SEMIPARAMETRIC INFERENCE IN SEASONAL AND CYCLICAL LONG MEMORY PROCESSES[*]

by

Josu Arteche
University of the Basque Country, Bilbao

and

Peter M Robinson
London School of Economics and Political Science

Contents:

Discussion Paper
No. EM/98/359
September 1998

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

[*] Research supported by ESRC grant R000235892. The first author acknowledges financial support from the Bank of Spain and UPV grant 038.321-HB039/97, and the second author a Leverhulme Personal Research Professorship. The data were provided by the Bank of England. We thank an associate editor and a referee for their comments.

Abstract

Several semiparametric estimates of the memory parameter in standard long memory time series are now available. They consider only local behaviour of the spectrum near zero frequency, about which the spectrum is symmetric. However, long-range dependence can appear as a spectral pole at any Nyqvist frequency (reflecting seasonal or cyclical long memory), where the spectrym need display no such symmetry. We introduce Seasonal/Cyclical Asymmetric Long Memory (SCALM) processes that allow differing rates of increase on either side of such a pole. To estimate the two consequent memory parameters we extend two semiparametric methods that were proposed for the standard case of a spectrum diverging at the origin, namely the log-periodogram and Gaussian or Whittle methods. We also provide three tests of symmetry. Monte Carlo analysis of finite sample behaviour and an empirical application to UK inflation data are included. Our models and methods allow also for the posssibility of negative dependence, described by a possibly asymmetric spectral zero.

Keywords: Semiparametric inference; long memory; seasonality.
JEL No.: C32

1 INTRODUCTION

In the analysis of time series the behaviour of the spectral density (spectrum) around zero frequency has attracted great interest, in particular with respect to the possibility of a spectral pole or zero. For a scalar covariance stationary process, x_t, $t = 0, \pm 1, \pm 2, ...$, assume absolute continuity of the spectral distribution function so that there exists a spectrum $f(\lambda)$, and lag-j autocovariance γ_j, satisfying

$$\gamma_j = E\{(x_t - Ex_t)(x_{t+j} - Ex_t)\} = \int_{-\pi}^{\pi} f(\lambda)\cos(j\lambda)\mathrm{d}\lambda. \tag{1.1}$$

A standard semiparametric model for local behaviour close to zero is $f(\lambda) \sim C|\lambda|^{-2d}$ as $\lambda \to 0$, for $0 < C < \infty$, where the memory parameter d satisfies $|d| < 1/2$, where $d < 1/2$ is inherent in stationarity and $d > -1/2$ ensures invertibility. When $d = 0$, x_t is often said to have short memory, when $d > 0$ long memory, and when $d < 0$ antipersistence or negative dependence. Various methods of estimating d are now available. For reviews see e.g. Beran (1994) and Robinson (1994a).

Interesting spectral behaviour is also possible at one or more other frequencies between 0 and π, corresponding to analogous concepts of memory. We can assume

$$f(\omega + \lambda) \sim C|\lambda|^{-2d} \quad \text{as} \quad \lambda \to 0 \tag{1.2}$$

for $0 < C < \infty$, $|d| < 1/2$ and $\omega \in (0, \pi)$. There is a spectral pole at ω if $d > 0$ and a zero if $d < 0$. (1.2) provides an alternative approach to standard methods of modelling seasonal or cyclic behaviour, where ω represents either one of the seasonal frequencies or the cycle.

Parametric and semiparametric models conforming to (1.2) and their estimation and testing have been discussed by such authors as Andel (1986), Carlin and Dempster (1989), Chung (1996), Giraitis and Leipus (1995), Gray et al. (1989,1994), Hassler (1994), Hosking (1984), Ooms (1995), Porter-Hudak (1990), Ray (1993) and Robinson (1994b). Models and methods for the standard case $\omega = 0$ extend fairly straightforwardly to (1.2) with known $\omega \in (0, \pi)$, but moving from $\omega = 0$ to $\omega \in (0, \pi)$ also broadens the scope for modelling because we can extend (1.2) to

$$\begin{aligned} f(\omega + \lambda) &\sim C_1\lambda^{-2d_1} \quad \text{as} \;\; \lambda \to 0^+, \\ f(\omega - \lambda) &\sim C_2\lambda^{-2d_2} \quad \text{as} \;\; \lambda \to 0^+, \end{aligned} \tag{1.3}$$

where $\omega \in (0, \pi)$,

$$0 < C_i < \infty, \quad |d_i| < \frac{1}{2}, \quad i = 1, 2, \tag{1.4}$$

and we permit

$$d_1 \neq d_2 \quad \text{and/or} \quad C_1 \neq C_2. \tag{1.5}$$

Since the spectrum is symmetric about frequencies zero and π, the possibility of (1.5) is excluded for $\omega = 0, \pi$, but for $\omega \in (0, \pi)$ any values of C_i and d_i satisfying (1.4) are possible. Clearly (1.3) and (1.4) nest (1.2) as a special case. We call a process with spectrum satisfying (1.3)-(1.5) a Seasonal/Cyclical Asymmetric Long Memory (SCALM) process. Note that both the intercept and the slope in the local linear relationship between $\log f(\omega + \lambda)$ and $\log \lambda$ can differ from those pertaining to $f(\omega - \lambda)$.

The following section introduces examples of SCALM processes. Sections 3 and 4 extend to SCALM processes some methods of estimation proposed for the case (1.2) when $\omega = 0$ (where the spectrum is symmetric), namely the log-periodogram and Gaussian semiparametric or local Whittle estimates, and introduce tests for spectral symmetry, $d_1 = d_2$. The behaviour of these tests in finite samples is analyzed in Section 5 through a small Monte Carlo study. An empirical application to a monthly UK inflation series is introduced in Section 6. Section 7 concludes. Technical details are placed in the Appendix.

2 SEASONAL/CYCLICAL ASYMMETRIC LONG MEMORY MODELS

Although the stress in the present paper is on semiparametric inference based on the "local" SCALM model (1.3)-(1.5), it is important to describe parametric models which define $f(\lambda)$ over the whole of the Nyqvist band while conforming to (1.3)-(1.5), thereby demonstrating that processes satisfying (1.3)-(1.5) exist (at least under Gaussianity) and providing ways of computer generation of such processes. We begin by describing existing models for the special case (1.2) and then indicate how we can build on them to cover the more general behaviour in (1.3)-(1.5).

Seasonal fractional noise is characterized by a spectral density

$$f(\lambda) = C|1 - \cos(s\lambda)| \sum_{n=-\infty}^{\infty} \left| n + \frac{s\lambda}{2\pi} \right|^{-2(1+d)}$$

and autocovariances

$$\gamma_j = \frac{Ex_t^2}{2}\left(\left|\frac{j}{s}+1\right|^{2d+1} - 2\left|\frac{j}{s}\right|^{2d+1} + \left|\frac{j}{s}-1\right|^{2d+1}\right),$$

where s is the number of observations per year (see for instance Jonas (1983), Carlin and Dempster (1989) or Ooms (1995)). Thus $f(\lambda)$ satisfies (1.2) at $[(s+1)/2]$ values of ω, namely at the origin and seasonal frequencies $\omega_j = \frac{2\pi j}{s}$, $j = 1, ..., [(s-1)/2]$, $[\cdot]$ denoting integer part. All spectral poles/zeros are symmetric in the sense of (1.2) and of the same magnitude.

