Statistica Sinica **6**(1996), 97-112

OPTIMAL SPECTRAL BANDWIDTH FOR LONG MEMORY

Miguel A. Delgado and Peter M. Robinson

Universidad Carlos III de Madrid and London School of Economics

Abstract. For long range dependent time series with a spectral singularity at frequency zero, a theory for optimal bandwidth choice in non-parametric analysis of the singularity was developed by Robinson (1994a). In the present paper, the optimal bandwidths are compared with those in case of a smooth spectrum. They are also analysed in case of fractional ARIMA models and calculated as a function of the self similarity parameter in some special cases. Feasible data-dependent approximations to the optimal bandwidth are proposed. We also include some applications using real data.

Key words and phrases: Long range dependence, spectrum, self similarity parameter, semiparametric estimation, optimal bandwidth choice, ARFIMA models.

1. Introduction

A theory of optimal bandwidth choice in nonparametric spectral estimation was developed many years ago (see e.g. Parzen (1957)). This theory, in large part, precedes the corresponding optimal bandwidth literature for nonparametric probability density and regression estimation, though it has not been developed to the same extent. There are considerable similarities between the two types of theory. In both cases, a nonparametric estimate of an unknown function at a given point of the domain borrows information from neighbouring points. The extent of such information is largely determined by a "bandwidth" number, and the choice of this considerably affects the estimate. Too large a bandwidth tends to be associated with a large bias, too small a bandwidth with a large variance. One usually seeks a bandwidth which balances bias and imprecision. A mathematically simple way of doing this consists of minimizing a form of mean squared error of the nonparametric estimate, either at a particular point of interest, or else averaged across an interval, even the whole domain. Typically, a closed form formula for an 'optimal' bandwidth results, depending on the precise way the nonparametric estimate has been implemented and on features of the nonparametric function, in particular, smoothness properties.

In the spectral estimation situation, as well as the probability density and regression situations, it is typically assumed that the unknown function is at least finite at all points at which it is estimated. This assumption may be controversial in case of spectral estimation. Some plots of spectral estimates exhibit sharp peaks (so that it has long been common practice to use a logarithmic scale), and this could be consistent with a singularity in the spectral density. Correspondingly, plots of sample autocorrelations are sometimes indicative of a slow rate of decay. Consequently there has been considerable study of 'long range dependent' parametric and nonparametric models which imply a singularity in the spectral density, typically at zero frequency. A recent literature survey is in Robinson (1994c).

Recently, Robinson (1994a) has developed some optimality theory for nonparametric frequency domain estimation in case of long range dependence. The present paper elaborates on and extends his work. The following section compares his results with those for 'short memory' time series with a smooth spectral density. In Section 3 these formulae are further analyzed and numerically illustrated for fractional ARIMA (ARFIMA) models. Feasible approximations of the optimal bandwidth are proposed in Section 4, and applied in Section 5, to the analysis of annual minimum water levels of the river Nile (which has also illustrated many other methods of long memory time series analysis), and to the analysis of inflation rate, using Spanish data.

2. Optimal Spectral Bandwidth

Denote by $X_t, t = 0, \pm 1, \pm 2, \ldots$, a discrete parameter covariance stationary time series; for the sake of simplicity we suppose X_t is also Gaussian, though our conclusions have more general relevance. Denote the lag-j autocovariance of X_t by $\gamma_j = E[(X_j - E(X_0))(X_0 - E(X_0))]$, $j = 0, \pm 1, \pm 2, \ldots$, so $f(\lambda)$, the spectral density of X_t, satisfies $\gamma_j = \int_{-\pi}^{\pi} f(\lambda) \cos j\lambda \, d\lambda$. For a realization of size n, introduce the periodogram $I(\lambda) = (2\pi n)^{-1} |\sum_{t=1}^{n} X_t e^{it\lambda}|^2$.

All estimates in the paper depend on $I(\lambda)$ computed at frequencies $\lambda_j = 2\pi j/n$ for integer j, where $1 \leq j < n$. Note that $E(X_0)$ is not assumed to be zero (or known) and for $j \neq 0(\mathrm{mod}(n))$, $I(\lambda_j)$ is invariant to location change in the X_t.

We focus on estimation around zero frequency when dealing with long range dependence, in which case $f(0) = \infty$. However, suppose first that $0 < f(0) < \infty$ and

$$f(\lambda)/f(0) = 1 + E_\alpha \lambda^\alpha + o(|\lambda|^\alpha), \text{ as } \lambda \to 0+, \tag{2.1}$$

for some $\alpha \in (0, 2]$, where $0 < |E_\alpha| < \infty$. This condition essentially says that, in a neighbourhood of $\lambda = 0$, $f(\lambda)$ satisfies a Lipschitz condition of degree α for $0 < \alpha \leq 1$, or $f(\lambda)$ is differentiable and its derivative satisfies a Lipschitz condition of degree $\alpha - 1$, and is zero at $\lambda = 0$, for $1 < \alpha \leq 2$. We have

$E_1 = d\log f(0)/d\lambda$ and $E_2 = (d^2 f(0)/d\lambda^2)/(2f(0))$. Consider estimating $f(0)$ (see e.g. Brillinger 1975, Robinson 1983) by

$$\hat{f}(0) = m^{-1}\sum_{j=1}^{m} I(\lambda_j), \tag{2.2}$$

where m is an integer between 1 and n such that

$$m^{-1} + mn^{-1} \to 0 \quad \text{as } n \to \infty. \tag{2.3}$$

Under (2.3) and additional regularity conditions the scaled mean squared error of $\hat{f}(0)$,

$$E\{[\hat{f}(0) - f(0)]/f(0)\}^2 \sim m^{-1} + E_\alpha^2 \lambda_m^{2\alpha}/(\alpha+1)^2, \tag{2.4}$$

and an optimal m, minimizing the right hand side is

$$m_{\text{opt}} = \left[\frac{(\alpha+1)^2}{2\alpha(2\pi)^{2\alpha}E_\alpha^2}\right]^{\frac{1}{2\alpha+1}} n^{\frac{2\alpha}{2\alpha+1}}. \tag{2.5}$$

Now consider processes with spectral density satisfying

$$f(\lambda) \sim g_H(\lambda) = G\lambda^{1-2H} \quad \text{as } \lambda \to 0+, \tag{2.6}$$

