and 75] gave some conditions on the α_j under which (2.5) obtains.

THEOREM 1. *Let Conditions* A, B *and* C *hold. Then*

$$\frac{\widehat{F}(\lambda_m)}{F(\lambda_m)} \to_p 1 \quad \text{as } n \to \infty.$$

PROOF. Introduce

$$v(\lambda_j) = (2\pi n)^{-1/2}\sum_{t=1}^{n} e_t e^{it\lambda_j}, \qquad J(\lambda) = |v(\lambda)|^2.$$

Then

$$\widehat{F}(\lambda_m) - F(\lambda_m) = \frac{2\pi}{n}\sum_{j=1}^{m}\{I(\lambda_j) - |\alpha(\lambda_j)|^2 J(\lambda_j)\} \tag{2.8}$$

$$+\frac{2\pi}{n}\sum_{j=1}^{m} f(\lambda_j)\left\{\frac{2\pi J(\lambda_j)}{\sigma^2} - 1\right\} \tag{2.9}$$

$$+\frac{2\pi}{n}\sum_{j=1}^{m} f(\lambda_j) - F(\lambda_m). \tag{2.10}$$

(2.9) and (2.10) are each $o_p(F(\lambda_m))$ from the first part of, respectively, Propositions 3 and 1. (All propositions are stated and proved in the next section). The term on the RHS of (2.8) is the real part of

$$\frac{2\pi}{n}\sum_{j=1}^{m}\left\{\{w(\lambda_j) - \alpha(\lambda_j)\, v(\lambda_j)\right\}\left\{\overline{w}(\lambda_j) + \overline{\alpha}(\lambda_j)\overline{v}(\lambda_j)\right\},$$

whose modulus is bounded by

$$2\pi\left\{\frac{1}{n}\sum_{j=1}^{m}\left|w(\lambda_j)-\alpha(\lambda_j)v(\lambda_j)\right|^2\right\}^{1/2}\left\{\frac{2}{n}\sum_{j=1}^{m}I(\lambda_j)+\frac{2}{n}\sum_{j=1}^{m}\left|\alpha(\lambda_j)\right|^2 J(\lambda_j)\right\}^{1/2}.$$

The first term in the second pair of braces is $O_p(F(\lambda_m))$ by Proposition 2, the second term is nonnegative with expectation $O(n^{-1}\Sigma_{j=1}^{m}f(\lambda_j)) = O(F(\lambda_m))$ by Proposition 1, because $E\{J(\lambda_j)\} = (2\pi)^{-1}\sigma^2$. The term in the first pair of braces is $o_p(F(\lambda_m))$ by the first part of Proposition 4. □

If Conditions A and C are strengthened it is possible to calculate a stochastic order of magnitude of $\widehat{F}(\lambda_m)/F(\lambda_m) - 1$, and this will actually be of practical use in Sections 4 and 5.

CONDITION A′. For some $\tau > 0$, either of the following hold: (i)

$$\alpha_i(\lambda) - \alpha_i L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H} = O\left(L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H+\tau}\right) \quad \text{as } \lambda \to 0+,\ i = 1,2, \tag{2.11}$$

or (ii)

$$\alpha_i(\lambda) - \left(\frac{2\pi}{\sigma^2}\right)^{1/2} L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H} = O\left(L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H+\tau}\right), \quad \text{as } \lambda \to 0+,\ i = 1 \text{ or } 2, \tag{2.12}$$

and

$$\alpha_j(\lambda) = O\left(L^{1/2}\left(\frac{1}{\lambda}\right)\lambda^{1/2-H+\tau}\right), \quad \text{as } \lambda \to 0+, j \neq i. \tag{2.13}$$

Also, $L(\lambda) = L_1(\lambda)(1 + O(\lambda^{-\tau}))$ as $\lambda \to \infty$, where $L_1(\lambda)$ is differentiable.

Condition A′ is satisfied with $\tau = 2$ in case of models (1.4) and (1.5). Generally it corresponds approximately to the representation $f(\lambda) = L(1/|\lambda|)\cdot|\lambda|^{1-2H}g(\lambda)$, where g is an even function which satisfies a Lipschitz condition of degree τ, for $0 < \tau \leq 1$, or is differentiable with derivative satisfying a Lipschitz condition of degree $\tau - 1$, for $1 < \tau \leq 2$.

CONDITION C′. (i) and (ii) of Condition C hold, plus, for some $\nu > 0$, the following hold:

(iii)′ $$\max_{t\geq 1} E\left|e_t\right|^{2+\nu} < \infty;$$

(iv)′ $$\sum_{t=2}^{n}\left\{E\left(e_t^2\left|\sum_{s=1}^{t-1}e_s^2\right.\right) - \sigma^2\right\} = O_p\left(n^{2/(2+\nu)}\right) \quad \text{as } n \to \infty.$$

Condition C′ strengthens the moment assumptions on x_t, and (iv)′ is again trivially satisfied if $e_t^2 - \sigma^2$ is a martingale difference or iid sequence.

THEOREM 2. *Under Conditions* A′, B *and* C′, *for* $\delta < \frac{1}{2}(1-H)/(2-H)$,

$$\frac{\widehat{F}(\lambda_m)}{F(\lambda_m)} - 1 = O_p\left(n^{-\nu/(2+\nu)} + m^{-\delta} + \left(\frac{m}{n}\right)^{\tau}\right), \quad \text{as } n \to \infty.$$

PROOF. Same as that of Theorem 1, but using the second part of Propositions 1, 3 and 4. □

3. Lemmas and propositions. This section states and proves the propositions referred to in Section 2. The proofs make use of some lemmas, which are presented first.

LEMMA 1. *For all* $\eta > 0$, *as* $\lambda \to 0+$,

(i) $$\sup_{\theta \geq \lambda} L\left(\frac{1}{\theta}\right)\theta^{-\eta} \sim L\left(\frac{1}{\lambda}\right)\lambda^{-\eta},$$

(ii) $$\sup_{\theta \leq \lambda} L\left(\frac{1}{\theta}\right)\theta^{\eta} \sim L\left(\frac{1}{\lambda}\right)\lambda^{\eta}.$$

PROOF. A trivial rewriting of, for example, property 4 of Seneta [(1976), pages 20–21]. □

LEMMA 2. *For any* $\varepsilon \in (0,1]$, $\eta > 0$, $\delta \in (0,\eta)$, *as* $\lambda \to 0+$,

$$L\left(\frac{1}{\varepsilon\lambda}\right)(\varepsilon\lambda)^{\eta} = O\left(\varepsilon^{\eta-\delta} L\left(\frac{1}{\lambda}\right)\lambda^{\eta}\right).$$

PROOF. From Lemma 1(ii),

$$L\left(\frac{1}{\varepsilon\lambda}\right)(\varepsilon\lambda)^{\eta} = (\varepsilon\lambda)^{\eta-\delta} L\left(\frac{1}{\varepsilon\lambda}\right)(\varepsilon\lambda)^{\delta} = O\left((\varepsilon\lambda)^{\eta-\delta} L\left(\frac{1}{\lambda}\right)\lambda^{\delta}\right).$$ □

LEMMA 3. *For all* $\eta > 0$, *as* $\lambda \to 0+$,

(i) $$\int_{\lambda}^{\Lambda} L\left(\frac{1}{\theta}\right)\theta^{-\eta-1}\,d\theta \sim \frac{1}{\eta} L\left(\frac{1}{\lambda}\right)\lambda^{-\eta}$$

if Λ *is so small that* $L(\lambda)$ *is locally bounded in* $[1/\Lambda, \infty)$, *and*

(ii) $$\int_{0}^{\lambda} L\left(\frac{1}{\theta}\right)\theta^{\eta-1}/d\theta \sim \frac{1}{\eta} L\left(\frac{1}{\lambda}\right)\lambda^{\eta}.$$