Another class of seasonal or cyclic long memory models such as those in Hosking (1984), Andel (1986), Chan and Wei (1988), Robinson (1994b), Chan and Terrin (1995) or Giraitis and Leipus (1995) is defined by

$$D(L)x_t = u_t \quad t = 1, 2, ..., \tag{2.1}$$

where L is the lag operator,

$$D(z) = (1-z)^{d_0}\{\prod_{j=1}^{r-1}(1 - 2z\cos\omega_j + z^2)^{d_j}\}(1+z)^{d_r}, \tag{2.2}$$

and u_t is a short memory process (for instance a stationary and invertible ARMA). (2.1) and (2.2) allow spectral poles/zeros of different magnitude at frequencies ω_j, $j = 0, ..., r$, where $\omega_0 = 0$ and $\omega_r = \pi$. If $D(z) = (1-z)^d$, (2.1) satisfies (1.2) with $\omega = 0$; $D(z) = (1+z)^d$ satisfies (1.2) with $\omega = \pi$; if $D(z) = (1-2z\cos\omega+z^2)^d$ and $\omega \in (0,\pi)$, (2.1) is the so called Gegenbauer process (Gray et al. (1989,1994)), satisfying (1.2) for the stated ω; $D(z) = (1-z^{12})^d$ is the seasonal model used by Porter-Hudak (1990) for USA monetary aggregates and (1.2) holds for $\omega_j = 2\pi j/12$, $j = 0, 1, ..., 6$; $D(z) = (1-z^3)^{d_3}(1-z^{12})^{d_{12}}$ is used by Ray (1993) for monthly IBM revenue data; if $\omega_j = 2\pi j/s$ for $j = 1, 2, ...[(s-1)/2]$, that is, if ω_j are seasonal frequencies based on s observations per year, then (2.1) has been called "flexible ARFISMA" (Hassler (1994)) or "flexible (seasonal) ARMA(p,d,q)$_s$" (Ooms (1995)).

One way of defining a spectrum that satisfies (1.3) and (1.5) is to extend the Gegenbauer process of Gray et al. (1989), considering

$$f(\lambda) = \begin{cases} \frac{\sigma_1^2}{2\pi}\theta(\lambda;\omega)^{-2d_1} & \text{if} \quad \omega < \lambda \leq \pi, \\ \frac{\sigma_2^2}{2\pi}\theta(\lambda;\omega)^{-2d_2} & \text{if} \quad 0 \leq \lambda \leq \omega, \end{cases} \tag{2.3}$$

where $\theta(\lambda;\omega) = |1-2e^{i\lambda}\cos\omega + e^{2i\lambda}|$. Clearly (1.3)-(1.5) hold under (2.3). The corresponding γ_j satisfy

$$\begin{aligned}\gamma_j &= \frac{\sigma_2^2}{2\sqrt{\pi}}(2\sin\omega)^{\frac{1}{2}-2d_2}\Gamma(1-2d_2)P_{j-\frac{1}{2}}^{2d_2-\frac{1}{2}}(\cos\omega) \\ &+ (-1)^j\frac{\sigma_1^2}{2\sqrt{\pi}}(2\sin\omega)^{\frac{1}{2}-2d_1}\Gamma(1-2d_1)P_{j-\frac{1}{2}}^{2d_1-\frac{1}{2}}(-\cos\omega), \qquad (2.4)\end{aligned}$$

where $P_a^b(x)$ are associated Legendre functions and $\Gamma(\cdot)$ is the gamma function. Equation (2.4) can be obtained as in Chung (1996) for the symmetric case, applying formula 3.663.1 in Gradshteyn and Ryzhik (1980). The asymptotic behaviour of γ_j is

$$\gamma_j \approx j^{2d_1-1}\sin(\pi d_1 - j\omega) + j^{2d_2-1}\sin(\pi d_2 + j\omega) \quad \text{as} \quad j \to \infty \qquad (2.5)$$

where $a \approx b$ if $\frac{a}{b} \to C$, where C is a finite nonzero constant (see the Appendix). The autocovariances (2.5) not only decrease hyperbolically as is typical of long-range dependent data, but are also affected by the cyclic behaviour of the sine function with period depending on ω. Expression (2.4) simplifies when $\omega = \pi/2$, i.e. $\cos\omega = 0$:

$$\gamma_j = \frac{\sigma_2^2}{2}\frac{\Gamma(1-2d_2)}{\Gamma(1-d_2-\frac{j}{2})\Gamma(1-d_2+\frac{j}{2})} + (-1)^j\frac{\sigma_1^2}{2}\frac{\Gamma(1-2d_1)}{\Gamma(1-d_1-\frac{j}{2})\Gamma(1-d_1+\frac{j}{2})}, \qquad (2.6)$$

$j = 0, \pm1, \pm2, \ldots$ (see the Appendix) . Gaussian x_t satisfying (2.6) can be readily generated, using for example the algorithm of Davies and Harte (1987).

To facilitate the understanding of SCALM processes consider further the case $\omega = \pi/2$. When $d_1 > d_2$, the spectrum (2.3) shows that cyclical components with period approaching 4 from below (frequency approaching $\pi/2$ from above) are stronger than cycles of period approaching 4 from above. The autocovariances (2.6) displayed in Figure 1 for $j = 86$ through $j = 100$ exhibit a corresponding pattern. When $d_1 = 0.4$ and $d_2 = 0.1$ (Figure 1a) γ_j is larger just above lag $4k$ than just below. When $d_1 = 0.1$ and $d_2 = 0.4$ (Figure 1b) the situation is reversed.

This phenomenon might help to explain a difference that might be observed in economic time series between autocorrelation of summer and autumn observations on the one hand and summer and spring ones on the other.

It is straightforward to extend (2.3) to a more general spectrum with asymmetric pole/zero by taking

$$f(\lambda) = \begin{cases} \frac{\sigma_1^2}{2\pi}\theta(\lambda;\omega)^{-2d_1}\rho(\lambda) & \omega < \lambda \leq \pi, \\ \frac{\sigma_2^2}{2\pi}\theta(\lambda;\omega)^{-2d_2}\rho(\lambda) & 0 < \lambda \leq \omega, \end{cases}$$

where $\rho(\lambda)$ is a nonnegative, even function, for example the spectrum of a stationary fractional ARIMA process. To write down a spectrum with two or more asymmetric poles/zeros on $(0,\pi)$ is more complicated. Suppose $f(\lambda)$ has poles/zeros at frequencies $0 < \omega_1 < \omega_2 ... < \omega_{r-1} < \pi$ and possibly at $\omega_0 = 0$ and $\omega_r = \pi$. Let $S_j = (\omega_{j-1}, \omega_j]$ for $j = 1, ..., r$. Write

$$\begin{aligned} h_0(\lambda) &= |1 - e^{i\lambda}|^{-d_0}, \\ h_r(\lambda) &= |1 + e^{i\lambda}|^{-d_r}, \\ h_{j,k}(\lambda) &= \theta(\lambda;\omega_j)^{-d_{j,k}} \quad , \quad j = 1, ..., r-1 \quad , \quad k = 1, 2, \\ g(\lambda) &= h_0(\lambda)^2 h_r(\lambda)^2 \prod_{j=1}^{r-1} h_{j,1}(\lambda) h_{j,2}(\lambda), \end{aligned}$$

where $d_0, d_r, d_{jk} \in (-1/2, 1/2)$. Let $g_j(\lambda)$, $j = 1, ..., r$, be even, positive and bounded functions in $[-\pi, \pi]$. Now specify

$$f(\lambda) = \begin{cases} g_1(\lambda) g(\lambda) \frac{h_{12}(\lambda)}{h_{11}(\lambda)} & \text{if} \quad \lambda \in S_1, \\ g_j(\lambda) g(\lambda) \frac{h_{j-1,1}(\lambda) h_{j,2}(\lambda)}{h_{j-1,2}(\lambda) h_{j,1}(\lambda)} & \text{if} \quad \lambda \in S_j, \qquad j = 2, 3, ..., r-1, \\ g_r(\lambda) g(\lambda) \frac{h_{r-1,1}(\lambda)}{h_{r-1,2}(\lambda)} & \text{if} \quad \lambda \in S_r \ . \end{cases} \tag{2.7}$$