$1/2 < H < 1$, $0 < G < \infty$. Because $f(0)$ is now infinite it is no longer meaningful to estimate it, but it is of interest to investigate the impact on optimal bandwidth in case one attempts to estimate $f(0)$ in the incorrect belief that it is finite. Additionally, Robinson (1994a) has shown that an optimal type of spectral bandwidth is relevant to the choice of bandwidth in the semiparametric estimate of H proposed by Robinson (1994b). The criterion (2.4) is no longer relevant, but Robinson (1994a) suggested the extended criterion

$$E\{(\hat{F}(\lambda_m) - F_H(\lambda_m))/F_H(\lambda_m)\}^2, \tag{2.7}$$

where

$$\hat{F}(\lambda_m) = \frac{2\pi m}{n}\hat{f}(0), \qquad F_H(\lambda) = G\lambda^{2(1-H)}/(2(1-H)). \tag{2.8}$$

To extend condition (2.1) it is assumed for $\alpha \in (0,2]$,

$$f(\lambda)/G\ \lambda^{1-2H} = 1 + E_\alpha(H)\lambda^\alpha + o(\lambda^\alpha) \text{ as } \lambda \to 0+, \tag{2.9}$$

where $0 < |E_\alpha(H)| < \infty$, $1/2 < H < 1$. In general $E_\alpha(H)$ depends on H as well as α, as will be illustrated subsequently. In case $1/2 < H < 3/4$, under (2.3) and additional conditions Robinson (1994a) established that

$$(2.2) \sim 4(1-H)^2\left[\frac{1}{(3-4H)m} + \left\{\frac{E_\alpha(H)}{2-2H+\alpha}\right\}^2 \lambda_m^{2\alpha}\right], \text{ as } n \to \infty, \tag{2.10}$$

and an optimal m is

$$m_{\rm opt}(H) = \left\{ \frac{(2-2H+\alpha)^2}{2\alpha(2\pi)^{2\alpha} E_\alpha^2(H)(3-4H)} \right\}^{1/(2\alpha+1)} n^{2\alpha/(2\alpha+1)}. \tag{2.11}$$

The rate of convergence in (2.11) is identical to that in (2.5), so that long range dependence affects only the multiplying factor in the optimal m. Note finally that the formulae (2.10) and (2.11) reduce to (2.4) and (2.5) on taking $H = 1/2$ and $G = f(0)$. Robinson (1994a) showed that (2.10) and (2.11) also hold when G in (2.6) is replaced by a function that varies slowly with λ.

When $3/4 < H < 1$, $f(\lambda)$ is no longer square-integrable on a neighbourhood of the origin, and under (2.3) and additional conditions Robinson (1994a) showed that, with $D_H = 2\Gamma(2(1-H))\cos((1-H)\pi)$,

$$(2.7) \sim A_1(2\pi m)^{4H-4} + A_2(2\pi m)^{2H-2+\alpha} n^{-\alpha} + A_3(2\pi m)^{2\alpha} n^{-2\alpha}, \tag{2.12}$$

where

$$A_1 = 2D_H^2(1-H)^2 \left\{ \frac{1}{(4H-3)(2H-1)} + \frac{1}{2H^2(2H-1)^2} - \frac{1}{H^2(4H-1)} - \frac{4\Gamma(2H-1)^2}{\Gamma(4H)} \right\},$$

$$A_2 = -\frac{4D_H E_\alpha(H)(1-H)^2}{H(2H-1)(2-2H+\alpha)}, \quad A_3 = \frac{4E_\alpha(H)^2(1-H)^2}{(2-2H+\alpha)^2},$$

and it is minimized with respect to m by

$$m_{\rm opt}(H) \sim \frac{n^{\frac{\alpha}{2-2H+\alpha}}}{2\pi} \Big\{ \frac{D_H(2-2H+\alpha)}{4\alpha} \Big[\frac{2H-2+\alpha}{E_\alpha(H)(2H-1)} + \frac{1}{|E_\alpha(H)|} \Big[\frac{(2-2H+\alpha)^2}{H^2(2H-1)^2} + 16\alpha(1-H) \Big\{ \frac{1}{(4H-3)(2H-1)} - \frac{1}{H^2(4H-1)} - \frac{4\Gamma(2H-1)^2}{\Gamma(4H)} \Big\} \Big]^{1/2} \Big] \Big\}^{\frac{1}{2-2H+\alpha}}. \tag{2.13}$$

3. Fractional ARIMAs

A fractional differencing representation is given by

$$f(\lambda) = |1 - e^{i\lambda}|^{1-2H} h(\lambda), \tag{3.1}$$

where $0 < h(0) < \infty$. In the ARFIMA$(p, H-1/2, q)$ model,

$$h(\lambda) = \frac{\sigma^2}{2\pi} \frac{|b(e^{i\lambda})|^2}{|a(e^{i\lambda})|^2}, \tag{3.2}$$

where

$$a(z) = 1 - \sum_{j=1}^{p} a_j z^j, \quad b(z) = 1 - \sum_{j=1}^{q} b_j z^j, \tag{3.3}$$

all zeros of a and b are outside the unit circle in the complex plane, and $\sigma^2 > 0$ (see e.g. Adenstedt (1974); Hosking (1981)). In general, and as is the case in the ARFIMA model, assume that $h(\lambda)$ has first derivative $h'(0) = 0$, and second derivative $h''(0)$. Then

$$\begin{aligned} \frac{f(\lambda)}{G\,\lambda^{1-2H}} &= \frac{h(\lambda)}{G}\left(\frac{\sin(\lambda/2)}{\lambda/2}\right)^{1-2H} \\ &\sim G^{-1}\Big\{h(0) + h''(0)\lambda^2/2\Big\}\,\Big\{1 - (\lambda/2)^2/6\Big\}^{1-2H} \\ &\sim G^{-1}\Big\{h(0) + h''(0)\lambda^2/2\Big\}\,\Big\{1 - (1-2H)\lambda^2/24\Big\} \\ &\sim 1 + \Big\{h''(0)/2h(0) + (2H-1)/24\Big\}\lambda^2, \end{aligned} \tag{3.4}$$

on taking $G = h(0)$. Thus

$$E_2(H) = h''(0)/2h(0) + (2H-1)/24. \tag{3.5}$$

The second component of $E_2(H)$ is positive and takes values zero when $H = 1/2$, $1/48$ when $H = 3/4$, and $1/24$ when $H = 1$. Note that $E_2(H) = (2H-1)/24$ in the ARFIMA$(0, H - 1/2, 0)$ case. The first component of (3.5) can be positive or negative and it can be large or small.

