9-15-2008

Confidence Intervals for Long Memory Regressions

Kyungduk Ko
Boise State University

Jaechoul Lee
Boise State University

Robert Lund
Clemson University

Confidence Intervals for Long Memory Regressions

Kyungduk Ko[a], Jaechoul Lee[a,*] and Robert Lund[b]

[a] Department of Mathematics, Boise State University

Boise, ID 83725-1555, USA

[b] Department of Mathematical Sciences, Clemson University

Clemson, SC 29634-0975, USA

SUMMARY

This paper proposes an accurate confidence interval for the trend parameter in a linear regression model with long memory errors. The interval is based upon an equivalent sum of squares method and is shown to perform comparably to a weighted least squares interval. The advantages of the proposed interval lies in its relative ease of computation and should be attractive to practitioners.

Keywords: Asymptotic Normality, Linear Regression, Long Memory, Ordinary Least Squares, Weighted Least Squares.

*Correspondence to: Jaechoul Lee, Department of Mathematics, Boise State University, 1910 University Dr, Boise, ID 83725-1555, USA. E-mail: jaechlee@math.boisestate.edu, Phone: 1-208-426-5630, Fax: 1-208-426-1356.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Statistics & Probability Letters, published by Elsevier. Copyright restrictions may apply. doi: 10.1016/j.spl.2008.01.057

1 Introduction

This paper examines confidence intervals for the parameters in the simple linear regression model

$$Y_t = \mu + \beta t + \varepsilon_t, \qquad t = 1, 2, \ldots, n, \tag{1}$$

where $\{Y_t\}$ is a stationary time series, μ and β are unknown regression parameters, and the errors $\{\varepsilon_t\}$ are a zero mean stationary series with long memory in that $\sum_{h=0}^{\infty} |\gamma(h)| = \infty$, where $\gamma(h) = \mathrm{Cov}(Y_t, Y_{t+h})$. Our task is to construct an accurate confidence interval for β with minimal computational burden.

Regression inference with long memory errors has been previously studied. Robinson and Hidalgo (1997) establish a central limit theorem for the weighted least squares (WLS) estimators in regressions with long memory errors. Ordinary least squares (OLS) estimators, though subefficient to WLS estimators in long memory settings, often perform well and were studied in Yajima (1988), who established their asymptotic normality and derived their explicit asymptotic variance. Yajima (1991) and Kleiber (2001) present efficiency relations between WLS and OLS estimators in long memory error settings.

To construct confidence intervals in the practical setting where the errors are governed by unknown long memory parameters, estimates of these parameters are needed. Whittle-type estimators are popular and have desirable properties even when the underlying error model is misspecified (see Taqqu and Teverovsky, 1997). Koul and Surgailis (2000) show that the asymptotic normality of the Whittle estimator depends on the rate of consistency of the regression parameter estimate. As a variant of maximum likelihood, Haslett and Raftery (1989) use the concentrated maximum log-likelihood in the time domain to estimate d. For other estimation methods, see Taqqu *et al.* (1995).

Our goal here is to obtain an approximate $(1 - \alpha) \times 100\%$ confidence interval for β. To do this, we will use an equivalent sum of squares method that employs a closed form expression for the OLS parameter variances to 'adjust' for the long-memory aspects in the series. The explicit computations presented here allow us to account for the effects of a finite sample size in confidence intervals and should have other uses. The performance of our method improves on the OLS methods of Yajima (1988, 1991) and bypasses the computationally demanding WLS methods of Robinson and Hidalgo (1997). Our method merely requires an estimate of the long memory parameters.

The rest of this paper proceeds as follows. We review simple long memory processes and Yajima's OLS asymptotic variance in Section 2. Section 3 identifies the exact variance of the OLS estimators in closed form and clarifies our proposed interval. This interval is studied by simulation in Section 4. Extensions of the methods to several common autoregressive fractionally integrated moving-average (ARFIMA) processes is presented in Section 5.

2 Preliminaries

2.1 Long memory processes

A long memory process is a time series that has a slow decay in its autocovariances: $|\gamma(h)| \sim ch^{-\kappa}$, for $0 < \kappa < 1$ and large h, with $c > 0$ a constant depending on the process. As a result, the autocovariances of long memory process are not absolutely summable (over all lags) and typical regularity conditions for time series limit theorems do not immediately apply.