For example consider the case $r = 3$ and for simplicity $d_0 = d_3 = 0$ so that there are two spectral poles/zeros in $(0, \pi)$ and none at 0 or π. Let $g_i(\lambda) = \frac{\sigma_i^2}{2\pi}$ for $i = 1, 2, 3$. Then

$$f(\lambda) = \begin{cases} \frac{\sigma_1^2}{2\pi}\theta(\lambda;\omega_1)^{-2d_{12}}\theta(\lambda;\omega_2)^{-d_{21}}\theta(\lambda;\omega_2)^{-d_{22}} & ,0 < \lambda \leq \omega_1, \\ \frac{\sigma_2^2}{2\pi}\theta(\lambda;\omega_1)^{-2d_{11}}\theta(\lambda;\omega_2)^{-2d_{22}} & ,\omega_1 < \lambda \leq \omega_2, \\ \frac{\sigma_3^2}{2\pi}\theta(\lambda;\omega_1)^{-d_{11}}\theta(\lambda;\omega_1)^{-d_{12}}\theta(\lambda;\omega_2)^{-2d_{21}} & ,\omega_2 < \lambda \leq \pi \ . \end{cases}$$

On the other hand if $d_{j1} = d_{j2} = d_j$ and $g_j(\lambda) = f_u(\lambda)$ for $0 \leq j \leq r$ then

$$f(\lambda) = \frac{f_u(\lambda)}{|1 - e^{i\lambda}|^{2d_0} |1 + e^{i\lambda}|^{2d_r} \prod_{j=1}^{r-1} \theta(\lambda;\omega_j)^{2d_j}} \tag{2.8}$$

which is the spectrum of the "symmetric" process (2.1) and (2.2). Note that there is not in general actual symmetry around a non-zero and non-π pole/zero in (2.8), except for example when

$$f(\lambda) = \frac{\sigma^2}{2\pi\theta(\lambda;\omega)^{2d}}, \quad \omega = \pi/2. \tag{2.9}$$

However (2.8) does possess the asymptotic symmetry property of (1.2) for $\omega = 0, \omega_1, ..., \omega_{r-1}, \pi$.

In general it is not possible to obtain an exact expression for the autocovariances corresponding to the SCALM spectrum (2.7) when there are two or more spectral poles/zeros in $(0,\pi)$. However when $0 < d_i < 1/2$ for $i = 0, 1, ..., r$, asymptotic behaviour is given by

$$\begin{aligned}\gamma_j &\approx j^{2d_0-1} + j^{2d_r-1}\sin(\pi d_r + j\pi) \\ &+ \sum_{k=1}^{r-1}\{j^{2d_{k1}-1}\sin(\pi d_{k1} - j\omega_k) + j^{2d_{k2}-1}\sin(\pi d_{k2} + j\omega_k)\}\end{aligned}$$

as $j \to \infty$, and so γ_j is ultimately governed by the largest of the d's.

Because SCALM processes are more readily represented in the frequency domain than in the time domain, a natural method of estimating parametric models is a discrete version of Whittle's method. On the basis of observations x_t, $t = 1, ..., n$, define the discrete Fourier transform

$$w(\lambda) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} x_t e^{it\lambda} \tag{2.10}$$

and the periodogram

$$I(\lambda) = |w(\lambda)|^2. \tag{2.11}$$

We consider $I(\lambda)$ for frequencies $\lambda = \lambda_j = 2\pi j/n$, $j = 1, ..., n-1$, where $I(\lambda)$ is invariant to location shift so that there is no need to estimate an unknown mean of x_t. Denote the parametric spectrum of x_t by $f(\lambda;\phi)$, where ϕ is an unknown $p \times 1$ parameter vector with true value ϕ_0. For example in case (2.3) $\phi = (d_1, \sigma_1^2, d_2, \sigma_2^2)'$ assuming ω is known. Write

$$Q(\phi) = \sum_j{}' \left\{\log f(\lambda_j;\phi) + \frac{I(\lambda_j)}{f(\lambda_j;\phi)}\right\}, \tag{2.12}$$

where $\sum_j'$ is a sum over $j = 1, ..., n-1$, excluding those λ_j coinciding with supposed poles/zeros in $f(\lambda;\phi)$, and define

$$\tilde{\phi} = \arg\min_{\phi\in\Phi} Q(\phi),$$

where Φ is some compact subset of R^p. Then under suitable regularity conditions, $\tilde{\phi}$ will be consistent for ϕ_0, $\sqrt{n}(\tilde{\phi} - \phi_0)$ will be asymptotically normal with mean zero and a consistently estimable covariance matrix, and for Gaussian x_t, $\tilde{\phi}$ will be asymptotically efficient. Asymptotic theory for parametric Gaussian estimates of long memory models has been set down by Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990), Heyde and Gay (1993), Hosoya (1996a,b) and others, albeit for time domain estimates or the continuous Whittle function, rather than (2.12). This theory does not directly include SCALM models, but, undoubtedly can be suitably extended. These remarks are based on the presumption that the location of the poles/troughs in $f(\lambda)$ is known, as is natural in case of seasonal processes. Asymptotic distribution theory for Gaussian estimates allowing for unknown poles/zeros has been considered in case of "symmetric" models such as (2.9), though rigorous proofs are not yet available, so a theory for SCALM models with unknown spectral poles/zeros is also not available. In any case, whether or not the location of the poles/zeros is known, the benefits of parametric estimation are offset by the disadvantage of inconsistency in case of misspecification of the parametric form. For example the d estimates, which are of principal interest in explaining behaviour near spectral poles/zeros, will be inconsistent even if f is misspecified only at frequencies where it is smooth. For this reason we focus on semiparametric methods which are based on the specification (1.3), (1.4) and allow valid inference under more general circumstances.

3 LOG-PERIODOGRAM REGRESSION

Due to their simplicity, perhaps the most popular semiparametric methods of estimating the memory parameter d in (1.2) with $\omega = 0$ are variants of the log-periodogram regression introduced by Geweke and Porter-Hudak (1983). Here d is estimated by a least squares regression of $\log I(\lambda_j)$ on $-2\log\lambda_j$ with an intercept. The regression is carried out for $j = 1, ..., m$, where m is an integer between 1 and $n/2$, called the bandwidth, satisfying at least

$$\frac{1}{m} + \frac{m}{n} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.1}$$

The original version of this approach, due to Geweke and Porter-Hudak (1983), uses instead the regressor $-\log\{4\sin^2(\lambda_j/2)\}$, but as indicated by Robinson (1995a), use of the simpler $-2\log\lambda_j$, which corresponds more naturally to (1.2), leads to equivalent first-order asymptotic properties.