We can get a more useful picture of the variability in $E_\alpha(H)$ by studying the ARFIMA case (3.1)-(3.3). Put

$$\begin{aligned} a &= a(1) = 1 - \sum_{j=1}^{p} a_j, & b &= b(1) = 1 - \sum_{j=1}^{q} b_j, \\ a' &= \frac{d}{d\lambda}\, a(e^{i\lambda})\Big|_{\lambda=0} = -i\sum_{j=1}^{p} j a_j, & b' &= \frac{d}{d\lambda}\, b(e^{i\lambda})\Big|_{\lambda=0} = -i\sum_{j=1}^{q} j b_j, \\ a'' &= \frac{d^2}{d\lambda^2}\, a(e^{i\lambda})\Big|_{\lambda=0} = \sum_{j=1}^{p} j^2 a_j, & b'' &= \frac{d^2}{d\lambda^2}\, b(e^{i\lambda})\Big|_{\lambda=0} = \sum_{j=1}^{q} j b_j^2. \end{aligned}$$

It is easily shown that

$$\frac{h''(0)}{h(0)} = \left(\frac{\bar{b}''}{\bar{b}} + \frac{b''}{b} + 2\frac{|b'|^2}{|b|^2}\right) - \left(\frac{\bar{a}''}{\bar{a}} + \frac{a''}{a} + 2\frac{|a'|^2}{|a|^2}\right),$$

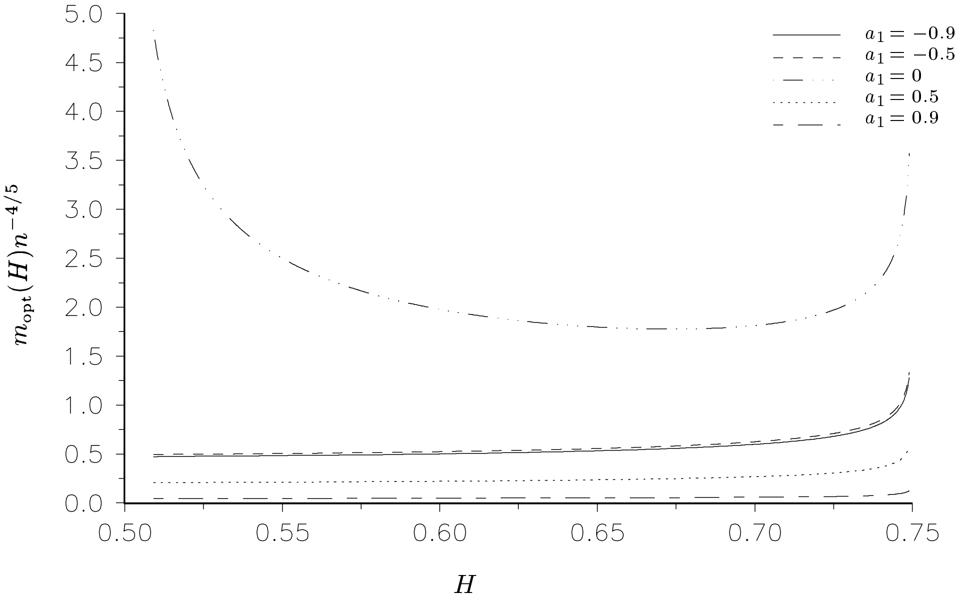

Figure 1. Plots of $m_{\text{opt}}(H)n^{-4/5}$ for the ARFIMA$(1, H-1/2, 0)$ model for $1/2 < H < 3/4$, $(1-a_1L)(1-L)^{H-1/2}X_t = \varepsilon_t$.

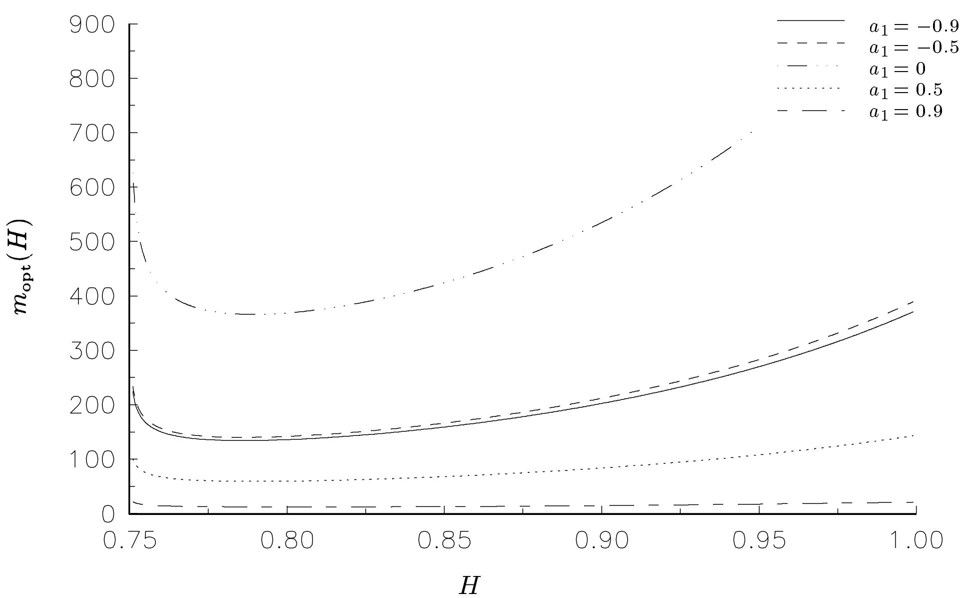

Figure 2. Plots of $m_{\text{opt}}(H)$ for the ARFIMA$(1, H-1/2, 0)$ model for $3/4 < H < 1$, $(1-a_1L)(1-L)^{H-1/2}X_t = \varepsilon_t$, $n = 800$.