A simple long memory model can be defined in terms of a fractional difference operator $(1-B)^d$, which is viewed as a general binomial series expansion. For $d \in (-0.5, 0.5)$, we define

$$(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j B^j,$$

with B the backward shift operator ($BX_t = X_{t-1}$) and the square summable coefficients

$$\binom{d}{j}(-1)^j = \frac{\Gamma(d+1)(-1)^j}{\Gamma(d-j+1)\Gamma(j+1)} = \frac{\Gamma(-d+j)}{\Gamma(-d)\Gamma(j+1)}.$$

Here $\Gamma(\cdot)$ denotes the gamma function defined as $\Gamma(v) = \int_0^\infty t^{v-1}e^{-t}dt$ for $v > 0$, $\Gamma(0) = \infty$, and by integration by parts for negative arguments: $v^{-1}\Gamma(1+v)$ for $v < 0$ (Brockwell and Davis 1991).

A fractionally differenced noise, or ARFIMA$(0, d, 0)$ process, $\{X_t\}$ is defined as the solution to the equation $(1-B)^d X_t = Z_t$, where $\{Z_t\}$ is white noise with zero mean and variance σ^2 (see Granger and Joyeux 1980 and Hosking 1981). An ARFIMA$(0,d,0)$ process is stationary and invertible when $-0.5 < d < 0.5$.

The autocorrelation $\rho(h) = \gamma(h)/\gamma(0)$ of an ARFIMA$(0,d,0)$ process is known explicitly as

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(1-d+h)\Gamma(d)}, \qquad h = 1, 2, \ldots, n-1. \tag{2}$$

and the process variance is

$$\gamma(0) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}. \tag{3}$$

For $0 < d < 0.5$, $\{X_t\}$ has long memory with $\kappa = 1 - 2d$ and the autocorrelation of $\{X_t\}$ is positive at every lag and decays hyperbolically to zero with increasing lag. When $d = 0$, $\{X_t\}$ is white noise and therefore has $\gamma(h) = 0$ for all $h \neq 0$. If $-0.5 < d < 0$, $\{X_t\}$ has short memory and the autocorrelations of the process are all negative (with the exception that $\rho(0) = 1$ (Hosking 1981)). Because of this, we work with an ARFIMA$(0,d,0)$ $\{\varepsilon_t\}$ with $0 < d < 0.5$. Generalizations to ARFIMA(p,d,q) error structures are studied in Section 5.

2.2 Asymptotic variance of the OLS estimator

Yajima (1988) derived the asymptotic variance of the OLS estimator of β (see also Koul and Surgailis 2000). For this, let D be a 2×2 diagonal matrix with elements $D_{j,j} = [\sum_{t=1}^{n} t^{2(j-1)}]^{1/2}$ for $j = 1, 2$. Suppose that $S = [s_{k,\ell}]_{k,\ell=1,2}$ and $R = [r_{k,\ell}]_{k,\ell=1,2}$, where

$$s_{k,\ell} = \frac{\sqrt{(2k-1)(2\ell-1)}\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} \int_0^1 \int_0^1 x^{k-1} y^{\ell-1} |x-y|^{2d-1} dx dy$$

and $r_{k,\ell} = \sqrt{(2k-1)(2\ell-1)}/[k+\ell-1]$. Yajima (1988) showed that the asymptotic variances of the OLS estimators are

$$\text{Var}_{\text{asy}}(\hat{\beta}_{\text{OLS}}) = n^{2d} \frac{A_{2,2}}{D_{2,2}^2}, \qquad \text{Var}_{\text{asy}}(\hat{\mu}_{\text{OLS}}) = n^{2d} \frac{A_{1,1}}{D_{1,1}^2},$$

where $A = \sigma^2 R^{-1} S R^{-1}$.

One can base confidence intervals on the above variances and asymptotic normality. In particular, a $(1-\alpha) \times 100\%$ confidence interval for β is simply

$$\hat{\beta}_{\text{OLS}} \pm z_{\alpha/2} \text{Var}_{\text{asy}}(\hat{\beta}_{\text{OLS}})^{1/2},$$

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Statistics & Probability Letters, published by Elsevier. Copyright restrictions may apply. doi: 10.1016/j.spl.2008.01.057

where z_α is the upper αth quantile of the standard normal distribution. This interval requires values for d and σ^2. As we will see later, raw estimates of d can be very biased. Moreover, this bias will degrade interval performance if ignored. We also seek to modify the interval to account for the effects of finite sample sizes.