Note that $I(\lambda_j)$ is an even function, so when $\omega = 0$ using $j = \pm1, ..., \pm m$, is equivalent to using just $j = 1, ..., m$. When $\omega \neq 0$ in the symmetric model (1.2), use of information on both sides of the spectral pole/zero can make a substantial difference. We estimate d by

$$\hat{d} = -\frac{1}{2}\frac{\sum_{j=\pm 1}^{\pm m} \nu_j \log I(\omega + \lambda_j)}{\sum_{j=\pm 1}^{\pm m} \nu_j^2} \tag{3.2}$$

where $\nu_j = \log|j| - \frac{1}{m}\sum_{l=1}^{m} \log l$. In the SCALM model (1.3), on the other hand, a natural suggestion is to run separate regressions on each side of ω, so

$$\hat{d}_1 = -\frac{1}{2}\frac{\sum_{j=1}^{m} \nu_j \log I(\omega + \lambda_j)}{\sum_{j=1}^{m} \nu_j^2} \quad , \quad \hat{d}_2 = -\frac{1}{2}\frac{\sum_{j=1}^{m} \nu_j \log I(\omega - \lambda_j)}{\sum_{j=1}^{m} \nu_j^2}. \tag{3.3}$$

Thus (3.2) can be regarded as a pooling of $\hat{d}_1$ and $\hat{d}_2$,

$$\hat{d} = \frac{\hat{d}_1 + \hat{d}_2}{2}, \tag{3.4}$$

which is valid only when $d_1 = d_2$ (but whether or not $C_1 = C_2$).

Work on estimating (1.2) when $\omega = 0$ suggests two possible modifications to this scheme. Due to anomalous behaviour of the periodogram very close to a spectral pole/zero, Künsch (1986) and Robinson (1995a) trimmed out very low frequencies from the regression. Recent work, under somewhat different conditions than those in Robinson (1995a), by Hurvich, Deo and Brodsky (1998), suggests that trimming may not be necessary in order to achieve basic asymptotic properties. However trimming seems harder to avoid in SCALM models, since, if $d_1 < d_2$, frequencies λ_j just before ω exert a relatively serious effect on those just after, producing some potential to contaminate $\hat{d}_1$. Since we are unlikely to know a priori whether $d_1 < d_2$ or $d_1 > d_2$ we trim both sides of ω. It is possible that this trimming might be avoidable, specially if we replace (2.10) by a tapered discrete Fourier transform. Note also that we might use different bandwidths m in $\hat{d}_1$ and $\hat{d}_2$, and in other estimates of d_1 and d_2 considered in this paper.

The second type of modification is an efficiency improvement suggested by Robinson (1995a) for (1.2) with $\omega = 0$ and based on pooling adjacent periodogram ordinates. To extend this idea to $\omega \in (0, \pi)$ consider for some integer $J \geq 1$,

$$\hat{I}_{\omega Jk} = \sum_{j=1}^{J} I(\omega - \lambda_{k+j-J}) \quad , \quad \tilde{I}_{\omega Jk} = \sum_{j=1}^{J} I(\omega + \lambda_{k+j-J}) \quad ,$$

and then write

$$\hat{d}_1^{(J)} = -\frac{1}{2}\frac{\sum_k' \nu_k \log \tilde{I}_{\omega Jk}}{\sum_k' \nu_k^2} \quad , \quad \hat{d}_2^{(J)} = -\frac{1}{2}\frac{\sum_k' \nu_k \log \hat{I}_{\omega Jk}}{\sum_k' \nu_k^2}, \tag{3.5}$$

where now $\sum_k'$ is a sum over $k = l + J, l + 2J, ..., m$, and $\nu_k = \log k - \frac{J}{m-l}\sum' \log j$, assuming $m - l$ is an integer multiple of J. Here we have implemented the trimming discussed above by means of the user-chosen number l. The integer J describes the extent of the efficiency improvement, when $J = 1$ (3.5) is equivalent to (3.3) apart from the trimming, asymptotic efficiency improving with increasing J.

Under some assumptions similar to those in Robinson (1995a) which we state in the Appendix we get the following result:

Theorem 1 *Let Assumptions* **A.1-A.4** *hold. Then for $J \geq 1$, as $n \to \infty$, $2\sqrt{m}(\hat{d}_i^{(J)} - d_i)$, $i = 1, 2$, converge to independent $N(0, J\psi'(J))$ random variables where $\psi'(z) = \frac{d}{dz}\psi(z)$ and $\psi(z) = \frac{d}{dz}\log\Gamma(z)$ is the digamma function. Thus under the null hypothesis of symmetry:*

$$H_0 : d_1 = d_2, \tag{3.6}$$

we have that

$$\left\{\frac{2m}{J\psi'(J)}\right\}^{\frac{1}{2}} (\hat{d}_1^{(J)} - \hat{d}_2^{(J)}) \xrightarrow{d} N(0, 1). \tag{3.7}$$

The proof of this theorem extends that of Theorem 3 of Robinson (1995a) in a relatively straightforward way, taking into account our Theorem 5 in the Appendix on the covariances of discrete Fourier transforms, and is thus omitted. For details see Arteche (1998). If (3.6) is not rejected we can estimate $d_1 = d_2 = d$ by

$$\hat{d}^{(J)} = \frac{1}{2}(\hat{d}_1^{(J)} + \hat{d}_2^{(J)}), \tag{3.8}$$

see (3.4).

The limit distributional properties of $\hat{d}^{(J)}$ are readily deduced from Theorem 1 in a manner that indicates both the inconsistency caused by an incorrect a priori assumption of symmetry (3.6) and the improvement that (3.8) affords over $\hat{d}_1^{(J)}$ and $\hat{d}_2^{(J)}$ under symmetry.

Corollary 1 *Let* **A.1-A.4** *hold. Then as $n \to \infty$,*

$$\sqrt{8m}(\hat{d}^{(J)} - \frac{1}{2}(d_1 + d_2)) \xrightarrow{d} N(0, J\psi'(J))$$

so that under symmetry (3.6),

$$\sqrt{8m}(\hat{d}^{(J)} - d) \xrightarrow{d} N(0, J\psi'(J))$$

where $d = d_1 = d_2$.

In view of the local character of $\hat{d}_1^{(J)}$ and $\hat{d}_2^{(J)}$ we can similarly estimate the left and right memory parameters at each of several spectral poles/zeros, ω_j, as permitted in the modelling of Section 2. It is clear from Theorem 5 in the Appendix that the asymptotic properties of the left and right d estimates will not vary across the ω_j, and moreover the estimates will be asymptotically independent across the ω_j, so that we can readily construct statistics for testing hypothesis across the ω_j, for example of equality of all the left or right memory parameters. In the interests of parsimony this would be a useful preliminary to parametric modelling.

4 GAUSSIAN SEMIPARAMETRIC ESTIMATION

To estimate d in (1.2) with $\omega = 0$, Robinson (1995b) showed that a local type of Whittle estimate, which he termed Gaussian semiparametric, has a smaller asymptotic variance than the estimates of the previous section. This estimate can readily be extended to (1.2) with $\omega \neq 0$, and also to SCALM models, though we have found it necessary to introduce some trimming (not employed at all by Robinson (1995b)) of frequencies close to ω in the asymmetric case.