where the overbar denotes complex conjugation. In the ARFIMA$(1, H-1/2, 0)$ case we have

$$\frac{h''(0)}{2h(0)} = -\left\{\frac{a_1}{1-a_1} + \left(\frac{a_1}{1-a_1}\right)^2\right\} = -\frac{a_1}{(1-a_1)^2}, \tag{3.6}$$

and in the ARFIMA$(2, H-1/2, 0)$ case

$$\frac{h''(0)}{2h(0)} = -\left\{\frac{a_1+4a_2}{1-a_1-a_2} + \left(\frac{a_1+2a_2}{1-a_1-a_2}\right)\right\}^2 = \frac{a_1-a_1a_2+4a_2}{(1-a_1-a_2)^2}. \tag{3.7}$$

Corresponding ARFIMA$(0, H-1/2, 2)$ formulae are obtained by replacing a's by b's and changing sign. For the ARFIMA$(1, H-1/2, 1)$ case

$$\frac{h''(0)}{2h(0)} = \frac{b_1}{(1-b_1)^2} - \frac{a_1}{(1-a_1)^2}. \tag{3.8}$$

Focusing on the ARFIMA$(1, H-1/2, 0)$ case, (3.6) indicates that $h''(0)/2h(0)$ approaches minus infinity when a_1 approaches 1; for example, it is -90 when $a_1 = 0.9$ and -990 when $a_1 = 0.99$. Thus, for large or moderate a_1, $E_2(H)$ will be dominated by the $h''(0)/2h(0)$ component. In Figure 1, we plot $m_{\text{opt}}(H)n^{-4/5}$, $\alpha = 2$, versus H, for $1/2 < H < 3/4$, using (2.11). When $a_1 = 0$, $E_2(H)$ is very small, $m_{\text{opt}}(H)n^{-4/5}$ takes very large values, and $E_2(H) \to 0$ as $H \to 1/2$ (i.e. $m_{\text{opt}}(H) \to \infty$ as $H \to 1/2$). For other a_1 values $m_{\text{opt}}(H)n^{-4/5}$ suffers little variation with respect to H, but as (2.11) indicates, for any a_1, $m_{\text{opt}}(H)$ increases quickly when H is close to $3/4$. Figure 2 shows plots of $m_{\text{opt}}(H)$ against H for $n = 800$ when $3/4 < H < 1$ using (2.12). The magnitudes of $m_{\text{opt}}(H)$ have the same ordering with respect to a_1 as in Figure 1. (qualitatively similar results were obtained here, and in subsequent numerical work we describe, for $n = 400$).

In the ARFIMA$(2, H-1/2, 0)$ case, for $a_1^2 + 4a_2 < 0$ the roots of the characteristic polynomial are complex conjugate, corresponding to a finite peak in $h(\lambda)$, and hence possibly in $f(\lambda)$, at a nonzero frequency. Figure 3 shows plots of $\log f(\lambda)$ for different H values. We present two examples where a peak in $f(\lambda)$ at $\lambda \neq 0$ is present. The precise location of such peak is affected by H. In the short memory case ($H = 1/2$), the peak is located at $\lambda = \pi/4$ if $a_1(a_2-1)/4a_2 = 1/\sqrt{2}$, which can happen if $a_1 = 1.172$ and $a_2 = -0.707$; while a peak is located at $\pi/6$ if $a_1(a_2-1)/4a_2 = \sqrt{3}/2$, which can happen if $a_1 = 1.268$ and $a_2 = -0.577$. However, for this latter (a_1, a_2) value, there is hardly a peak for $\lambda > 0$ for the larger H. Nevertheless, when there is a peak, and if m is chosen sufficiently large so that λ_m is to the right of the peak, then an estimate of H based on the $I(\lambda_j)$ for $1 \le j \le m$ (such as that discussed in the following section) might have a serious negative bias.

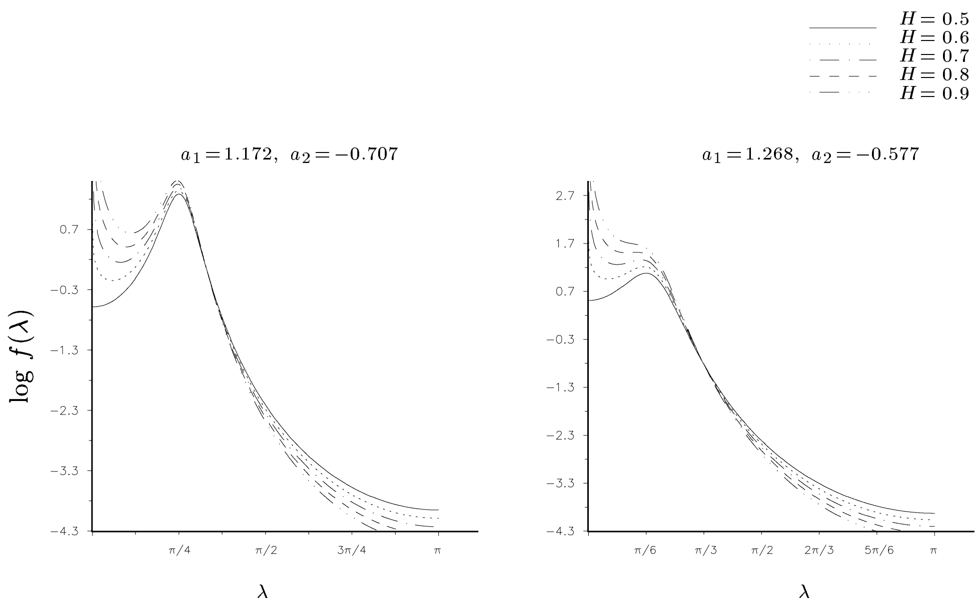

Figure 3. Log spectral density of the ARFIMA$(2, H-1/2, 0)$ model with a peak at $\lambda \neq 0$, $(1-L)^{H-1/2}(1-a_1L-a_2L^2)X_t = \varepsilon_t$.

4. Feasible Approximations to the Optimal Bandwidth

In order to approximate the optimal bandwidth, we need an estimate of H. Robinson (1994b) has suggested the estimate

$$\hat{H}_{mq} = 1 - \frac{\log\{\hat{F}(q\lambda_m)/\hat{F}(\lambda_m)\}}{2\log q}, \tag{4.1}$$

where $q \in (0,1)$. This estimate is consistent for H even when G in (2.6) is replaced by a function that varies slowly at $\lambda = 0$. As noted by Robinson (1994b), we always have $H_{mq} \leq 1$, so (as with Yule-Walker estimation of autoregressive coefficients) a "stationary" estimate will almost certainly result even if the data come from a nonstationary process (e.g. one with a unit root). There are a number of tests for a unit root that can be applied at an initial stage. The bulk of these are directed against autoregressive alternatives, but one that is directed against fractional alternatives, and may thus be more relevant in our setting, is a special case of the class of Robinson (1994d).