PROOF. Trivially rewrite (1–32′) and (1–32″) of Yong [(1974), page 20]. □

LEMMA 4. *Under Condition* B, *if* $0 < \eta < 1$,

$$\frac{1}{n}\sum_{j=1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(L\left(\frac{1}{\lambda_m}\right)\lambda_m^{1-\eta}\right) \quad as\ n \to \infty.$$

PROOF. Because $m/n \to 0$, Lemma 1(i) gives

$$\limsup_{n\to\infty}\max_{1\le j\le m}\sup_{\lambda\le\lambda_j}\left\{L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta}\Big/L\left(\frac{1}{\lambda}\right)\lambda^{-\eta}\right\}\le 1, \tag{3.1}$$

and thus

$$\limsup_{n\to\infty}\max_{1\le j\le m}\left\{L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta}\Big/\frac{n}{2\pi}\int_{\lambda_{j-1}}^{\lambda_j}L\left(\frac{1}{\lambda}\right)\lambda^{-\eta}\,d\lambda\right\}\le 1, \tag{3.2}$$

so, as $n \to \infty$,

$$\frac{1}{n}\Sigma_{j=1}^{m}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(\int_0^{\lambda_m}L\left(\frac{1}{\lambda}\right)\lambda^{-\eta}\,d\lambda\right).$$

Now apply Lemma 3(ii). □

LEMMA 5. *Under Condition* B, *if* $1 < \eta < 2$, *then, for some* $\delta > 0$,

$$\frac{1}{n^2}\sum_{j=1}^{m}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(m^{-\delta}L\left(\frac{1}{\lambda_m}\right)\lambda_m^{2-\eta}\right) \quad as\ n\to\infty.$$

PROOF. The LHS is of order

$$\frac{1}{n}\sum_{j=1}^{s}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-\eta} + \frac{1}{sn}\sum_{j=s+1}^{m}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-\eta},$$

for $1 \le s < m$, which by Lemma 4 is

$$O\left(L\left(\frac{1}{\lambda_s}\right)\lambda_s^{2-\eta} + \frac{1}{s}L\left(\frac{1}{\lambda_m}\right)\lambda_m^{2-\eta}\right) = O\left(\left\{\left(\frac{s}{m}\right)^{2-\eta-\varepsilon} + \frac{1}{s}\right\}L\left(\frac{1}{\lambda_m}\right)\lambda_m^{2-\eta}\right),$$

where $0 \le \varepsilon < 2-\eta$, using Lemma 2. Now choose $s \sim m^\delta$, where $\delta = \frac{2-\eta-\varepsilon}{3-\eta-\varepsilon}$. □

LEMMA 6. *Under Condition* B, *for* $\eta > 1$, *for any positive integer* $u < m$,

$$\frac{1}{n}\sum_{j=u+1}^{m}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(L\left(\frac{1}{\lambda_u}\right)\lambda_u^{1-\eta}\right) \quad as\ n\to\infty.$$

PROOF. By (3.1) and (3.2) (which hold for any $\eta > 0$),

$$\frac{1}{n}\sum_{j=u+1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(\int_{\lambda_u}^{\lambda_m} L\left(\frac{1}{\lambda}\right)\lambda^{-\eta}\,d\lambda\right) \quad \text{as } n \to \infty.$$

The result follows from Lemma 3(i) because $L(\lambda)$ is locally bounded in $[\Lambda, \infty)$ for large enough Λ, and $m/n \to 0$. □

LEMMA 7. *Under Condition* B, *and assuming also that L is differentiable with derivative $L'(\lambda) = O(L(\lambda)/\lambda)$ as $\lambda \to \infty$, for $0 < \eta < 1$,*

$$\frac{1}{n}\sum_{j=1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta}\cos t\lambda_j = O\left(L(t)t^{\eta-1}\right) \quad \text{as } n \to \infty, \tag{3.3}$$

uniformly in t such that $n/m < tn/2$.

PROOF. Choose s such that $s \sim n/t$, so that $s < m$ for n large enough. Now

$$\left|\frac{1}{n}\sum_{j=1}^{s} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta}\cos t\lambda_j\right| \le \frac{1}{n}\sum_{j=1}^{s} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} = O\left(L\left(\frac{1}{\lambda_s}\right)\lambda_s^{1-\eta}\right), \tag{3.4}$$

as $n \to \infty$, by Lemma 4. On the other hand, by summation by parts,

$$\begin{aligned}
&\frac{1}{n}\sum_{j=s+1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta}\cos t\lambda_j \\
&\quad = \frac{1}{n}\sum_{j=s+1}^{m-1}\left\{L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta} - L\left(\frac{1}{\lambda_{j+1}}\right)\lambda_{j+1}^{-\eta}\right\}\sum_{\ell=s+1}^{j}\cos t\lambda_\ell \\
&\qquad + \frac{1}{n}L\left(\frac{1}{\lambda_m}\right)\lambda_m^{-\eta}\sum_{j=s+1}^{m}\cos t\lambda_j.
\end{aligned}$$

Because $\sum_{j=a}^{b}\cos j\lambda = O(\lambda^{-1})$ uniformly in a and b for $0 < \lambda \le \pi$, this is

$$\begin{aligned}
&O\left(\frac{1}{t}\left[\sum_{j=s+1}^{m-1}\left\{\left|L\left(\frac{1}{\lambda_j}\right) - L\left(\frac{1}{\lambda_{j+1}}\right)\right|\lambda_j^{-\eta} + L\left(\frac{1}{\lambda_{j+1}}\right)\left|\lambda_j^{-\eta} - \lambda_{j+1}^{-\eta}\right|\right\}\right.\right. \\
&\qquad\qquad \left.\left. + L\left(\frac{1}{\lambda_m}\right)\lambda_m^{-\eta}\right]\right) \\
&\quad = O\left(\frac{1}{t}\left[\frac{1}{n}\sum_{j=s+1}^{m-1}\left\{\left|L'\left(\frac{1}{\theta_j}\right)\right|\theta_j^{-2}\lambda_j^{-\eta} + \theta_j^{-\eta-1}L\left(\frac{1}{\lambda_j}\right)\right\} + L\left(\frac{1}{\lambda_m}\right)\lambda_m^{-\eta}\right]\right),
\end{aligned}$$

as $n \to \infty$, by the mean value theorem, where $\lambda_j \le \theta_j \le \lambda_{j+1}$. This is

$$O\left(\frac{1}{t}\left\{\frac{1}{n}\sum_{j=s+1}^{m-1} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-\eta-1} + L\left(\frac{1}{\lambda_m}\right)\lambda_m^{-\eta}\right\}\right) = O\left(\frac{1}{t}L\left(\frac{1}{\lambda_s}\right)\lambda_s^{-\eta}\right),$$

as $n \to \infty$, using Lemma 1(i) and Lemma 6. Then (3.3) follows by substitution and (1.2). □

PROPOSITION 1. *As $n \to \infty$,*

$$\frac{2\pi}{n}\sum_{j=1}^{m} f(\lambda_j) - F(\lambda_m) = o(F(\lambda_m)) \quad \text{(under Conditions A and B)}$$

(3.5)
$$= O\left(\left\{m^{-\delta} + \left(\frac{m}{n}\right)^{\tau}\right\}F(\lambda_m)\right) \quad \text{(under Conditions A}' \text{ and B)},$$

for $\delta < 2(1-H)$.