Consider, for $0 < \omega < \pi$, the functions,

$$\begin{aligned} Q_1(C,d) &= \frac{1}{m-l}\sum_{j=l+1}^{m}\left\{\log C\lambda_j^{-2d} + \frac{\lambda_j^{2d}}{C}I(\omega+\lambda_j)\right\}. \\ Q_2(C,d) &= \frac{1}{m-l}\sum_{j=l+1}^{m}\left\{\log C\lambda_j^{-2d} + \frac{\lambda_j^{2d}}{C}I(\omega-\lambda_j)\right\}. \end{aligned}$$

We estimate (C_1, d_1) and (C_2, d_2) by minimizing Q_1 and Q_2 respectively. Eliminating C, we have the estimates

$$\tilde{d}_i = \arg\min_{\Theta} R_i(d) \quad , \quad i = 1, 2,$$

of d_1, d_2, where

$$R_i(d) = \log \tilde{C}_i(d) - \frac{2d}{m-l}\sum_{l+1}^{m}\log\lambda_j,$$

$$\tilde{C}_1(d) = \frac{1}{m-l}\sum_{j=l+1}^{m}\lambda_j^{2d}I(\omega+\lambda_j)\ ,\ \tilde{C}_2(d) = \frac{1}{m-l}\sum_{j=l+1}^{m}\lambda_j^{2d}I(\omega-\lambda_j), \tag{4.1}$$

and Θ is a closed subset of $(-1/2, 1/2)$. Under assumptions stated in the Appendix and similar to those in Robinson (1995b) we obtain:

Theorem 2 *If Assumptions* **B.1**, **A.2**, **B.3** *and* **B.4** *hold, then as* $n \to \infty$, $\tilde{d}_i \xrightarrow{p} d_i$ *for* $i = 1, 2$. *If Assumptions* **B.1**, **C.2** *and* **C.3** *and either* **C.4** *or* **C.5-C.6** *hold, then as* $n \to \infty$, $2\sqrt{m}(\tilde{d}_i - d_i)$, $i = 1, 2$, *converge to independent* $N(0,1)$ *random variables. Thus under symmetry* (3.6),

$$\sqrt{2m}(\tilde{d}_1 - \tilde{d}_2) \xrightarrow{d} N(0,1). \quad (4.2)$$

The proof is similar to that in Robinson (1995b) and is thus omitted; for a full proof see Arteche (1998). Because $J\psi'(J) > 1$, it seems by comparison with Theorem 1 that $\tilde{d}_1 - \tilde{d}_2$ produces a locally more powerful test of (3.6) than $\tilde{d}_1^{(J)} - \tilde{d}_2^{(J)}$ for any J. However $\tilde{d}_1$ and $\tilde{d}_2$ are not defined in closed form unlike $\tilde{d}_1^{(J)}$ and $\tilde{d}_2^{(J)}$. It is possible to alleviate this problem by means of a Lagrange Multiplier (LM) test which entails only estimation of a single parameter under the null hypothesis. This estimation would in any case be of interest if the test based on (4.2) fails to reject (3.6), so we discuss it first.

Define

$$\bar{d} = \arg\min_{\Theta}\{R_1(d) + R_2(d)\}. \quad (4.3)$$

Theorem 3 *Let Assumptions* **B.1**, **C.2** *and* **C.3** *and either* **C.4** *or* **C.5** *and* **C.6** *hold. Then under symmetry* (3.6), *as* $n \to \infty$,

$$\sqrt{8m}(\bar{d} - d) \xrightarrow{d} N(0,1).$$

Now consider an LM test of the hypothesis

$$\Delta = 0 \quad (4.4)$$

based on the criterion

$$R(\delta, \Delta) = R_1(\delta + \Delta) + R_2(\delta),$$

where δ here represents any admissible value of the common $d = d_1 = d_2$ under (4.4), which is thus equivalent to (3.6).

Theorem 4 *Let Assumptions* **B.1**, **C.2** *and* **C.3** *and either* **C.4** *or* **C.5** *and* **C.6** *hold. Then under symmetry* (3.6), *as* $n \to \infty$,

$$LM = 2m\frac{L_1^2}{L_0^2} \xrightarrow{d} \chi_1^2, \quad (4.5)$$

where $L_k = \frac{1}{m-l}\sum_{l+1}^{m} \nu_j^k \lambda_j^{2\tilde{d}} I(\omega+\lambda_j)$, $\nu_j = \log j - \frac{1}{m-l}\sum_{l+1}^{m}\log j$ *and* $\tilde{d}$ *is defined by (4.3). The test based on rejecting (3.6) if* $LM > \chi^2_{1\alpha}$ *at* $100\alpha\%$ *significance level is consistent.*

5 FINITE SAMPLE BEHAVIOUR

In this section we study via Monte Carlo analysis the finite sample performance of the Wald statistics (3.7) and (4.2) and LM statistic (4.5) for the hypothesis of spectral symmetry (3.6). We first generated two independent Gaussian processes $\{\epsilon_{1,t}\}$ and $\{\epsilon_{2,t}\}$ with zero means and lag-j autocovariances

$$\begin{aligned}\gamma_{1j} &= \sigma_1^2\left(\delta_{j0} - \frac{\sin(j\omega)}{\pi j}\right),\\ \gamma_{2j} &= \sigma_2^2\frac{\sin(j\omega)}{\pi j},\end{aligned}$$

respectively, where $\delta_{j0} = 1$ if $j = 0$ and 0 otherwise. It follows that $\epsilon_{1,t}$ and $\epsilon_{2,t}$ have spectra

$$f_{\epsilon_1}(\lambda) = \begin{cases} 0, & 0 \le \lambda < \omega, \\ \frac{\sigma_1^2}{2\pi}, & \omega \le \lambda \le \pi, \end{cases} \tag{5.1}$$

and

$$f_{\epsilon_2}(\lambda) = \begin{cases} \frac{\sigma_2^2}{2\pi}, & 0 \le \lambda < \omega, \\ 0, & \omega \le \lambda \le \pi. \end{cases} \tag{5.2}$$

Now define the processes $\{x_{j,t}\}$, $j = 1, 2$, by

$$(1 - 2L\cos\omega + L^2)^{d_j} x_{j,t} = \epsilon_{j,t}\ ,\ \ j = 1,2, \quad t = 0 \pm 1 \ldots . \tag{5.3}$$

Thus the $\{x_{j,t}\}$ have spectra

$$f_{x_j}(\lambda) = \frac{f_{\epsilon_j}(\lambda)}{|1 - e^{i\lambda}\cos\omega + e^{2i\lambda}|^{2d_j}}\ ,\ \ 0 \le \lambda < \pi\ ,\ \ j = 1, 2,$$

and in view of (5.1) and (5.2) and independence of the $\{\epsilon_{j,t}\}$, $j = 1, 2$, $x_t = x_{1,t} + x_{2,t}$ has spectrum

$$f(\lambda) = f_{x_1}(\lambda) + f_{x_2}(\lambda)$$

which is seen to be identical to (2.3). In order to generate realizations of x_t, from Gray et al. (1989) we can rewrite (5.3) as

$$\sum_{s=0}^{\infty} C_s^{(d_j)}(\cos\omega) x_{j,t-s} = \epsilon_{j,t}\ ,\ \ j = 1, 2,\ \ t = 0, \pm 1, \ldots, \tag{5.4}$$

where the Gegenbauer polynomials $C_s^{(d)}(\eta)$ are of the form

$$C_s^{(d)}(\eta) = \sum_{k=0}^{[s/2]} \frac{(-1)^k \Gamma(s-k-d)(2\eta)^{s-2k}}{\Gamma(k+1)\Gamma(s-2k+1)\Gamma(-d)}.$$

We truncate the sum in (5.4) so that actually our generated $x_{j,t}$ are

$$x_{j,t} = -\sum_{s=1}^{1500} C_s^{(d_j)}(\cos\omega) x_{j,t-s} + \epsilon_{j,t}, \tag{5.5}$$

where $x_{j,t} = 0$ for $t \leq 0$. We prefer an autoregressive truncation over the moving average one of Gray et al. (1989) because autoregressive coefficients decay faster. The Gegenbauer functions are obtained via the recursion

$$C_s^{(d)}(\eta) = 2\eta\left(\frac{-d+s-1}{s}\right) C_{s-1}^{(d)}(\eta) - \left(\frac{-2d+s-2}{s}\right) C_{s-2}^{(d)}(\eta).$$

(see formula 8.933.1 in Gradshteyn and Ryzhik (1980)). This method permits the approximate generation of processes with an asymmetric spectral pole/zero at any frequency between 0 and π. For $\omega = \pi/2$ an exact procedure is possible by means of the algorithm of Davies and Harte (1987). Comparison of exact and approximate procedures on the basis of sample autocovariance plots indicated little difference in performance.