In order to illustrate the behaviour of $\hat{H}_{mq}$ evaluated at the optimal bandwidth values, we performed a small Monte Carlo experiment, generating data according to a Gaussian ARFIMA$(1, H-1/2, 0)$ with $a_1 = 0.5$. Figure 4 presents plots of sample root mean squared errors (RMSE) and biases of $\hat{H}_{mq}$ $(q = 1/2)$ from 5000 replications against m, for various values of H and $n = 1000$. Biases can be positive or negative, increasing with H and decreasing with m. The m which minimizes the MSE differs from $m_{\text{opt}}(H)$. Even the theoretical MSE of $\hat{H}_{mq}$ will differ from that of $m_{\text{opt}}(H)$, depending, among other things, on q.

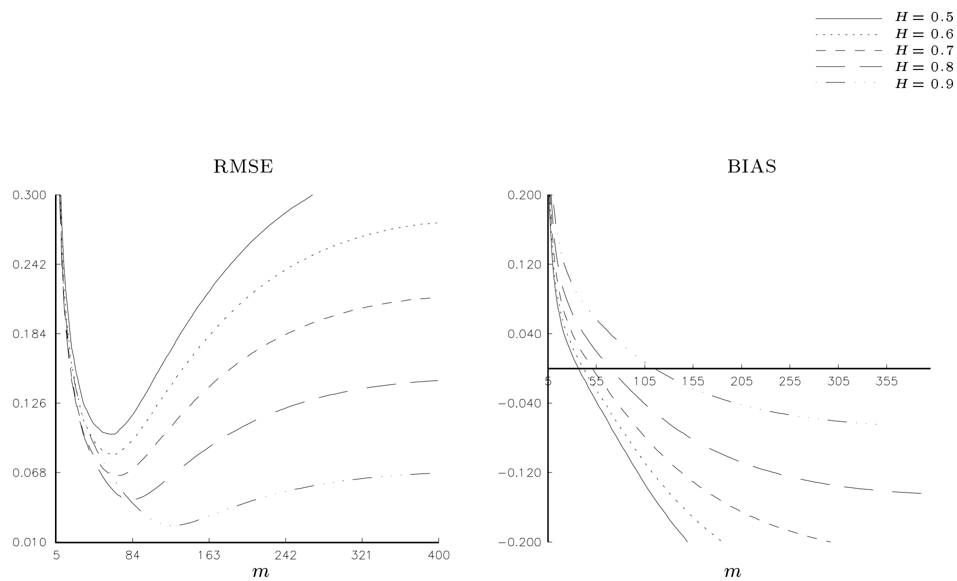

Figure 4. Monte Carlo Bias and RMSE of H estimates in ARFIMA$(1, H - 1/2, 0)$, $a_1 = 0.5$. Sample size $n = 800$, based on 5000 replications.

Table 1 below compares m values minimizing Monte Carlo RMSE of $\hat{H}_{mq}$, $\tilde{m}$ say, and corresponding $m_{\text{opt}}(H)$, for different values of H. The $\tilde{m}$ are significantly greater than $m_{\text{opt}}(H)$, but the RMSEs of $\hat{H}_{m_{\text{opt}}(H)q}$ are fairly close to the minimum achievable RMSE.

Table 1. m values minimizing the RMSE of $\hat{H}_{mq}$ in the Monte Carlo, $\tilde{m}$, versus $m_{\text{opt}}(H)$ in the ARFIMA$(1, H - 1/2, 0)$ with $a_1 = 0.5$, $q = 1/2$, (Monte Carlo RMSE in parenthesis.)

$n = 1000$

	$\tilde{m}$	$m_{\text{opt}}(H)$
0.5	67 (0.0921)	51 (0.1021)
0.6	69 (0.0770)	56 (0.0839)
0.7	86 (0.0623)	67 (0.0634)
0.8	96 (0.0415)	73 (0.0490)
0.9	148 (0.0216)	83 (0.0498)

Once H has been estimated we need to approximate $E_\alpha(H)$, which in general depends on H and, possibly, the parameters explaining the short memory part of the model, as (3.5) indicates. Given $\hat{H}_{mq}$ and a preliminary value of $h''(0)/2h(0)$, $E_2(H)$ can be estimated according to

$$E_2(\hat{H}_{mq}) = h''(0)/2h(0) + (2\hat{H}_{mq} - 1)/24. \tag{4.1}$$

Starting from a pilot value of m, $\hat{m}^{(0)}$ say, $m_{\text{opt}}(H)$ and H can be estimated by the following iterative procedure,

$$\hat{H}_q^{(k+1)} = \hat{H}_{\hat{m}^{(k)}q}, \text{ where } \hat{m}^{(k+1)} = m_{\text{opt}}(\hat{H}_q^{(k+1)}),\ k = 0, 1, \ldots, \tag{4.2}$$

and where (4.1) is used in the computation of $m_{\text{opt}}(H)$ by (2.11) and (2.13).

Tables 2 and 3 below summarize Monte Carlo results for the iterative procedure (4.2), taking $h''(0)/2h(0) = -a_1/(1-a_1)^2$ as known. Convergence is typically achieved after two iterations. The estimates of $m_{\text{opt}}(H)$ values are fairly close to the true ones, and the RMSE are also close to the minimum achievable ones. However, the procedure is not truly automatic since the true value $h''(0)/2h(0)$ is unknown. Table 3 also includes results for a modified version of the estimate of Geweke and Porter-Hudak (1983). The version we use is that of Robinson (1995),

$$\tilde{H} = \frac{1}{2}\left[1 - \frac{\sum_{j=\ell+1}^{m} \log I(\lambda_j)\{\log j - \frac{1}{m-\ell}\sum_{j=\ell+1}^{m}\log j\}}{\sum_{j=\ell+1}^{m}\{\log j - \frac{1}{m-\ell}\sum_{j=\ell+1}^{m}\log j\}^2}\right],$$

where ℓ is a trimming number introduced in this setting by Künsch (1986). We took $\ell = 1, 2, 3$ in the computations for Table 3. We consider $\tilde{H}$ with $m = [n^{1/2}]$, a simple "rule-of-thumb" choice sometimes appearing in the applied literature, in order merely to provide some comparison of our approach with methods popular in that literature. The iterated averaged periodogram is often more biased than $\tilde{H}$ computed with the arbitrary m, but its RMSE's are always less.