2.3 Effective sample sizes

For correlated series, effective sample sizes measure the number of independent observations containing a preset quantity of information for a fixed parameter. Such ideas go back to Laurmann and Gates (1977) and are perhaps best illustrated with a simple example. Suppose that $\{Y_t\}$ is stationary in time t with mean μ, variance $\gamma_Y(0)$, and autocorrelation $\rho_Y(h)$ at lag h. Then the sample mean $\bar{Y} = n^{-1}\sum_{t=1}^n Y_t$ has variance

$$\mathrm{Var}(\bar{Y}) = \frac{\gamma_Y(0)}{n}\left[1 + 2\sum_{h=1}^{n-1}\left(1 - \frac{h}{n}\right)\rho_Y(h)\right].$$

The variance of $\bar{Y}$ in the case of independent observations is $\gamma_Y(0)/n$. The ratio of these two variances is set equal to the ratio of the effective sample size, call it n_e, to the sample size n:

$$\frac{\mathrm{Var}(\bar{Y}_{\mathrm{IID}})}{\mathrm{Var}(\bar{Y}_{\mathrm{CORR}})} = \left[1 + 2\sum_{h=1}^{n-1}\left(1 - \frac{h}{n}\right)\rho_Y(h)\right]^{-1} := \frac{n_e}{n},$$

where the subscripts IID and CORR indicate variances in independent and identically distributed and correlated settings, respectively. The idea is that the variance of $\bar{Y}$ is exactly the same in two cases: 1) a series with n_e independent observations each with variance $\gamma_Y(0)$, and 2) a series of n dependent (but stationary) observations with variance $\gamma_Y(0)$ and autocorrelation $\rho_Y(h)$ at lag h. Additional heuristics are developed in Thiébaux and Zwiers (1984) and Lee and Lund (2007).

3 A calibrated OLS interval

Let $Y = (Y_1, Y_2, \ldots, Y_n)'$ denote the data vector and let X be the design matrix with (t, k)th element $x_{t,k} = t^{k-1}$ for $k = 1, 2$. The OLS estimator of $\theta = (\mu, \beta)'$ is

$$\hat{\theta}_{\mathrm{OLS}} = (X'X)^{-1}X'Y,$$

and has variance

$$\text{Var}(\hat{\theta}_{\text{OLS}}) = (X'X)^{-1}X'CX(X'X)^{-1},$$

where C is the $n \times n$ variance/covariance matrix of $\{\varepsilon_t\}_{t=1}^n$ with (i,j)th element $C_{i,j} = \gamma(|i-j|)$ for $i, j = 1, 2, \ldots, n$.

The exact variance/covariance structure of the OLS estimators of μ and β with stationary ARFIMA$(0, d, 0)$ errors is a tedious computation. However, Lee and Lund (2004) present closed form expressions for the variances of $\hat{\mu}_{\text{OLS}}$ and $\hat{\beta}_{\text{OLS}}$ when $\{\varepsilon_t\}$ is a general stationary process. Plugging $\gamma(0)$ and $\rho(h)$ in (2) and (3) into Lee's and Lund's (2004) variance formulas produces

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{12\sigma^2\Gamma(1-2d)}{n(n+1)(n-1)\Gamma^2(1-d)} \left[1 + 2\sum_{h=1}^{n-1} u_h \left\{ \prod_{j=1}^{h} \frac{j-1+d}{j-d} \right\}\right] \tag{4}$$

and

$$\text{Var}_{\text{exact}}(\hat{\mu}_{\text{OLS}}) = \frac{2\sigma^2\Gamma(1-2d)}{n\Gamma^2(1-d)} \left[2 + \frac{3}{n-1} + 2\sum_{h=1}^{n-1} w_h \left\{ \prod_{j=1}^{h} \frac{j-1+d}{j-d} \right\}\right], \tag{5}$$

where $\{u_h\}$ and $\{w_h\}$ are, for $h = 1, 2, \ldots, n-1$,

$$u_h = \frac{(n-h)(n^2-2hn-2h^2-1)}{n(n+1)(n-1)}, \qquad w_h = \frac{(n-h)[2n^2-(3h+1)n-3h^2-1]}{n(n-1)^2}.$$

We have been unable to locate the explicit forms in (4) and (5) elsewhere; these variances margins will have uses beyond confidence intervals. Notice that the weights $\{u_h\}_{h=1}^{n-1}$ and $\{w_h\}_{h=1}^{n-1}$ do not depend on d and σ^2; the variance of $\hat{\beta}_{\text{OLS}}$ increases as d increases. These computations are exact and apply to every sample size n.