REMARK. Here we could relax Condition A′ to

$$f(\lambda) - L\left(\frac{1}{\lambda}\right)\lambda^{1-2H} = O\left(L\left(\frac{1}{\lambda}\right)\lambda^{1-2H+\tau}\right) \quad \text{as } \lambda \to 0+.$$

PROOF OF PROPOSITION 1. For n sufficiently large, the LHS of (3.5) is dominated by

$$\left|\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}\left\{L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-2H} - L\left(\frac{1}{\lambda}\right)\lambda^{1-2H}\right\}d\lambda\right| + a,$$

where

$$a = o\left(\frac{1}{n}\sum_{j=1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-2H} + \int_0^{\lambda_m} L\left(\frac{1}{\lambda}\right)\lambda^{1-2H}\,d\lambda\right) \tag{3.6}$$

under Condition A, and

$$a = O\left(\frac{1}{n}\sum_{j=1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-2H+\tau} + \int_0^{\lambda_m} L\left(\frac{1}{\lambda}\right)\lambda^{1-2H+\tau}\,d\lambda\right) \tag{3.7}$$

under Condition A′. By Lemmas 3(ii) and 4 and by (2.3), (3.6) is $o(F(\lambda_m))$ whereas (3.7) is $O((m/n)^{\tau}F(\lambda_m))$. On the other hand,

$$\begin{aligned}
&\left|\int_0^{\lambda_1}\left\{L\left(\frac{1}{\lambda_1}\right)\lambda_1^{1-2H} - L\left(\frac{1}{\lambda}\right)\lambda^{1-2H}\right\}d\lambda\right| \\
&\quad \le L\left(\frac{1}{\lambda_1}\right)\lambda_1^{2(1-H)} + \int_0^{\lambda_1} L\left(\frac{1}{\lambda}\right)\lambda^{1-2H}\,d\lambda \\
&\quad = O(F(\lambda_1)) = O(m^{-\delta}F(\lambda_m)),
\end{aligned}$$

as $n \to \infty$, by Lemmas 3(ii) and 2, while

$$\sum_{j=2}^{m} L\left(\frac{1}{\lambda_j}\right) \int_{\lambda_{j-1}}^{\lambda_j} \left\{\lambda_j^{1-2H} - \lambda^{1-2H}\right\} d\lambda$$
$$= O\left(\sum_{j=2}^{m} L\left(\frac{1}{\lambda_j}\right) \lambda_j^{-2H} \int_{\lambda_{j-1}}^{\lambda_j} (\lambda_j - \lambda)\, d\lambda\right)$$
$$= O\left(\frac{1}{n^2} \sum_{j=1}^{m} L\left(\frac{1}{\lambda_j}\right) \lambda_j^{-2H}\right) = O(m^{-\delta} F(\lambda_m)),$$

by Lemma 5, while, from Seneta [(1976), page 7], we can choose L in Conditions A and A′ to be differentiable and such that $\lambda L'(\lambda)/L(\lambda) \to 0$ as $\lambda \to \infty$, so

$$\sum_{j=2}^{m} \int_{\lambda_{j-1}}^{\lambda_j} \lambda^{1-2H} \left\{L\left(\frac{1}{\lambda_j}\right) - L\left(\frac{1}{\lambda}\right)\right\} d\lambda = O\left(\frac{1}{n^2} \sum_{j=2}^{m} L\left(\frac{1}{\lambda_{j-1}}\right) \lambda_{j-1}^{-2H}\right)$$
$$= O(m^{-\delta} F(\lambda_m)),$$

by proceeding as before and in the proof of Lemma 6. □

PROPOSITION 2. *Under Conditions* A *and* B,

$$E\left\{\frac{2\pi}{n} \sum_{j=1}^{m} I(\lambda_j)\right\} = O(F(\lambda_m)) \quad \text{as } n \to \infty. \tag{3.8}$$

PROOF. Because x_t can be replaced by $x_t - \mu$ in the LHS of (3.8), by Herglotz's representation [Brillinger (1975), page 25] and periodicity of f and of Fejér's kernel

$$K(u) = \frac{1}{n} \left| \sum_{t=1}^{n} e^{itu} \right|^2,$$

the LHS of (3.8) is $S[-\pi, \pi]$, where

$$S[A_j] = \frac{1}{n} \sum_{j=1}^{m} \int_{A_j} K(u) f(u + \lambda_j)\, du.$$

(Of course $S[A_j]$ does not depend on j, with A_j referring to the range of integration in its jth summand). From, for example, Zygmund [(1977), page 90],

$$K(u) = O\left(\frac{n}{1 + n^2 u^2}\right), \qquad 0 < |u| \le \pi, \qquad \int_{-\pi}^{\pi} K(u)\, du = 2\pi. \tag{3.9}$$

For any $\varepsilon > 0$, (3.9) gives

$$S[(-\pi, -\varepsilon] \cup [\varepsilon, \pi)] = O\left(\frac{m}{n^2 \varepsilon^2} \int_{-\pi}^{\pi} f(\lambda)\, d\lambda\right) = O\left(\frac{m}{n^2}\right), \tag{3.10}$$

because x_t has finite variance. For small enough ϵ and n sufficiently large, (1.2), Condition A and Lemma 1(i) imply that, as $\lambda \to 0+, f(a\lambda) = O(f(\lambda)) = O(f(\theta))$, for all $a \neq 0$ and $\theta \in (0, \lambda)$. Applying also (3.9) and Condition B, as $n \to \infty$

$$S[(-\varepsilon, -2\lambda_m] \cup [2\lambda_m, \varepsilon)] = O\left(\frac{m}{n} f(\lambda_m) \int_{2\lambda_m}^{\pi} K(u)\, du\right) = O(n^{-1} f(\lambda_m)), \tag{3.11}$$

$$S\left[\left(-2\lambda_m, -\frac{1}{2}\lambda_j\right] \cup \left[\frac{1}{2}\lambda_j, 2\lambda_m\right)\right] = O\left(\frac{1}{n^2} \sum_{j=1}^{m} \frac{1}{\lambda_j^2} f(u + \lambda_j)\, du\right) = O\left(\lambda_m f(\lambda_m) \sum_{j=1}^{\infty} j^{-2}\right), \tag{3.12}$$

$$S\left[\left(-\frac{1}{2}\lambda_j, \frac{1}{2}\lambda_j\right)\right] = O\left(\frac{1}{n} \sum_{j=1}^{m} f\left(\frac{1}{2}\lambda_j\right) \int_{-\pi}^{\pi} K(u)\, du\right) = O\left(\frac{1}{n} \sum_{j=1}^{m} f(\lambda_j)\right). \tag{3.13}$$

By Conditions A and B and Proposition 1, (3.10)–(3.13) are each $O(F(\lambda_m))$. □

PROPOSITION 3. *As* $n \to \infty$,

$$\begin{aligned} &\frac{1}{n} \sum_{j=1}^{m} f(\lambda_j) \left\{ \frac{2\pi J(\lambda_j)}{\sigma^2} - 1 \right\} \\ &= o_p(F(\lambda_m)) \quad (\textit{under Conditions A, B and C}) \\ &= O_p\left(\left\{ n^{-\nu/(2+\nu)} + m^{(H-1)/(5-4H)} + \left(\frac{m}{n}\right)^{\tau} \right\} F(\lambda_m)\right), \end{aligned} \tag{3.14}$$

under Conditions A′, B *and* C′.