Table 2. Monte Carlo mean values of $\hat{m}^{(k)}$ for (4.2) based on 2000 replications of the ARFIMA$(1, H-1/2, 0)$ with $a_1 = 0.5$ and $h''(0)/2h(0)$ known, $\hat{m}^{(0)} = n^{4/5}$, $q = 1/2$.

$n = 1000$

H	0.5	0.6	0.7	0.8	0.9
$\hat{m}^{(1)}$	88	74	79	90	104
$\hat{m}^{(2)}$	55	62	77	78	83
$\hat{m}^{(3)}$	54	61	75	78	81
$\hat{m}^{(\infty)}$	53	60	75	76	81
$m_{\text{opt}}(H)$	51	56	67	73	83

Table 3. Monte Carlo RMSE and BIAS of $\hat{H}_q^{(k)}$ for (4.2) and $\tilde{H}$ based on 2000 replications of the ARFIMA$(1, H-1/2, 0)$ with $a = 0.5$ and $h''(0)/2h(0)$ known, $\hat{m}^{(0)} = n^{4/5}$, $q = 1/2$.

$n = 1000$

		H	0.5	0.6	0.7	0.8	0.9
$\hat{H}_q^{(1)}$		RMSE	0.2561	0.2134	0.1648	0.1085	0.0436
		BIAS	−0.255	−0.219	−0.163	−0.076	−0.0421
$\hat{H}_q^{(2)}$		RMSE	0.1077	0.0792	0.0619	0.0432	0.0379
		BIAS	−0.076	−0.039	−0.025	−0.009	0.021
$\hat{H}_q^{(3)}$		RMSE	0.1017	0.0878	0.0795	0.0455	0.0550
		BIAS	−0.034	−0.027	−0.024	0.002	0.037
$\hat{H}_q^{(\infty)}$		RMSE	0.0991	0.0861	0.0740	0.0499	0.0572
		BIAS	−0.021	−0.019	−0.018	0.001	0.039
	$\ell = 1$	RMSE	0.1591	0.1591	0.1593	0.1594	0.1604
		BIAS	−0.018	−0.018	−0.019	−0.021	−0.024
$\tilde{H}$	$\ell = 2$	RMSE	0.1863	0.1863	0.1868	0.1861	0.1847
		BIAS	−0.021	−0.020	−0.021	−0.022	−0.027
	$\ell = 3$	RMSE	0.2146	0.2145	0.2144	0.2138	0.2119
		BIAS	−0.020	−0.020	−0.020	−0.021	−0.024

It is possible to obtain a more "automatic" m by using an expansion of the semiparametric spectral density

$$f(\lambda) = |1 - e^{i\lambda}|^{1-2H} h(\lambda).$$

Given a pilot m value $\hat{m}^{(0)}$, estimate H by $\hat{H} = \hat{H}_{\hat{m}^{(0)}q}$. Then perform the least squares regression

$$I(\lambda_j) = \sum_{k=0}^{2} Z_{jk}(\hat{H})\hat{\beta}_k + \hat{\varepsilon}_j, \qquad j = 1, \ldots, \hat{m}^{(0)}, \tag{4.3}$$

where $Z_{jk}(H) = |1 - e^{i\lambda} j|^{1-2H} \lambda_j^k / k!$. $\hat{\beta}_0$ and $\hat{\beta}_2$ are estimates of $h(0)$ and $h''(0)$ respectively. Thus $h''(0)/2h(0)$ is estimated by $\hat{\beta}_2/2\hat{\beta}_0$. This estimate is plugged in (4.1) in order to implement the iterative procedure (4.2).

Tables 4 and 5 summarize Monte Carlo results for the feasible estimates of $m_{\text{opt}}(H)$ and corresponding H estimates based on the algorithm (4.2). The $h''(0)/2h(0)$ estimate given by (4.3) is not updated at each iteration. We tried updating it but this made matters worse. The $m_{\text{opt}}(H)$ estimates in Table 4 are more biased than those using the infeasible procedure (Table 2). The H estimates in Table 5 are less efficient than those in Table 3. They are again more efficient than $\tilde{H}$ using the arbitrary $m = [n^{1/2}]$, and usually less biased (see Table 3). The Monte Carlo results do not seem so bad as to eliminate our automatic iterative procedure from practical consideration, and they suggest that further study be directed at theoretically justifying and refining it.

Table 4. Monte Carlo mean values of $\hat{m}^{(k)}$ in procedure (5.5) based on 2000 replications of the ARFIMA$(1, H-1/2, 0)$ with $a_1 = 0.5$ and $h''(0)/2h(0)$ estimated by $\hat{\beta}_2/2\hat{\beta}_0$, and starting value $\hat{m}^{(0)} = n^{4/5}$.

$n = 1000$

H	0.5	0.6	0.7	0.8	0.9
$\hat{m}^{(1)}$	61	53	71	127	227
$\hat{m}^{(2)}$	39	45	70	121	228
$\hat{m}^{(3)}$	39	45	72	124	228
$\hat{m}^{(\infty)}$	38	44	68	118	180
$m_{\text{opt}}(H)$	51	56	67	73	83

Table 5. Monte Carlo RMSE and BIAS of $\hat{H}^{(k)}$ in (4.2) based on 2000 replications of the ARFIMA$(1, H-1/2, 0)$ with $a_1 = 0.5$ and $h''(0)/2h(0)$ estimated by $\hat{\beta}_2/2\hat{\beta}_0$, and starting value $\hat{m}^{(0)} = n^{4/5}$.