For the parameter β, the equivalent sample size is

$$n_e = n\left[\frac{\text{Var}(\hat{\beta}_{\text{OLS, IND}})}{\text{Var}(\hat{\beta}_{\text{OLS, CORR}})}\right] = n\left[1 + 2\sum_{h=1}^{n-1} u_h \left\{ \prod_{j=1}^{h} \frac{j-1+d}{j-d} \right\}\right]^{-1}. \tag{6}$$

A similar n_e could be derived for μ with (5). Observe that $n_e \to \infty$ as $n \to \infty$ and that $n_e \leq n$ when $d \in (0, 0.5)$.

When d and σ^2 are known, we hence propose the interval

$$\hat{\beta}_{\text{OLS}} \pm z_{\alpha/2}\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}})^{1/2} \tag{7}$$

as a good interval. When d and σ^2 are unknown, estimates of these parameters can be substituted into (4) and (5) to obtain estimated variances of the OLS estimators, which we denote as $\widehat{\mathrm{Var}}_{\mathrm{exact}}(\hat{\beta}_{\mathrm{OLS}})$ and $\widehat{\mathrm{Var}}_{\mathrm{exact}}(\hat{\mu}_{\mathrm{OLS}})$. Because μ and β are estimated, the z margins in (7) are modified to t-percentiles. In short, our 'calibrated' interval is

$$\hat{\beta}_{\mathrm{OLS}} \pm t_{\alpha/2,\hat{n}_e-2}\widehat{\mathrm{Var}}_{\mathrm{exact}}(\hat{\beta}_{\mathrm{OLS}})^{1/2}, \tag{8}$$

where $t_{\alpha,\hat{n}_e-2}$ is the upper αth quantile of the Student's t-distribution with $\hat{n}_e-2$ degrees of freedom, and $\hat{n}_e$ is as in (6) with estimates of d and σ^2 plugged in for their true values. This interval uses the exact OLS variance for each sample size n, which should help performance for smaller sample sizes. The proposed interval also avoids the computational demanding task of calculating the inverse of the covariance matrix of a long memory error series, which is needed to obtain a WLS estimate. In fact, taking this inverse may not be numerically feasible for large n. As long memory is a ubiquitous property for series that are sampled frequently in time (in which case n is usually large), this is an important practical point. As we will see in the next section, the interval in (8) performs just as well as WLS intervals in simulations.

4 A Simulation Study

We now study the performance of our interval for the sample sizes $n = 50, 100, 200, 500, 1000$, and the long memory parameters $d = 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45$. The other parameters in the regression model were taken as $\mu = 3.5$, $\beta = 0.5$, and $\sigma^2 = 1$.

To proceed further, we need an estimate of the long memory parameter d and the innovation variance σ^2. While we do not wish to favor any particular estimation method, we will examine both the Haslett and Raftery (1989) and Whittle-type estimators. The Haslett-Raftery estimator maximizes approximate concentrated log-likelihoods and the Whittle estimator involves the periodogram of the series (see (3.2) and (3.3) in Taqqu and Teverovsky 1997). These estimates enjoy practical popularity and are easy to compute — the one line commands fracdiff (or arima.fracdiff in S-Plus) and farisma or whittleFit in R are such ways. Unfortunately, these estimators of d are also biased. Table 1 shows sample average biases and standard deviations, computed from ten thousand independent simulations. The biases are all negative and some are quite large, particularly for small n and d slightly below 0.5. Overall, the Whittle estimates are less biased, but have a slightly larger standard deviation. We have found that neglecting this bias will produce untrustable confidence intervals, regardless of the method the interval is based upon (OLS or WLS); this point is illustrated below. An explicit expression for this bias is not immediately obvious to us, as in fact the concentrated log-likelihood and Whittle methods require a numerical optimization, which is typically done by Newton's method. However, it is a simple task to bias adjust the estimators of d with the values in Table 1. In the event that the bias adjusted estimate of d exceeds 0.5, we simply interpret our estimate of d as 0.5. After d is estimated, σ^2 is estimated as the sample variance in the fractionally differenced OLS residuals $\{(1-B)^{\hat{d}}(Y_t - \hat{\mu}_{\text{OLS}} - \hat{\beta}_{\text{OLS}} t)\}$. Estimating σ^2 is not as problematic as estimating d.