PROOF. The LHS of (3.14) is

$$\left\{ \frac{1}{n} \sum_{j=1}^{m} f(\lambda_j) \right\} \frac{1}{n\sigma^2} \sum_{t=1}^{n} \left(e_t^2 - \sigma^2\right) + \frac{2}{n\sigma^2} \sum_{s<t} \sum e_t e_s \gamma_{t-s,m}, \tag{3.15}$$

where

$$\gamma_{tm} = \frac{1}{n} \sum_{j=1}^{m} f(\lambda_j) \cos t\lambda_j.$$

Now

$$\frac{1}{n}\sum_{t=1}^{n}\left\{e_t^2 - E(e_t^2|e_s^2,\, s<t)\right\} = o_p(1), \qquad \text{under Condition C(iii)},$$

$$\frac{1}{n}\sum_{t=2}^{n}\left\{e_t^2 - E\left(e_t^2\middle|\sum_{s=1}^{t-1}e_s^2\right)\right\} = O_p(n^{-\nu/(2+\nu)}), \quad \text{under Condition C}'\text{(iii)}'.$$

applying, respectively, Heyde and Seneta [(1972), Theorem 1] and von Bahr and Esseen [(1965), Theorem 2], so the same results hold for $n^{-1}\Sigma_{t=1}^{n}(e_t^2-\sigma^2)$ on applying Condition C(iv) or C′(iv)′. Proposition 1 then implies that the first term of (3.15) is, respectively, $o_p(F(\lambda_m))$ and $O_p(n^{-\nu/(2+\nu)}F(\lambda_m))$. Under Condition C(ii), the second term of (3.15) has variance that is

$$O\left(\frac{1}{n}\sum_{t=1}^{n-1}\gamma_{tm}^2\right) = O\left(\frac{r}{n}F^2(\lambda_m) + \max_{r<t\le n/2}\gamma_{tm}^2\right),$$

for $n/m < r \le n/2$, where the first term on the right follows from $\gamma_{tm} = O(F(\lambda_m))$, due to Proposition 1. On the other hand, under Condition A,

$$|\gamma_{tm}| \le \left|\frac{1}{n}\sum_{j=1}^{m}L\left(\frac{1}{\lambda_j}\right)\lambda_j^{1-2H}\cos t\lambda_j\right| + a, \tag{3.16}$$

where, as in the proof of Proposition 1, $a = o(F(\lambda_m))$ under Condition A and $O((m/n)^{\tau}F(\lambda_m))$ under Condition A′. Lemma 7 bounds the first term on the RHS of (3.16) by $O(L(t)t^{2(H-1)})$ as $n\to\infty$, and, by Lemma 1(ii),

$$\max_{r<t\le n/2}|\gamma_{tm}| = O(L(r)r^{2(H-1)}) \quad \text{as } r\to\infty.$$

Now pick $r \sim nm^{4(H-1)/(5-4H)}/2\pi$, so that $n\to\infty$ and Condition B imply $r > 1/\lambda_m \to \infty$. Then the preceding calculations give

$$\frac{1}{n}\sum_{t=1}^{n-1}\gamma_{tm}^2 = O\left(m^{4(H-1)/(5-4H)}\left\{L^2\left(\frac{1}{\lambda_m}\right) + L^2(r)\right\}\lambda_m^{4(1-H)}\right)$$

and Lemma 1(ii) and Condition A imply this is $O(m^{2(H-1)/(5-4H)}F^2(\lambda_m)) = o(F^2(\lambda_m))$. □

PROPOSITION 4. *As* $n\to\infty$,

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{m}|w(\lambda_j) - \alpha(\lambda_j)v(\lambda_j)|^2 \\
&= o_p(F(\lambda_m)) \qquad (\textit{under Conditions } \text{A, B } \textit{and } \text{C}) \\
&= O_p\left(\left\{m^{-\delta} + \left(\frac{m}{n}\right)^{2\tau}\right\}F(\lambda_m)\right) \quad (\textit{under Conditions } \text{A}', \text{B}' \textit{ and } \text{C}'),
\end{aligned} \tag{3.17}$$

for $\delta < (1-H)/(2-H)$.

PROOF. The LHS of (3.17) is nonnegative so it will suffice to show that its expectation has orders corresponding to those on the RHS. The expectation is

$$(3.18)\quad \frac{1}{2\pi n^2}\sum_{j=1}^{m}\sum_{s,t=1}^{n}\exp\left[i(s-t)\lambda_j\right]\left\{\mathrm{Cov}(x_s,x_t)-\alpha(\lambda_j)E(x_se_t) - \overline{\alpha}(\lambda_j)E(e_sx_t)+\left|\alpha(\lambda_j)\right|^2E(e_se_t)\right\}.$$

By periodicity, for any λ,

$$\begin{aligned}\mathrm{Cov}(x_s,x_t) &= \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}|\alpha(u)|^2\exp[i(t-s)u]\,du\\ &= \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}|\alpha(u+\lambda)|^2\exp[i(t-s)(u+\lambda)]\,du,\\ E(x_se_t) &= \sigma^2\alpha_{s-t} = \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\alpha(\lambda)\exp[i(t-s)u]\,du\\ &= \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\alpha(u+\lambda)\exp[i(t-s)(u+\lambda)]\,du,\end{aligned}$$

so (3.18) is $S_m[(-\pi,\pi)]$, where, for any set A and integer r,

$$S_r[A] = \frac{\sigma^2}{(2\pi)^2 n}\sum_{j=1}^{r}\int_A K(u)|\alpha(\lambda_j)-\alpha(\lambda_j+u)|^2\,du.$$

For any A,

$$S_r[A] = O\left(\frac{1}{n}\sum_{j=1}^{r}f(\lambda_j)\int_A K(u)\,du + \frac{1}{n}\sum_{j=1}^{r}\int_A K(u)f(\lambda_j+u)\,du\right).$$

Thus, for any $\varepsilon > 0$,

$$(3.19)\quad \begin{aligned}S_m\left[(-\pi,-\varepsilon]\cup[\epsilon,\pi)\right] &= O\left(\frac{1}{n^2\varepsilon}\sum_{j=1}^{m}f(\lambda_j) + \frac{1}{n^2\varepsilon^2}\sum_{j=1}^{m}\int_{-\pi}^{\pi}f(\lambda_j+u)\,du\right)\\ &= O\left(n^{-1}F(\lambda_m)+n^{-2}m\right) = O(n^{-1}F(\lambda_m)),\end{aligned}$$

as $n\to\infty$, using (3.9), the integrability of f and Proposition 1. For small enough ε,

$$(3.20)\quad \begin{aligned}S_m\left[(-\varepsilon,-2\lambda_m]\cup[2\lambda_m,\varepsilon)\right] &= O\left(\frac{1}{n^2\lambda_m}\sum_{j=1}^{m}f(\lambda_j)+\frac{1}{n}f(\lambda_m)\right)\\ &= O(m^{-1}F(\lambda_m)),\end{aligned}$$

as $n \to \infty$, using (3.11). For a sequence ε_n such that $0 < \varepsilon_n < 1, 2\varepsilon_n m$ is an integer, and $\varepsilon_n \to 0$, $\varepsilon_n m \to \infty$ as $n \to \infty$, $S_m\left[(-2\lambda_m, -\varepsilon_n\lambda_m] \cup [\varepsilon_n\lambda_m, 2\lambda_m)\right]$ is

$$O\left(\frac{1}{n^2\varepsilon_n\lambda_m}\sum_{j=1}^{m}\left\{f(\lambda_j) + \frac{1}{\varepsilon_n\lambda_m}\int_{-2\lambda_m}^{2\lambda_m} f(u+\lambda_j)\,du\right\}\right)$$

$$(3.21) \qquad = O\left(\frac{1}{\varepsilon_n^2 m}F(\lambda_m)\right) \quad \text{as } n \to \infty.$$

Proceeding as in (3.12) and (3.13), $S_{2\varepsilon_n m}\left[(-\varepsilon_n\lambda_m, \varepsilon_n\lambda_m)\right]$ is

$$O\left(\frac{1}{n}\sum_{j=1}^{2\varepsilon_n m}\left\{f(\lambda_j) + \int_{\lambda_j/4<|u|\le\varepsilon_n\lambda_m} K(u)f(u+\lambda_j)\,du\right.\right.$$

$$(3.22) \qquad \left.\left. + \int_{-\lambda_j/4}^{\lambda_j/4} K(u)f(u+\lambda_j)\,du\right\}\right)$$

$$= O\left(F(\varepsilon_n\lambda_m) + \sum_{j=1}^{2\varepsilon_n m}\left\{\frac{1}{j^2}\int_{-\varepsilon_n\lambda_m}^{\varepsilon_n\lambda_m} f(u+\lambda_j)\,du + \frac{1}{n}f\left(\frac{3}{4}\lambda_j\right)\right\}\right)$$

$$= O(F(\varepsilon_n\lambda_m)),$$

as $n \to \infty$. It remains to consider

$$(3.23) \qquad \frac{1}{n}\sum_{j=2\varepsilon_n m+1}^{m}\int_{-\varepsilon_n\lambda_m}^{\varepsilon_n\lambda_m} K(u)\mid\alpha(\lambda_j) - \alpha(\lambda_j+u)\mid^2 du.$$