We carried out simulations for $\omega = \pi/4, \pi/2, 3\pi/4$ but report results only for $\omega = \pi/2$ because these are fairly typical. We took $d_1, d_2 = \{-0.4, -0.2, 0, 0.2, 0.4\}$ and $\sigma_1^2 = \sigma_2^2 = c$ where c was taken to be 1 with no loss of generality. Differing σ_1^2 and σ_2^2 would not affect our results as our earlier comments concerning robustness to differing C_1, C_2 indicate. Assumptions **A.1** with $\alpha = 2$, **C.2**(**A.2**), **C.3** and **C.5** since ϵ_{1t} and ϵ_{2t} are Gaussian, are satisfied. The Gaussian semiparametric estimates needed to construct (4.2) and (4.5) were obtained by a simple golden section search on the first derivative of the objective function. The minimization was carried out over $\Theta = [-0.499, 0.499]$. The log-periodogram estimates were calculated with $J = 1$. Two sample sizes were analyzed, $n = 256$ and $n = 512$. For each three different bandwidths were used, $m = n/16$, $n/8$ and $n/4$. If one of the $\epsilon_{j,t}$ were replaced by a short range dependent process with a peak near ω, the choice of m would become more delicate (see Robinson (1995b)). Performance was found to worsen with exclusion of frequencies close to ω, because although trimming seems hard to avoid in asymptotic theory it can worsen estimates in finite samples, unless the difference between d_2 and d_1 is very large (see also Arteche (1998)). The tables contain sizes (along the NW-SE diagonal) and powers at nominal 5% significance level. The number of replications was 1000 and all calculations were done using GAUSS-386i.

Tables 1 and 2 concern Wald tests based on log-periodogram estimates with $J = 1$ (within parentheses) and Gaussian semiparametric estimates, introduced in Theorems 1 and 2 respectively. We performed two-tailed tests, squaring the test statistics in Theorems 1 and 2, and comparing them with the χ_1^2 5% critical values. As expected power tends to increase with m and n. Monte Carlo size is always larger than nominal size but tends to it as m and n increase. It is also noticeable that size is smaller for extreme values of d_1 and d_2.

The behaviour of the LM test introduced in Theorem 4 is described in Tables 3 and 4. Power and size tend to increase with d_1 and d_2, reflecting a more conservative behaviour of this test under antipersistence than under persistence. The LM test tends to be more conservative than the Wald, with markedly lower powers for the smaller n and m, and sizes that are always less than nominal ones, though they increase with m and n.

6 APPLICATION TO UK INFLATION

In this section we apply the techniques introduced in this paper to a monthly UK inflation series. The series analysed is $x_t = \log p_t - \log p_{t-1}$ from May 1915 to April 1996, where p_t is the logged Retail Price Index. Thus $n = 972$. All calculations and figures were done using S-Plus 3.1.

Figure 2 displays the periodogram of x_t. Of course this is not a consistent estimate of the spectral density, but the sharp peaks at the origin, and to varying extents at seasonal frequencies, suggest the possibility of low-frequency as well as seasonal long memory. We extract the same conclusion from the plot of the first 150 sample autocorrelations in Figure 3. Oscillations decay very slowly, like those encountered in the theoretical study of seasonal/cyclical long memory models in Section 2.

Seasonality has usually been treated either by including seasonal dummies or seasonal differencing. The unsuitability of the former treatment, so far as UK inflation is concerned, has been pointed out by Hassler and Wolters (1995). Figure 4 shows the periodogram of the seasonally differenced series $(1 - L^{12})x_t$. The deep troughs at the origin and at seasonal frequencies suggest possible overdifferencing, so a milder, fractional, differencing could be more appropriate. Moreover Figure 2 suggests that peaks may be of differing magnitudes, and thus the possibility of different persistence parameters at the origin and across seasonal frequencies. Table 5 shows the untrimmed Gaussian semiparametric and log-periodogram ($J = 1$) estimates of the memory parameters at the origin and at seasonal frequencies. We use a small bandwidth, $m = 30$, to avoid the influence of neighbouring spectral peaks. Since the asymptotic standard deviations of log-periodogram and Gaussian semiparametric estimates of d_1 and d_2, both with $m = 30$, are 0.117 and 0.091 respectively (see Theorems 1 and 2), no significant differences from zero at the 5% level are found for frequencies $5\pi/6$ and π, whereas the origin and remaining seasonal frequencies have at least one significant right or left estimate. We analyse the possibility of asymmetric spectral poles by testing (3.6). Table 6 displays the log-periodogram Wald (Wlp), Gaussian semiparametric Wald (Wgs) and LM test statistics with $m = 30$ for $\omega_j = 2\pi j/12$, $j = 1, 2, 3, 4, 5$, where Wlp and Wgs are the squares of the statistics in Theorems 1 and 2. On the basis of our asymptotic theory we do not reject spectral symmetry at the 5% level for any of the ω_j. Of course the exact values of Wlp, Wgs and LM depend on m. Figures 5, 6 and 7 show the various test statistics with $m = 11, ..., 50$, for the hypothesis of spectral symmetry at $\pi/6$, $\pi/3$ and $2\pi/3$, these implying cycles of periods 12, 6 and 3 respectively. We do not find strong evidence of spectral asymmetry, although the null hypothesis tends to be rejected at frequency $\pi/6$ for small m and at $\pi/3$ for large m, in the latter case rejection only occurring using the log-periodogram Wald test. We observe again the relatively more conservative behaviour of the LM test found in the Monte Carlo analysis of Section 5.

7 CONCLUSION

Most research done to date on long memory focuses on spectral behaviour at zero frequency. The various attempts to extend this concept to seasonal or cyclical long memory tend to assume, with loss of generality, an asymptotic symmetry around spectral poles/zeros. Consequently tests of spectral symmetry seem relevant as a prior step to parametric modelling, while semiparametric estimates which incorrectly assume symmetry will be inconsistent. In this paper we introduce three such semiparametric tests against an alternative SCALM class of processes, whose properties and generation are also discussed.

Many parametric seasonal long memory models impose also an equality of the memory parameters at the origin and seasonal frequencies (e.g. Porter-Hudak (1990) or Ray (1993)). Tests between memory parameters across different frequencies can be defined similarly, which might also be a useful preliminary to parametric modelling.

8 Appendix

The asymptotic behaviour of γ_j in (2.5) is obtained by applying equation 8.721.3 in Gradshteyn and Ryzhik (1980), namely

$$P_a^b(\cos\theta) = \frac{2}{\sqrt{\pi}}\frac{\Gamma(a+b+1)}{\Gamma(a+\frac{3}{2})}\frac{\cos[(a+\frac{1}{2})\theta - \frac{\pi}{4} + \frac{b\pi}{2}]}{\sqrt{2\sin\theta}}\left[1 + O\left(\frac{1}{a}\right)\right],$$

and Stirling's formula in (2.4). To deduce γ_j in (2.6) we apply equation 8.756.1 in Gradshteyn and Ryzhik (1980), namely

$$P_a^b(0) = \frac{2^b\sqrt{\pi}}{\Gamma(\frac{a-b}{2}+1)\Gamma(\frac{-a-b+1}{2})}.$$

Assumptions for log-periodogram regression:

A.1: For a frequency $\omega \in (0,\pi)$ there exists $\alpha \in (0,2]$ such that, as $\lambda \to 0^+$,

$$\begin{aligned} f(\omega+\lambda) &= C_1\lambda^{-2d_1}(1+O(\lambda^\alpha)),\\ f(\omega-\lambda) &= C_2\lambda^{-2d_2}(1+O(\lambda^\alpha)), \end{aligned}$$

where $C_1, C_2 \in (0,\infty)$ and $d_1, d_2 \in (-1/2, 1/2)$.