Tables 2 and 3 summarize our simulations. In both tables, 'CAL' refers to the proposed 'calibrated' interval using equivalent sample sizes, 'ASY' as Yajima's asymptotic OLS interval, and 'WLS' as the gold standard best linear estimation method based on asymptotic normality. To generate the tables, ten thousand simulations were run for various d and n, and confidence intervals were computed for 1) our 'calibrated' interval in (8), 2) Yajima's asymptotically normal interval described in Section 2.2, and 3) an optimal WLS interval. Haslett-Raftery and Whittle estimates of d were obtained from the OLS residuals $\{Y_t - \hat{\mu}_{\text{OLS}} - \hat{\beta}_{\text{OLS}} t\}$ and the bias corrected $\hat{d}$ was used in all three intervals.

Table 2 summarizes the coverage probabilities of the three methods for a 95% interval. With the bias corrected estimate of d, all three methods are working well, with our calibrated method producing empirical coverage probabilities slightly closer to the target level of 0.95 than the other methods, especially for small n. This is attributable to the finite sample size correction (6) with the exact computations in (4). As a convention, we assume that any interval where the bias adjusted estimate of d exceeds 0.5 includes the trend parameter β. This is because the data is suggesting that the error margin in (4) is infinite. The bias corrections are very important. For example, when

$d = 0.45$ and $n = 200$, Haslett-Raftery estimate bias corrected and non-biased corrected intervals have coverage probabilities (CAL 0.9514/0.8852), (ASY 0.9385/0.8617), and (WLS 0.9387/0.8693) respectively. The coverages of the non-biased corrected intervals are far too low to be considered reliable. Confidence intervals based on the bias-corrected Whittle estimator perform slightly worse than intervals based on the bias-corrected Haslett-Raftery estimate, but not drastically so. Because we are bias-correcting the estimate of d, it is not surprising that the estimator of d with a smaller variance is preferable. In cases where the Whittle estimate of d was too small (which due to its higher variability happens slightly more frequently than with the Haslett-Raftery estimator), the error margin of the interval was underestimated and the target parameter β was not in the interval as frequently as it needed to be.

Table 3 reports the average length of the confidence intervals in Table 2, conditional on the bias adjusted estimate of d being less than 0.5. Of course, the WLS interval has the shortest length; however, the length of the calibrated interval is quite competitive across the board. Finally, Table 4 reports the percent of times the bias corrected d exceeds 0.5. Of course, this percent decreases as n increases. These percentages are slightly higher for the Whittle estimator, which is attributed to its larger variance.

Overall, the calibrated interval appears to function well in small and moderate sample size settings, and retains very good asymptotic properties. It bypasses the need to take an inverse of an $n \times n$ long memory variance/covariance matrix, which is required to obtain the WLS interval. The accuracy of the WLS method is, however, approximately retained.

5 Extensions

Although we have focused on a simple linear regression model with ARFIMA$(0, d, 0)$ errors, the proposed method can be extended to polynomial or multiple regression models with ARFIMA$(0, d, 0)$ errors. We will not pursue this here.

General ARFIMA(p, d, q) long memory errors can also be handled. Sowell (1992) derives the autocovariance function of such models as follows. If $\{X_t\}$ is a solution to the ARFIMA(p,d,q) difference equation

$$\Phi(B)(1-B)^d X_t = \Theta(B) Z_t,$$

where $\Phi(B) = 1-\phi_1 B - \cdots - \phi_p B^p$, $\Theta(B) = 1+\theta_1 B + \cdots + \theta_q B^q$, and $\{Z_t\}$ is zero mean white noise with variance σ^2 (assume that Φ and Θ are causal and have no common zeroes), then Sowell (1992) derives $\gamma(h)$ explicitly and Doornik and Ooms (2003) express it in the numerically stable form