Now $|\alpha(\lambda) - \alpha(\lambda+u)|^2 = \left\{\alpha_1(\lambda) - \alpha_1(\lambda+u)\right\}^2 + \left\{\alpha_2(\lambda) - \alpha_2(\lambda+u)\right\}^2$. For i satisfying (2.5), (2.6), (2.11) or (2.12), as $\lambda, |\lambda+u| \to 0$,

$$\left\{\alpha_i(\lambda) - \alpha_i(\lambda+u)\right\}^2 = O(|L^{1/2}(\lambda^{-1})\lambda^{1/2-H} - L^{1/2}(|\lambda+u|^{-1})|\lambda+u|^{1/2-H}|^2 + a^2),$$

where, under (2.5) or (2.6), $a = o(L^{1/2}(\lambda^{-1})\lambda^{1/2-H} + L^{1/2}(|\lambda+u|^{-1})|\lambda+u|^{1/2-H})$, and under (2.11) or (2.12), $a = O(L^{1/2}(\lambda^{-1})\lambda^{1/2-H+\tau} + L^{1/2}(|\lambda + u|^{-1})|\lambda + u|^{1/2-H+\tau})$. The contribution of the a^2 term to (3.23) is thus $o(F(\lambda_m))$ under Condition A, and $O((m/n)^{2\tau}F(\lambda_m))$ under Condition A′, in view of (2.3). The contribution of $L(\lambda_j^{-1})(\lambda_j^{1/2-H} - |\lambda_j+u|^{1/2-H})^2$ to (3.23) is

$$O\left(\frac{1}{n}\int_{-\varepsilon_n\lambda_m}^{\varepsilon_n\lambda_m} u^2K(u)\,du\sum_{2\varepsilon_n m+1}^{m} L\left(\frac{1}{\lambda_j}\right)\lambda_j^{-2H-1}\right)$$

$$= O\left(\frac{1}{n}L\left(\frac{1}{\varepsilon_n\lambda_m}\right)(\varepsilon_n\lambda_m)^{1-2H}\right) = O\left(\frac{F(\varepsilon_n\lambda_m)}{\varepsilon_n\lambda_m}\right),$$

as $n \to \infty$, by Lemma 6 and (3.9). Likewise the contribution of

$$\{L^{1/2}(\lambda_j^{-1}) - L^{1/2}(|\lambda_j+u|^{-1})\}(\lambda_j+u)^{1-2H}$$

to (3.23) is

$$O\left(\frac{1}{n^2}\int_{-\varepsilon_n\lambda_m}^{\varepsilon_n\lambda_m}\sum_{2\varepsilon_n m+1}^{m} L\left(\frac{1}{\delta_{ju}}\right)\delta_{ju}^{-2}|\lambda_j+u|^{1-2H}\,du\right) \quad \text{as } n\to\infty,$$

for δ_{ju} between λ_j and λ_j+u, arguing as in the proof of Lemma 7. Now $\delta_{ju}\geq\frac{1}{2}\lambda_j$ for relevant j and u, so by Lemma 6 the last displayed expression is

$$O\left(\frac{L[1/\varepsilon_n\lambda_m]}{\varepsilon_n\lambda_m}(\varepsilon_n\lambda_m)^{1-2H}\frac{\varepsilon_n\lambda_m}{n}\right)=O\left(\frac{F(\varepsilon_n\lambda_m)}{\varepsilon_n m}\right) \quad \text{as } n\to\infty. \tag{3.24}$$

Now take $\varepsilon_n = m^{-1/2(2-H)}$, so that (3.19)–(3.22) and (3.24) are all $O(m^{-\delta}F(\lambda_m))$ for $\delta < (1-H)(2-H)$, applying Lemma 2. It remains only to consider the contribution of the term in $\alpha_j(\lambda)-\alpha_j(\lambda+u)$ under (2.7) and (2.13), and (2.3) implies this is $o(F(\lambda_m))$ and $O((m/n)^{2\tau}F(\lambda_m))$, respectively. □

4. Estimation of H and G. The models (1.4) and (when the orders of a and b are given) (1.5) are two of many possible parametric models for $f(\lambda)$. Gaussian estimates for long-memory parametric models have now been rigorously justified, by Fox and Taqqu (1986) and Dahlhaus (1989) under Gaussianity, and by Solo (1989) and Giraitis and Surgailis (1990) under linearity. When the parametric form of $f(\lambda)$ is misspecified, in general these estimates will be inconsistent. There is therefore interest in estimating an unknown H in the semiparametric model (1.1), which is agnostic about short- and medium-run behavior. A closed-form semiparametric estimate also has some computational advantage over Gaussian estimates of H in models such as (1.4) and (1.5), which have to be obtained by numerical methods. One application of a semiparametric estimate of H is in the estimation of the fractional ARIMA model (1.5) when the degrees of a and b are unknown: The x_t can be approximately fractionally differenced using the semiparametric estimate of H and the binomial theorem, and then standard order determination methods can be employed.

Two closed-form semiparametric estimates of H were proposed by Janacek (1982) and Geweke and Porter-Hudak (1983), assuming L in (1.1) is constant. Janacek employed a certain parametric f in which the parameter space increases slowly with n, but the justification he provided was heuristic and, in particular, no consistency proof was given. Geweke and Porter-Hudak regressed $\log I(\lambda_j)$ on $-2\,\log|1-\exp(i\lambda_j)|$, where $j=1,\dots,m$, where m satisfies at least Condition B, but discussed asymptotic properties only in case $0<H<\frac{1}{2}$ [when $f(0)=0$] and even here their discussion was heuristic. Recently, Robinson (1992) has established asymptotic properties for a modified version of Geweke and Porter-Hudak's estimate proposed by Künsch (1986), which omits the contributions for the very lowest λ_j, allowing $0<H<1$ but assuming L is constant and x_t is Gaussian.

We present a new semiparametric estimate of H which we show to be consistent under the mild assumptions of Theorem 1, and even when the slowly varying function L in (1.1) is of unknown form. To motivate the estimate, observe that (1.2) and (2.3) imply, for any $q > 0$,

$$\frac{F(q\lambda)}{F(\lambda)} \sim q^{2(1-H)} \frac{L(1/q\lambda)}{L(1/\lambda)} \sim q^{2(1-H)}, \tag{4.1}$$

as $\lambda \to 0+$, suggesting the estimate

$$\widehat{H}_{mq} = 1 - \frac{\log\left\{\widehat{F}(q\lambda_m)/\widehat{F}(\lambda_m)\right\}}{2\log q}. \tag{4.2}$$

Because $\widehat{H}_{mq} \equiv \widehat{H}_{mq,1/q}$ we can restrict q to the interval (0, 1). $\widehat{H}_{mq}$ nearly always lies in the stationarity region $(-\infty, 1)$; it cannot exceed 1, and it equals 1 only if $\widehat{F}(q\lambda_m) = \widehat{F}(\lambda_m)$. $\widehat{H}_{mq}$ is scale- and location-invariant due to the ratio in the numerator of (4.2) and the invariance of $\widehat{F}$ to location.

THEOREM 3. *Under Conditions* A, B *and* C, *for any* $q \in (0,1)$,

$$\widehat{H}_{mq} \to_p H \quad \text{as } n \to \infty.$$

PROOF. We have

$$\begin{aligned} &\widehat{H}_{mq} - \left[1 - \frac{\log\left\{F(q\lambda_m)/F(\lambda_m)\right\}}{2\log q}\right] \\ &\quad = \frac{1}{2\log q}\left[\log\left\{\frac{\widehat{F}(\lambda_m)}{F(\lambda_m)}\right\} - \log\left\{\frac{\widehat{F}(q\lambda_m)}{F(q\lambda_m)}\right\}\right], \end{aligned} \tag{4.3}$$

which is $o_p(1)$ as $n \to \infty$, by Slutzky's theorem and Theorem 1. Now apply (4.1) and Condition B. □

For the purposes of the following theorem and the following section, we find it useful to establish also a rate of convergence. We introduce the following condition.