A.2: In a neighbourhood $(-\delta,0)\cup(0,\delta)$ of ω $f(\lambda)$ is differentiable and, as $\lambda \to 0^+$,

$$\begin{aligned} \left|\frac{\mathrm{d}}{\mathrm{d}\lambda}f(\omega+\lambda)\right| &= O(\lambda^{-1-2d_1}),\\ \left|\frac{\mathrm{d}}{\mathrm{d}\lambda}f(\omega-\lambda)\right| &= O(\lambda^{-1-2d_2}). \end{aligned}$$

A.3: $\{x_t, t = 0, \pm1, \pm2, ...\}$ is a Gaussian process.

A.4: As $n \to \infty$,

$$\frac{\sqrt{m}n^{2|d_1-d_2|}\log m}{l^{1+2|d_1-d_2|}} + \frac{l(\log n)^2}{m} + \frac{m^{1+\frac{1}{2\alpha}}}{n} \to 0.$$

No assumption on $f(\lambda)$ outside a neighbourhood of the frequency ω is imposed, apart from integrability implied by covariance stationarity. Assumption **A.1** strengthens (1.3) by imposing a rate of convergence of $f(\omega+\lambda)/(C_1\lambda^{-2d_1})$ and $f(\omega-\lambda)/(C_2\lambda^{-2d_2})$ to 1. We could have generalized **A.1** allowing for different $\alpha's$ before and after ω but this would have complicated the notation and the results obtained would have been similar. To see the implications of **A.4** take $m \sim n^\theta$ and $l \sim n^\phi$. Thus **A.4** entails

$$2|d_1-d_2|+\frac{1}{2}\theta-\phi(1+2|d_1-d_2|)<0\ ,\ \phi<\theta\ ,\ \theta\left(1+\frac{1}{2\alpha}\right)<1. \tag{8.1}$$

The first two conditions imply $\theta>\phi>4|d_1-d_2|/(1+4|d_1-d_2|)$, and incorporating the last condition in (8.1) indicates that we must have $\alpha>2|d_1-d_2|$. Because $|d_1-d_2|<1$, **A.4** can be satisfied for any d_1, d_2 if $\alpha=2$. Under (3.6), Assumption **A.4** is the corresponding condition of Robinson (1995a) (indeed it may be shown that the latter condition suffices for the limit distributional properties of $\tilde{d}_i^{(J)}$ if $d_i \geq d_j$).

As shown by Robinson (1995a,b), the asymptotic properties of log-periodogram and Gaussian semiparametric estimates are substantially dependent on those of discrete Fourier transforms. Put $v_{1j}=v_{1j}(\omega)=w(\omega+\lambda_j)/(C_1^{\frac{1}{2}}\lambda_j^{-d_1})$, $v_{2j}=v_{2j}(\omega)=w(\omega-\lambda_j)/(C_2^{\frac{1}{2}}\lambda_j^{-d_2})$, $g_{hi}(j)=n^{d_h-d_i}j^{\frac{1}{2}+d_i-d_h}$, $i=1,2$, $h=1,2,m$, where $d_m=\max(d_1,d_2)$, and denote $\bar{v}_{ij}$ the complex conjugate of v_{ij}.

Theorem 5 *Let assumption* **A.1** *and* **A.2** *hold. Then for any sequences of positive integers $j=j(n)$ and $k=k(n)$ such that $j>k$ and $\frac{j}{n}\to 0$ as $n\to\infty$, for $i=1,2$,*

a) $E|v_{ij}|^2=1+O\left(\log j\, g_{mi}(j)^2+\left(\frac{j}{n}\right)^\alpha\right)$

b) $Ev_{ij}^2=O(\log j\, g_{mi}^2(j))$

c) $Ev_{ij}\bar{v}_{ik}=O(\log j\, g_{mi}(j)g_{mi}(k)))$

d) $Ev_{ij}v_{ik}=O(\log j\, g_{mi}(j)g_{mi}(k))$

e) $Ev_{1j}\bar{v}_{2j}=O(\log j(g_{12}(j)+g_{21}(j))$

f) $Ev_{1j}v_{2j}=O\left(\frac{\log j}{j}\right)$

g) $Ev_{1j}\bar{v}_{2k}=O(\log j(g_{21}(j)+g_{12}(k)))$

h) $Ev_{1j}v_{2k} = O(\log j(g_{21}(j) + g_{12}(k)))$.

The proof of Theorem 5 is omitted because it is based on that of Theorem 2 in Robinson (1995a), but it can be seen in Arteche (1998).

Assumptions for Gaussian semiparametric estimation:

B.1: **A.1** holds with $d_i \in \Theta$, $i = 1, 2$.

B.3: $x_t - Ex_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$ and $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ where $E[\epsilon_t|F_{t-1}] = 0$, $E[\epsilon_t^2|F_{t-1}] = 1$ a.s. for $t = 0, \pm 1, \pm 2, ...$, F_t is the σ-field generated by ϵ_s, $s \leq t$, and there exists a random variable ϵ such that $E\epsilon^2 < \infty$ and for all $\eta > 0$ and some $\kappa < 1$, $P(|\epsilon_t| > \eta) \leq \kappa P(|\epsilon| > \eta)$.

B.4: As $n \to \infty$,

$$\frac{m}{n} + \frac{l}{m} \log m + \frac{n^{2|d_1 - d_2|}}{l^{1+2|d_1-d_2|}} (\log m)^3 \to 0.$$

C.2: In a neighbourhood $(-\delta, 0) \cup (0, \delta)$ of ω, $\alpha(\lambda) = \sum_{k=0}^{\infty} \alpha_k e^{ik\lambda}$ is differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \alpha(\omega \pm \lambda) = O\left(\frac{|\alpha(\omega \pm \lambda)|}{\lambda}\right) \quad \text{as } \lambda \to 0^+.$$

C.3: Assumption **B.3** holds and

$$E(\epsilon_t^3|F_{t-1}) = \mu_3 \quad \text{and} \quad E(\epsilon_t^4|F_{t-1}) = \mu_4, \quad t = 0, \pm 1, ...,$$

for finite constants μ_3 and μ_4.

C.4: As $n \to \infty$,

$$\frac{(\log m)^3}{l^2} + \frac{l^3}{m} (\log m)^4 + \frac{n^{2|d_1 - d_2|}}{l^{1+2|d_1-d_2|}} \log m + \frac{m^{1+2\alpha}}{n^{2\alpha}} (\log m)^2 \to 0.$$

Taking $m \sim n^\theta$, $l \sim n^\phi$ as before, it follows that because $|d_1 - d_2| < 1$, $1 > \theta > \phi > 2/3$ will suffice to satisfy **B.4** for all d_1, d_2. Condition **C.4** is more restrictive and can only be satisfied if $|d_1 - d_2| < \alpha/(3 + 4\alpha)$ where the upper bound is $2/11$ for $\alpha = 2$. However we can relax **C.4** by strengthening **C.3**. We thus consider:

C.5: The fourth cumulant of ϵ_t is zero for all t.