$n = 1000$

	H	0.5	0.6	0.7	0.8	0.9
$\hat{H}_q^{(1)}$	RMSE	0.2561	0.2135	0.1648	0.1086	0.0437
	BIAS	−0.254	−0.2122	−0.164	−0.151	−0.042
$\hat{H}_q^{(2)}$	RMSE	0.1125	0.1002	0.0944	0.0946	0.1220
	BIAS	−0.031	−0.003	−0.005	0.008	0.033
$\hat{H}_q^{(3)}$	RMSE	0.1321	0.1171	0.1057	0.0965	0.1171
	BIAS	−0.005	−0.004	−0.008	0.014	0.039
$\hat{H}_q^{(\infty)}$	RMSE	0.1281	0.1170	0.1121	0.1071	0.1104
	BIAS	−0.007	−0.016	−0.016	0.019	0.037

5. Empirical Examples

The minimum water levels of the Nile River measured at the Roda Gorge near Cairo during years 622 through 1284 (see Toussoun (1925)) have been used in several studies dealing with long-range dependence. The periodogram of these data is presented in Figure 6. It is very large at frequencies near zero. In fact using the "automatic" procedure discussed in last section we obtain the estimate $\hat{H} = \hat{H}_{0.5}^{(\infty)} = 0.845$ with optimal $\hat{m} = \hat{m}^{(\infty)} = 70$. Graf (1983) obtained H estimates between 0.828 and 0.847 using a robust but parametric estimate, and Robinson (1994b) obtained semiparametric estimates (using (4.1)) between 0.832 and 0.859 for $m = 20j$, $j = 1, \ldots, 9$.

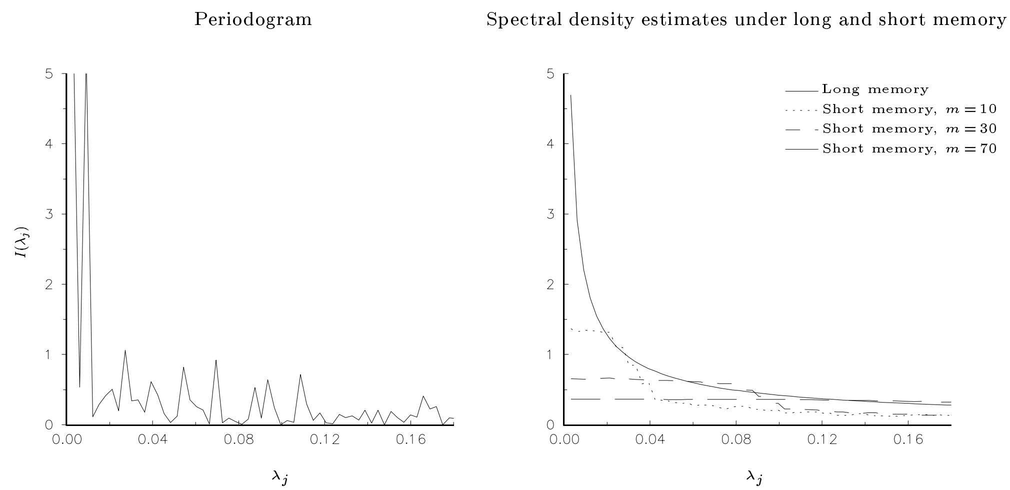

Figure 5. Peridogram and spectral density estimates for the Nile River data.

To estimate $f(\lambda)$ near $\lambda = 0$ we use $\hat{H}$ and $\hat{m}$ in the estimate

$$\hat{G} = \hat{G}_{\hat{H},\hat{m}} = 2(1 - \hat{H})\hat{F}(\lambda_{\hat{m}})\lambda_{\hat{m}}^{2(\hat{H}-1)}, \tag{5.1}$$

proposed by Robinson (1994b). Plugging $\hat{G}$ and $\hat{H}$ into (2.6), we plot in Figure 7

$$\hat{f}(\lambda) = \hat{G}\lambda^{1-2\hat{H}}. \tag{5.2}$$

Now suppose that we proceed more conventionally by estimating $f(\lambda)$ under the assumption that X_t has short memory, that is $f(\lambda)$ is smooth at $\lambda = 0$, so $H = 1/2$ is assumed in (2.6). We estimate $f(\lambda)$ by

$$\tilde{f}(\lambda) = (1 + 2m)^{-1} \sum_{j=-m}^{m} I(\lambda + \lambda_j), \tag{5.3}$$

for $m = 10, 30$ and 70, and plot these estimates in Figure 7. We have used Daniell weights for the sake of comparability with $\hat{f}(\lambda)$; notice that replacing $\hat{H}$ in (5.1) and (5.2) by $1/2$, and $\hat{m}$ by m gives (2.2) and nearly leads to (5.3) for $\lambda = 0$, specifically $\hat{f}(0) = (1 + 1/2m)\tilde{f}(0)$. Although the m used in (5.3) are not larger than $\hat{m}$, and in two cases much smaller, the resolution achieved by (5.2) is much greater, and (5.2) seems the more plausible approximation, bearing in mind Figure 5.

There is evidence of long-memory in economic aggregates. Robinson (1978) and Granger (1980) found that aggregates of Markov processes can, under certain assumptions, have long-memory. Granger and Joyeux (1980) found evidence of long-memory in monthly food inflation rate for the USA economy, where inflation rate is defined as the first difference of the logarithms of consecutive observed price indeces. They estimated H parametrically, proposing different ARFIMAs, and found an H estimate close to 1. Geweke and Porter-Hudak (1983) illustrated their semiparametric estimation procedure (an alternative to (4.1)) by means of an application to monthly consumer price indices. They found strong evidence of long-memory in the inflation rate, with an H estimate close to 1. Geweke and Porter-Hudak (1983) illustrated their semiparametric estimation procedure (an alternative to (4.1)) by means of an application to monthly consumer price indices. They found strong evidence of long-memory in the inflation rate, with an H estimate of 0.923 for the food inflation rate and 1.201 for the aggregate inflation rate, using USA data. Delgado and Robinson (1994) analyzed the monthly aggregate consumer price index for the Spanish economy from July 1939 to October 1991, so $n = 628$. Different semiparametric H estimators were compared, and they always provided values between 0.78 and 0.85 for m values between 75 and 244.