CONDITION B′. For some φ and ψ such that $0 < \varphi < \psi < 1$,

$$\frac{n^{\varphi}}{m} + \frac{m}{n^{\psi}} \to 0 \quad \text{as } n \to \infty.$$

Condition B′ is true if, for example, $m \sim \Gamma n^{\gamma}$ for a finite positive constant Γ and $\gamma \in (0,1)$.

THEOREM 4. *Under Conditions* A′, B′ *and* C′, *for some* $\delta > 0$,

$$\widehat{H}_{mq} - H = O_p(n^{-\delta}) \quad \text{as } n \to \infty.$$

PROOF. Combining Condition B′ with Theorem 2 gives $\widehat{F}(\lambda_m)/F(\lambda_m) - 1 = O_p(n^{-\delta})$, so that (4.3) is $O_p(n^{-\delta})$. Condition A′ implies that $f(\lambda) - L(\lambda^{-1})\lambda^{1-2H} = O(L(\lambda^{-1})\lambda^{1-2H+\tau})$, as $\lambda \to 0+$, thus

$$F(\lambda) - L(\lambda^{-1})\lambda^{2(1-H)}/2(1-H) = O\big(L(\lambda^{-1})\lambda^{2(1-H)+\tau}\big) \quad \text{as } \lambda \to 0+,$$

and thus

$$1 - \frac{\log\{F(q\lambda_m)/F(\lambda_m)\}}{2\log q} - H = O\left(\left(\frac{m}{n}\right)^{\tau}\right) = O(n^{-\delta}) \quad \text{as } n \to \infty. \qquad \square$$

Theorems 3 and 4 require no knowledge of the functional form of $L(\lambda)$. However under (1.3) we can consistently estimate G. Because (2.4) implies

$$\text{(4.4)} \qquad \frac{2(1-H)F(\lambda)}{M(1/\lambda)\lambda^{2(1-H)}} \sim G \quad \text{as } \lambda \to 0+,$$

we estimate G by

$$\text{(4.5)} \qquad \widehat{G}_{mg} = \frac{2(1-\widehat{H}_{mq})\widehat{F}(\lambda_m)\lambda_m^{2(\widehat{H}_{mq}-1)}}{M(1/\lambda_m)}.$$

Note that $\widehat{G}_{mq} \geq 0$.

THEOREM 5. *Under* (1.3) *and Conditions* A′, B′ *and* C′, *for any* $q \in (0,1)$,

$$\widehat{G}_{mq} \to_p G \quad \text{as } n \to \infty.$$

PROOF. From (4.4), $\widehat{G}_{mq}/G$ has the same probability limit as

$$\widehat{G}_{mq}\frac{M(1/\lambda_m)\lambda_m^{2(1-H)}}{2(1-H)F(\lambda_m)} = \left\{\frac{1-\widehat{H}_{mq}}{1-H}\frac{\widehat{F}(\lambda_m)}{F(\lambda_m)}\right\}\lambda_m^{2(\widehat{H}_{mq}-H)}.$$

The factor in braces converges to 1 in probability, by Theorems 1 and 3, and by Theorem 4, as $n \to \infty$

$$\begin{aligned}\left|\lambda_m^{2(\widehat{H}_{mq}-H)} - 1\right| &= \left|\exp\left\{2(\widehat{H}_{mq} - H)\log \lambda_m\right\} - 1\right| \\ &\leq 4(\log n)|\widehat{H}_{mq} - H| \to_p 0. \qquad \square\end{aligned}$$

We now report a small Monte Carlo investigation of finite-sample properties. Using the algorithm of Davies and Harte (1987), 30,000 Gaussian series of length $n = 64$ were generated from model (1.4) with $\sigma^2 = 1$, for each of five values of H, $H = 0.55(0.1)0.95$. For each of the $5 \times 30{,}000$ series, estimates $\widehat{H}_{mq}$ and $\widehat{G}_{mq}$ of H and $G = \Gamma(2H+1)\sin(\pi H)/2\pi$ [with the convention that

TABLE 1

Bias and MSE of $\hat{H}_{mq}$ in (1.4) for $n = 64$

	m						
H	6	10	14	18	22	26	30
0.55	−0.079 0.109	−0.047 0.057	−0.032 0.038	−0.022 0.028	−0.015 0.022	−0.012 0.018	−0.009 0.016
0.65	−0.107 0.097	−0.067 0.048	−0.045 0.031	−0.031 0.022	−0.022 0.017	−0.015 0.014	−0.012 0.012
0.75	−0.144 0.092	−0.096 0.043	−0.070 0.027	−0.052 0.018	−0.039 0.013	−0.030 0.010	−0.025 0.009
0.85	−0.185 0.091	−0.133 0.044	−0.105 0.027	−0.086 0.019	0.070 0.013	0.060 0.010	0.053 0.008
0.95	−0.232 0.100	−0.175 0.050	−0.147 0.033	−0.128 0.024	−0.113 0.018	−0.102 0.015	−0.095 0.012

the constant $M(\lambda)$ in (1.3) is 1] were calculated with $q = \frac{1}{2}$ and 7 values of $m, m = 6(4)30$. The results for $\hat{H}_{m,1/2}$ and $\hat{G}_{m,1/2}$ are presented in Tables 1 and 2, respectively. In each cell the top number is the Monte Carlo bias, the bottom one the Monte Carlo mean square error (MSE). One might not expect good results for so small a sample size, and the estimates are seriously biased. $\hat{H}_{m,1/2}$ always underestimates H, the bias increasing in H and decreasing in m [(1.4) decays monotonically but with real data there is a danger in choosing m too large in case of a peak in the spectrum at a small nonzero frequency]. The MSE of $\hat{H}_{m,1/2}$ is U-shaped in H but unsurprisingly decreases in m. $\hat{G}_{m,1/2}$ overestimates for the smallest two values of m but otherwise underestimates, the smallest biases being recorded for $m = 10$ and 14. The MSE of $\hat{G}_{m,1/2}$ decreases in both m and H.

Graf (1983) proposed an interesting form of robustified M-estimate of H based on a parametric model for f and applied various versions of it to a time series of annual minimum water levels of the River Nile measured at the Roda Gorge near Cairo during the years 622 through 1284 [see Toussoun (1925)], obtaining estimates between 0.828 and 0.847. We computed $\hat{H}_{m,1/2}$ for $m = 20(20)180$. The estimates were very stable in m, lying between 0.832 and 0.859.