C.6: As $n \to \infty$,

$$\frac{(\log m)^3}{l^2} + \frac{l^2}{m} (\log m)^2 + \frac{n^{2|d_1 - d_2|}}{l^{1+2|d_1-d_2|}} \log m + \frac{m^{1+2\alpha}}{n^{2\alpha}} (\log m)^2 \to 0.$$

Assumption **C.5** is implied by Gaussianity, and **C.6** entails $|d_1 - d_2| < \alpha/(2 + 2\alpha)$ where the upper bound is $1/3$ when $\alpha = 2$. This requirement is not much stronger than $|d_1 - d_2| <$

1/2 which is implied if there is both a left and a right spectral pole at ω. If $d_i \geq d_j$ then, as in Robinson (1995b), no trimming is in fact needed for the asymptotics of $\tilde{d}_i$, and $1/m + m/n \to 0$ and $m^{-1} + m^{1+2\alpha}(\log m)^2/n^{2\alpha} \to 0$ as $n \to \infty$, suffice for consistency and asymptotic normality respectively of $\tilde{d}_i$.

Proof of Theorem 4: Consider

$$\frac{\partial}{\partial \Delta} R(\delta, \Delta)\Big|_{(\tilde{d},0)} = \frac{\partial}{\partial \Delta} R_1(\delta + \Delta)\Big|_{(\tilde{d},0)} = 2\left\{\frac{D_1(\tilde{d})}{C_1(\tilde{d})} - \frac{1}{m-l}\sum_{j=l+1}^{m} \log \lambda_j\right\}, \tag{8.2}$$

where $\tilde{d}$ is given by (4.3), $C_1(d)$ by (4.1) and we define

$$D_1(\delta) = \frac{1}{m-l}\sum_{l+1}^{m} \lambda_j^{2\delta}(\log \lambda_j) I(\omega + \lambda_j) \;\; , \;\; D_2(\delta) = \frac{1}{m-l}\sum_{l+1}^{m} \lambda_j^{2\delta}(\log \lambda_j) I(\omega - \lambda_j).$$

By the mean value theorem (8.2) is

$$2\frac{D_1(d)}{C_1(d)} - \frac{2}{m-l}\sum_{j=l+1}^{m} \log \lambda_j + 2\frac{\mathrm{d}}{\mathrm{d}\delta}\left\{\frac{D_1(\delta)}{C_1(\delta)}\right\}\Big|_{\delta=\bar{d}} (\tilde{d} - d)$$

for $|\bar{d} - d| \leq |\tilde{d} - d|$. Proceeding as in the proof of Theorem 2 of Robinson (1995b), as $n \to \infty$,

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\left\{\frac{D_1(\delta)}{C_1(\delta)}\right\}\Big|_{\delta=\bar{d}} \xrightarrow{p} 2,$$

and

$$\sqrt{m}(\tilde{d} - d) = -\frac{\sqrt{m}}{4}\left\{\frac{D_1(d)}{C_1(d)} + \frac{D_2(d)}{C_2(d)} - \frac{2}{m-l}\sum_{j=l+1}^{m} \log \lambda_j\right\}(1 + o_p(1)).$$

Thus we have, again as in Robinson (1995b) (see also Arteche (1998)),

$$\begin{aligned}\sqrt{m}\frac{\partial}{\partial \Delta} R(\delta, \Delta)\Big|_{(\tilde{d},0)} &= \sqrt{m}\left\{\frac{D_1(d)}{C_1(d)} - \frac{1}{m-l}\sum_{j=l+1}^{m} \log \lambda_j\right\} - \sqrt{m}\left\{\frac{D_2(d)}{C_2(d)} - \frac{1}{m-l}\sum_{j=l+1}^{m} \log \lambda_j\right\} + o_p(1) \\ &\xrightarrow{d} N(0,2) \qquad \text{as} \quad n \to \infty,\end{aligned} \tag{8.3}$$

since the terms in braces are asymptotically independent (see Arteche (1998)) and each of them converge to $N(0,1)$ random variables as shown in Robinson (1995b). Noting that $2L_1/L_2 = \partial R(\tilde{d}, 0)/\partial \Delta$ and (8.3) the proof of (4.5) is straightforward. We omit the proof of consistency to save space. For details see Arteche (1998).

References

[1] ANDEL, J. (1986) Long-Memory Time Series Models. *Kybernetika.* 22, 105-123.

[2] ARTECHE, J. (1998) *Seasonal and Cyclical Long Memory in Time Series.* PhD Thesis, University of London.

[3] BERAN, J. (1994) *Statistics for Long-Memory Processes.* Monographs on Statistics and Applied Probability 61. Chapman & Hall, New York.

[4] CARLIN, J.B. and DEMPSTER, A.P. (1989) Sensitivity analysis of seasonal adjustments: empirical case studies. *J. Am. Statist. Assoc.* 84, 6-20.

[5] CHAN, N.H. and TERRIN, N. (1995) Inference for unstable long-memory processes with applications to fractional unit root autoregressions. *Ann. Statist.* 23, 1662-1683.

[6] CHAN, N.H. and WEI, C.Z. (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* 16, 367-401.

[7] CHUNG, C.F. (1996) A generalized fractionally integrated autoregressive moving average process. *J. Time Ser. Anal.* 17, 111-140.

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

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

[10] 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.

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

[12] GIRAITIS, L. and LEIPUS, R. (1995) A generalized fractionally differencing approach in long-memory modelling. *Liet. Matem. Rink.* 35, 65-81.

[13] 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. Th. Rel. Fields.* 86, 87-104.

[14] GRADSHTEYN, J.S. and RYZHIK, I.W. (1980) *Table of Integrals, Series and Products.* Orlando, Florida: Academic Press.

[15] GRAY, H.L., ZHANG, N.F. and WOODWARD, W.A. (1989) On generalized fractional processes. *J. Time Ser. Anal.* 10, 233-257.

[16] GRAY, H.L., ZHANG, N.F. and WOODWARD, W.A. (1994) On generalized fractional processes- a correction. *J. Time Ser. Anal.* 15, 561-562.

[17] HASSLER, U. (1994) (Mis)specification of long-memory in seasonal time series. *J. Time Ser. Anal.* 15, 19-30.

[18] HASSLER, U. and WOLTERS, J. (1995) Long memory in inflation rates: international evidence. *J. Bus. Econ. Statist.* 13, 37-45.

[19] HEYDE, C.C. and GAY, R. (1993) Smoothed periodogram asymptotics for processes and fields with possible long-range dependence. *Stoch. Proc. Appl.* 45, 169-182.

[20] HOSKING, J.R.M. (1984) Modelling persistence in hydrological time series using fractional differencing. *Water Resour. Res.* 20, 1898-1908.

[21] HOSOYA, Y. (1996a) A limit theory for long-range dependence and statistical inference on related fields. Preprint.

[22] HOSOYA, Y. (1996b) The quasi-likelihood approach to statistical inference on multiple time series with long-range dependence. *J. Econometrics.* 73, 217-236.

[23] HURVICH, C.M., DEO, R. and BRODSKY, J. (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long memory time series. *J. Time Ser. Anal.* 19, 19-46.

[24] JONAS, A.J. (1983) *Persistent Memory Random Processes*, PhD Thesis, Department of Statistics, Harvard University.

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

[26] OOMS, M. (1995) Flexible seasonal long-memory and economic time series. Preprint.

[27] PORTER-HUDAK, S. (1990) An application of seasonal fractionally differenced model to the monetary aggregates. *J. Am. Statist. Assoc.* 85, 338-344.

[28] RAY, B.K. (1993) Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model. *Int. J. Forecasting.* 9, 255-269.

[29] ROBINSON, P.M. (1994a) Time series with strong dependence. *Advances in Econometrics*, Vol.1. Cambridge: Cambridge University Press, 47-95.

[30] ROBINSON, P.M. (1994b) Efficient tests of non-stationary hypothesis. *J. Am. Statist. Assoc.* 89, 1420-1437.

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

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