Using the consumer price index data of Delgado and Robinson (1994), H was re-estimated using the automatic method of the previous section, giving $\hat{H} = 0.831$ and $\hat{m} = 140$. Figure 6 presents the periodogram and the estimates (5.2) and (5.3) for this series. The broad comments made about the Nile river data example apply here: the short-memory estimates do not seem sensible in view of Figure 6. We also re-examined the Spanish food inflation rate for the same period, considered by Delgado and Robinson (1994). This more dissaggregated price index produces, as might be expected, a smaller H estimate, $\hat{H} = 0.712$. Here $\hat{m} = 179$, for $n = 628$. Figure 7 presents the periodogram and spectrum estimates for this series, where once again (5.3) fails to reflect the magnitude of the periodogram near zero.

Acknowledgement

This article is based on research funded by the Economic and Social Research Council (ESRC) reference number: R000233609, and by the Spanish "Dirección General de Investigación Científica y Técnica", reference number: PB92-0247. We are grateful for the comments of the referee and associate editor.

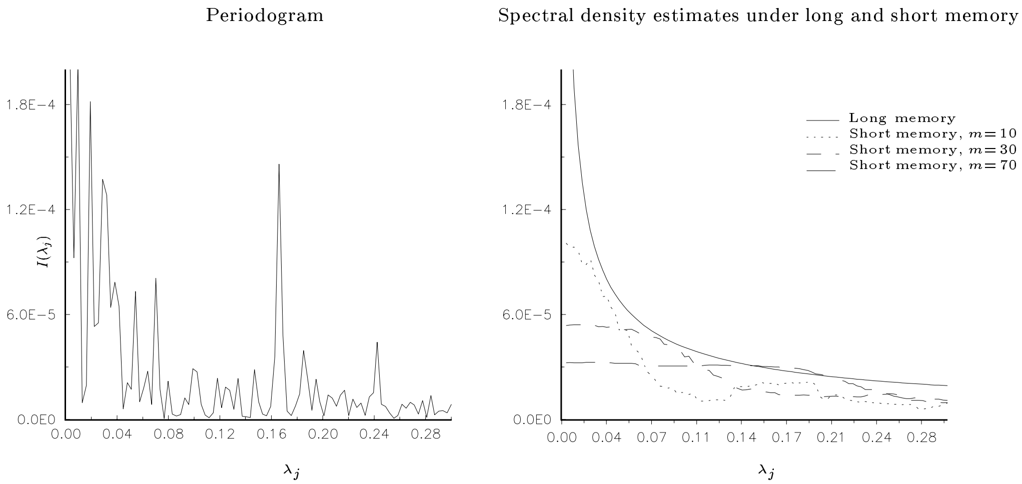

Figure 6. Periodogram and spectral density estimates for aggregate inflation rate (Spain 1939-1991)

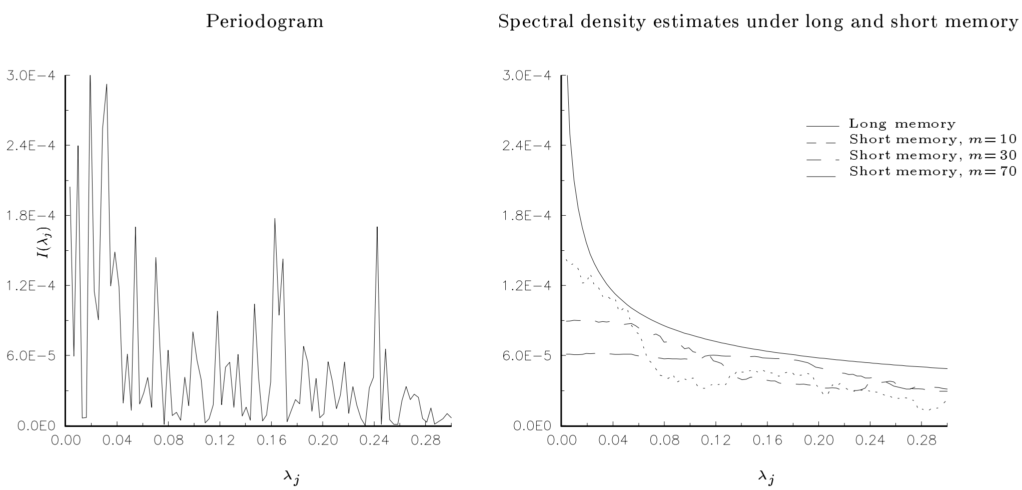

Figure 7. Periodogram and spectral density estimates for food inflation rate (Spain 1939-1991)

References

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

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

Delgado, M. A. and Robinson, P. M. (1994). New methods for the analysis of long-memory: Application to Spanish inflation. J. Forecasting **13**, 97-107.

Geweke, J. and Porter-Hudak S. (1983). The estimation and application of long memory time series models. J. Time Series Anal. **4**, 221-238.

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

Granger, C. W. J. (1980). Long-memory relationships and the aggregation of dynamic models. J. Econom. **14**, 227-238.

Granger, C. W. J. and Joyeux R. (1980). An introduction to long-memory time series models and fractional differencing. J. Time Series Anal. **1**, 15-29.

Hosking, J. R. M. (1981). Fractional differencing. Biometrika **68**, 165-176.

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

Lobato, I. and Robinson P. M. (1995). Average periodrogram estimation of long-memory. J. Econom. forthcoming.

Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. Ann. Math. Statist. **28**, 329-348.

Robinson, P. M. (1978). Statistical in ference for a random coefficient autoregressive model. Scand. J. Statist. **5**, 163-168.

Robinson, P. M. (1983). Review of various approaches to power spectrum estimation. In Handbook of Statistic, Vol. 3, (Edited by D. R. Brillinger and P. R. Krishnaiah), 343-368, Elsevier Science Publishers B.V., New York.

Robinson, P. M. (1994a). Rates of convergence and optimal spectral bandwidth for long range dependence. Probab. Theory Related Fields **99**, 443-473.

Robinson, P. M. (1994b). Semiparametric analysis of long-memory time series. Ann. Statist. **22**, 515-539.

Robinson, P. M. (1994c). Time series with strong dependence. In Advances in Econometrics. Six World Congress, Vol. 1, (Edited by C. A. Sims), 47-95, Cambridge University Press, Cambridge.

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

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

Toussoun, O. (1925), Memoire sur l'histoire du Nil. Mémoires de I'Institut d'Egypte, **9**.

University Carlos III de Madrid, C./ Madrid 126, Getafe, Madrid, Spain.

London School of Economics, Economics Department, Houghton Street, London WC2A 2AE, U.K.

(Received December 1993; accepted September 1995)