5. Further applications.

5.1. *Autocorrelation-consistent standard errors.* A long-used model for time series y_t, observed at $t = 1, \dots, n$, is the polynomial regression

$$y_t = \sum_{j=1}^{p} \beta_j t^{j-1} + x_t, \tag{5.1}$$

TABLE 2

Bias and MSE of $\widehat{G}_{mq}$ in (1.4) for $n = 64$

	m						
H	6	10	14	18	22	26	30
0.55	0.104	0.011	−0.004	−0.009	−0.011	−0.011	−0.011
	0.189	0.011	0.004	0.002	0.001	0.001	0.001
0.65	0.099	0.008	−0.010	−0.016	−0.019	−0.021	−0.021
	0.163	0.010	0.004	0.002	0.002	0.001	0.001
0.75	0.092	0.007	−0.011	−0.019	−0.024	−0.026	−0.027
	0.128	0.008	0.003	0.002	0.002	0.001	0.001
0.85	0.072	0.007	−0.009	−0.017	−0.022	−0.025	−0.026
	0.066	0.004	0.002	0.001	0.001	0.001	0.001
0.95	0.033	0.004	−0.004	−0.008	−0.010	−0.012	−0.013
	0.010	0.001	0.000	0.000	0.000	0.000	0.000

where x_t is an unobservable covariance stationary process, with $E(x_t) = 0$ and spectral density $f(\lambda)$, and $\beta = (\beta_1, \dots, \beta_p)'$ is unknown. For $f(\lambda)$ continuous and positive at $\lambda = 0$, Grenander (1954) showed that the ordinary least square (OLS) estimate of β is asymptotically as efficient as generalized least squares (GLS). Under (1.1), and with $p = 1$ in (5.1), results of Adenstedt (1974), Beran and Künsch (1985), Samarov and Taqqu (1988) and Yajima (1988) indicate that the efficiency of OLS may still be very good. However, standard errors with the appropriate consistency properties in the presence of unknown H must be attached, to ensure consistency of interval estimates and asymptotic validity of test statistics. Although Beran (1989) employed robustified M-type estimates of H based on a parametric long-memory time series model in constructing a modified t-ratio when $p = 1$ in (5.1), semiparametric estimates of H are especially suited because it is only behavior of f near zero frequency that is relevant and only a slow rate of convergence of the H estimate is required. For general p in (5.1), define

$$\widehat{\beta} = \left(\sum_{t=1}^{n} z_t z_t'\right)^{-1} \sum_{t=1}^{n} z_t y_t \quad \text{where } z_t = \left(1, t, \dots, t^{-1}\right)',$$

and

$$\widehat{x}_t = x_t - (\widehat{\beta} - \beta)' z_t, \quad \widehat{w}(\lambda) = (2\pi n)^{-1} \sum_{t=1}^{n} \widehat{x}_t e^{it\lambda}.$$

Denote by $\widetilde{H}_{mq}$ and $\widetilde{G}_{mq}$ the estimates $\widehat{H}_{mq}$ and $\widehat{G}_{mq}$ with the x_t replaced by the $\widehat{x}_t$, where (1.3) is assumed. Introduce the $p \times p$ matrices

$$D = \operatorname{diag}\left\{ n^{1/2}, \left(\sum_1^n t^2\right)^{1/2}, \ldots, \left(\sum_1^n t^{2(p-1)}\right)^{1/2} \right\},$$

$$Q = \{(2i-1)^{1/2}(2j-1)^{1/2}(i+j-1)^{-1}\},$$

$$R(H) = \Big\{ 2\Gamma(2(1-H))\cos((1-H)\pi) \times (2i-1)^{1/2}(2j-1)^{1/2} \int_0^1 \int_0^1 x^{i-1}y^{j-1}|x-y|^{2(H-1)}\,dx\,dy \Big\},$$

where the (i,j)-th elements of Q and $R(H)$ are indicated. Then we can show that

$$\text{(5.2)} \qquad \frac{n^{1/2-\widetilde{H}_{mq}}}{\widetilde{G}_{mq}^{1/2} M^{1/2}(n)} R^{-1/2}(\widetilde{H}_{mq}) QD(\widehat{\beta}-\beta) \to_d N(0, I_p) \quad \text{as } n \to \infty,$$

under Conditions A′, B′ and C′ and conditions ensuring that (5.2) holds with $\widetilde{G}_{mq}$ and $\widetilde{H}_{mq}$ replaced by G and H [see Yajima (1988), (1991)]. To establish (5.2) it suffices to show that

$$\text{(5.3)} \qquad (\log n)(\widetilde{H}_{mq} - H) \to_p 0, \qquad \widetilde{G}_{mq} \to_p G \quad \text{as } n \to \infty,$$

$$\text{(5.4)} \qquad R(H) \text{ is continuous in } H \text{ on } \left(\frac{1}{2}, 1\right),$$

where the convergence rate for $\widetilde{H}_{mq}$ in (5.3) is due to the requirement that $n^{\widetilde{H}_{mq}-H} \to_p 1$. Theorems 4 and 5 imply that $(\log n)\{\widehat{H}_{mq} - H\} \to_p 0$, $\widehat{G}_{mq} \to_p G$, as $n \to \infty$, and their proofs indicate that (5.3) will follow if, for some $\varepsilon > 0$,

$$n^{-1} \sum_{j=1}^m \{\widehat{I}(\lambda_j) - I(\lambda_j)\} = O_p(n^{-\varepsilon} F(\lambda_m)) \quad \text{as } n \to \infty,$$

where $\widehat{I}(\lambda) = |\widehat{w}(\lambda)|^2$. Now

$$\text{(5.5)} \qquad \frac{1}{n}\left| \sum_{j=1}^m \left\{ I(\lambda_j) - \widehat{I}(\lambda_j) \right\} \right| \le \frac{1}{n} \left[2 \sum_{j=1}^n |w(\lambda_j) - \widehat{w}(\lambda_j)|^2 \sum_{j=1}^m \left\{ I(\lambda_j) + \widehat{I}(\lambda_j) \right\} \right]^{1/2},$$

and the first sum in the square brackets is straightforwardly seen to equal

$$\frac{1}{2\pi} \operatorname{tr}\left\{ \sum_{t=1}^n z_t z_t' (\widehat{\beta}-\beta)(\widehat{\beta}-\beta)' \right\} = O_p\left(\operatorname{tr}\left\{ D^{-1} \sum_{t=1}^n z_t z_t' D^{-1} \right\} n^{2H-1} L(n) \right)$$
$$= O_p\left(n^{2H-1} L(n) \right),$$

as $n \to \infty$. By the triangle inequality and Proposition 2, the left hand side of (5.5) is

$$O_p\Big(\{n^{2H-2}L(n)F(\lambda_m)\}^{1/2}\Big) = O_p\left(m^{H-1}\frac{L(n)^{1/2}}{L\big(\lambda_m^{-1}\big)^{1/2}}F(\lambda_m)\right) = O_p(n^{-\varepsilon}F(\lambda_m)),$$

using Condition B′ and $L(\lambda) + L(\lambda)^{-1} = O(\lambda^{\eta})$ as $\lambda \to \infty$ for any $\eta > 0$. To check (5.4), note that $\Gamma(2(1-H))$ and $\cos((1-H)\pi)$ are continuous in H, and, for $\delta > 0$,

$$\left|\int_0^1\int_0^1 x^{i-1}y^{j-1}|x-y|^{2(H-1)}\,dx\,dy - \int_0^1\int_0^1 x^{i-1}y^{j-1}|x-y|^{2(H-1)+\delta}\,dx\,dy\right|$$
$$\le \int_0^1\int_0^1\Big(|x-y|^{2(H-1)} - |x-y|^{2(H-1)+\delta}\Big)\,dx\,dy$$

because x^a decreases in a, for $0 < x < 1$. However, $\int_0^1\int_0^1 |x-y|^{2(H-1)}\,dx\,dy = 1/H(2H-1)$ is continuous in H on $(\frac{1}{2}, 1)$.

5.2. *Feasible generalized least squares estimates.* Assuming (5.1) and (1.1) with L constant, Dahlhaus (1992) has suggested weighted least squares estimates of β which he showed to have the same asymptotic efficiency as GLS. The weights depend only on H, and Dahlhaus also showed that inserting an n^{ρ}-consistent estimate, for any $\rho > 0$, maintains the efficiency; the preceding discussion indicates that $\widetilde{H}_{mq}$ satisfies this requirement under Conditions A′, B′, and C′.

5.3. *Estimation of cointegrating coefficients.* Model (5.1) implies that a certain linear combination of a nonstationary dependent variable and certain nonstochastic nonstationary regressors is stationary. When the nonstationary regressors are stochastic, OLS is often consistent even when the errors have nonzero mean or are not orthogonal to the regressors, so the regression is incompletely specified. Consider the equation

$$y_t = \beta z_t + x_t, \tag{5.6}$$

linking the scalar stochastic variates x_t, y_t and z_t, where only y_t and z_t are observed, at $t = 1,\dots,n$. As $n \to \infty$, one sufficient condition that OLS $\widehat{\beta} = \sum_{t=1}^n y_t z_t / \sum_{t=1}^n z_t^2 \to_p \beta$ is $\sum_{t=1}^n x_t^2 / \sum_{t=1}^n z_t^2 \to_p 0$ (by the Cauchy inequality), which indicates, along with the "cointegrating" relation (5.6), that there is a linear combination of y_t and z_t which is stationary or "less nonstationary" than y_t and z_t individually and holds in the special case, stressed in much recent econometric literature [e.g., Johansen (1988)], in which z_t has a unit root and x_t is stationary with bounded spectrum, for example, y_t and z_t represent consumption and income. The insistence on a unit root—weakly dependent distinction between z_t and x_t is somewhat arbitrary and one can also consider a notion of cointegration when z_t is long-memory stationary while x_t is stationary with shorter memory, so one can think of an "equilibrium" relation $y_t = \beta z_t$ on low frequencies only. The problem of consistently estimating β is then more challenging, because $\hat{\beta}$ can be consistent only for $\beta + E(z_t x_t)/E(z_t^2)$, and OLS with an intercept can be consistent only for $\beta + \mathrm{Cov}(z_t, x_t)/\mathrm{Var}(z_t)$, neither of which equal β when x_t and z_t are not orthogonal. We suggest a more delicate approach, carrying out the regression in the frequency domain over only a degenerating band of low frequencies. Define $w_y(\lambda) = (2\pi n)^{-1/2}\sum_{t=1}^{n} y_t e^{it\lambda}$, $w_z(\lambda) = (2\pi n)^{-1/2}\sum_{t=1}^{n} z_t e^{it\lambda}$, $I_{yz}(\lambda) = w_y(\lambda)\overline{w}_z(\lambda)$, $I_z(\lambda) = |w_z(\lambda)|^2$, $\widehat{F}_{yz}(\lambda_m) = (2\pi/n)\sum_{j=1}^{m} I_{yz}(\lambda_j)$, $\widehat{F}_z(\lambda_m) = (2\pi/n)\sum_{j=1}^{m} I_z(\lambda_j)$, $\widetilde{\beta} = \mathrm{Re}\{\widehat{F}_{yz}(\lambda_m)\}/\widehat{F}_z(\lambda_m)$. Denote by $f_z(\lambda)$ the spectrum of z_t and assume that $f_z(\lambda) \sim N(\lambda^{-1})\lambda^{1-2J}$, as $\lambda \to 0+$, for $\frac{1}{2} < J < 1$ and $N(\lambda)$ a slowly varying function at infinity. Also assume that $z_t = E(z_t) + \sum_{j=0}^{\infty}\rho_j \epsilon_{t-j}$, $\sum_{j=0}^{\infty}\rho_j^2 < \infty$, where the white noise innovations ϵ_t have variance τ^2 and Condition C holds for the ϵ_t after replacing e by ϵ and σ by τ. Defining $\widehat{F}_{xz}$ analogously to $\widehat{F}_{yz}$, and with $\widehat{F}$ as in Section 2,

$$|\widetilde{\beta} - \beta| \leq \left|\frac{\mathrm{Re}\{\widehat{F}_{xz}(\lambda_m)\}}{\widehat{F}_z(\lambda_m)}\right| \leq \frac{|\widehat{F}_{xz}(\lambda_m)|}{\widehat{F}_z(\lambda_m)} \leq \left\{\frac{\widehat{F}(\lambda_m)}{\widehat{F}_z(\lambda_m)}\right\}^{1/2},$$

by the Cauchy inequality. Assuming Conditions A, B and C, but not assuming orthogonality between x_t and z_t or that $E(x_t) = 0$ or $E(z_t) = 0$, it follows from Theorem 1 that

$$\widetilde{\beta} \to_p \beta \quad \text{as } n \to \infty,$$

if $\frac{1}{2} < H < J$, or if $\frac{1}{2} < H = J$ and $L(\lambda)/N(\lambda) \to 1$ as $\lambda \to \infty$.

Acknowledgments. I am grateful to three referees and an Associate Editor for helpful comments, to Joris Pinkse for some help with the computations, and to Ignacio Lobato for a careful reading of the paper.

REFERENCES

ADENSTEDT, R. (1974). On large-sample estimation for the mean of a stationary random sequence. *Ann. Statist.* **2** 1095–1107.

BERAN, J. (1989). A test of location for data with slowly decaying serial correlations. *Biometrika* **76** 261–269.

BERAN, J. and KÜNSCH, H. R. (1985). Location estimators for processes with long-range dependence. Research Report 40, Seminar für Statistic, ETH Zürich.

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

DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766.

DAHLHAUS, R. (1992). Efficient location and regression estimation for long range dependent regression models. Unpublished manuscript.

DAVIES, R. B. and HARTE, D. S. (1987). Tests for Hurst effect. *Biometrika* **74** 95–101.

FOX, R. and TAQQU, M. S. (1985). Non-central limit theorems for quadratic forms in random variables having long-range dependence. *Ann. Probab.* **13** 428–446.

FOX, R. and TAQQU, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** 517–532.

GEWEKE, J. and PORTER-HUDAK, S. (1983). The estimation and application of long memory time series models. *J. Time Ser. Anal.* **4** 221–238.

GIRAITIS, L. and SURGAILIS, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probab. Theory Related Fields* **86** 87–104.

GRAF, H.-P. (1983). Long-range correlations and estimation of the self-similarity parameter. Dissertation, ETH Zürich.

GRENANDER, U. (1954). On the estimation of regression coefficients in the case of autocorrelated disturbance. *Ann. Math. Statist.* **25** 252–272.

HEYDE, C. C. and SENETA, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **9** 235–256.

IBRAGIMOV, I. A. (1963). On estimation of the spectral function of a stationary Gaussian process. *Theory Probab. Appl.* **8** 366–401.

JANACEK, G. J. (1982). Determining the degree of differencing for time series via the log spectrum. *J. Time Ser. Anal.* **3** 177–183.

JOHANSEN, S. (1988). Statistical analysis of cointegration vectors. *J. Econom. Dynamics Control* **12** 231–254.

KÜNSCH, H. R. (1986). Discrimination between monotonic trends and long-range dependence. *J. Appl. Probab.* **23** 1025–1030.

ROBINSON, P. M. (1991). Rates of convergence and optimal spectral bandwidth for long range dependence. *Probab. Theory Related Fields.* To appear.

ROBINSON, P. M. (1992). Log-periodogram regression for time series with long range dependence. Unpublished manuscript.

SAMAROV, A. and TAQQU, M. S. (1988). On the efficiency of the sample mean in long memory noise. *J. Time Ser. Anal.* **9** 191–200.

SENETA, E. (1976). *Regularly Varying Functions.* Springer, Berlin.

SINAI, YA. G. (1976). Self-similar probability distributions. *Theory Probab. Appl.* **21** 64–80.

SOLO, V. (1989). Asymptotics of quadratic forms for time series. Unpublished manuscript.

TOUSSOUN, O. (1925). Mémoire sur l'histoire du Nil. *Mémoires de l'Institut d'Egypte* **9**.

VON BAHR, B. and ESSEEN, C. G. (1965). Inequalities for the r-th absolute moments of a sum of random variables $1 \leq r \leq 2$. *Ann. Math. Statist.* **36** 299–303.

YAJIMA, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Ann. Statist.* **16** 791–807.

YAJIMA, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Ann. Statist.* **19** 158–177.

YONG, C. H. (1974). *Asymptotic Behavior of Trigonometric Series.* Chinese Univ., Hong Kong.

ZYGMUND, A. (1977). *Trigonometric Series.* Cambridge Univ. Press.

DEPARTMENT OF ECONOMICS
LONDON SCHOOL OF ECONOMICS
AND POLITICAL SCIENCE
HOUGHTON STREET
LONDON WC2A 2AE
ENGLAND