$$\gamma(h) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)} \sum_{k=-q}^{q} \sum_{j=1}^{p} \psi_k \tilde{\zeta}_j \tilde{C}(d, p+k-h, \rho_j) \frac{(d)_{p+k-h}}{(1-d)_{p+k-h}},$$

where $\psi_k = \sum_{s=|k|}^{q} \theta_s \theta_{s-|k|}$ $(\theta_0 = 1)$, $\rho_1, \ldots, \rho_p$ are the p roots of the AR polynomial Φ,

$$\tilde{\zeta}_j^{-1} = \prod_{i=1}^{p} (1-\rho_i \rho_j) \prod_{\substack{m=1 \\ m \neq j}}^{p} (\rho_j - \rho_m),$$

$(a)_i$ is Pochhammer's symbol defined as $(a)_i = \Gamma(a+i)/\Gamma(a)$, and

$$\tilde{C}(d,l,\rho) = \rho^{2p} G(d+l; 1-d+l; \rho) + \rho^{2p-1} + G(d-l; 1-d-l; \rho)$$

with $G(a;b;\rho) = \sum_{i=0}^{\infty} (a)_{i+1} \rho^i / (b)_{i+1}$.

The autocovariance functions of several ARFIMA models can be explicitly extracted from the above. In particular, if $\{X_t\}$ is an ARFIMA$(0,d,q)$ series, then Sowell (1992) obtains

$$\gamma(h) = \frac{\sigma^2 \Gamma(1-2d)}{\Gamma^2(1-d)} \sum_{k=-q}^{q} \psi_k \frac{(d)_{k-h}}{(1-d)_{k-h}}.$$

In the special case where $q = 1$, we have

$$\gamma(h) = \frac{\sigma^2 (1+\theta_1^2) \Gamma(1-2d)}{\Gamma^2(1-d)} \left\{ 1 + \frac{2\theta_1}{1+\theta_1^2} \left(\frac{d(1-d)-h^2}{(1-d)^2 - h^2} \right) \right\} \frac{(d)_h}{(1-d)_h}.$$

If $\{X_t\}$ is ARFIMA$(p,d,0)$, then

$$\gamma(h) = \frac{\sigma^2 \Gamma(1-2d)}{\Gamma^2(1-d)} \frac{(d)_{p-h}}{(1-d)_{p-h}} \sum_{j=1}^{p} \tilde{\zeta}_j \tilde{C}(d, p-h, \rho_j).$$

When $p = 1$, we obtain

$$\gamma(h) = \frac{\sigma^2 \Gamma(1-2d)}{(1-\phi_1^2)\Gamma^2(1-d)} \frac{(d)_{1-h}}{(1-d)_{1-h}} \tilde{C}(d, 1-h, \phi_1).$$

These results can be used to derive the exact variance of $\hat{\beta}_{\text{OLS}}$. In the regression model (1) with $\{\varepsilon_t\}$ as ARFIMA(p, d, q), use of the OLS variance expressions in Lee and Lund (2004) gives

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{\sigma^2\Gamma(1-2d)}{\sum_{t=1}^{n}(t-\bar{t})^2\Gamma^2(1-d)} \sum_{h=0}^{n-1} \tilde{u}_h \sum_{k=-q}^{q} \psi_k \left\{ \sum_{j=1}^{p} \tilde{\zeta}_j \tilde{C}(d, p+k-h, \rho_j) \right\} \frac{(d)_{p+k-h}}{(1-d)_{p+k-h}},$$

where $\bar{t} = (n+1)/2$, $\tilde{u}_0 = 1$ and $\tilde{u}_h = 2u_h$ for $1 \le h \le n-1$.

For the errors as an ARIMA$(0, d, q)$ series, this gives

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{\sigma^2\Gamma(1-2d)}{\sum_{t=1}^{n}(t-\bar{t})^2\Gamma^2(1-d)} \sum_{h=0}^{n-1} \tilde{u}_h \left\{ \sum_{k=-q}^{q} \psi_k \frac{(d)_{k-h}}{(1-d)_{k-h}} \right\},$$

which can be manipulated into

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{\sigma^2(1+\theta_1^2)\Gamma(1-2d)}{\sum_{t=1}^{n}(t-\bar{t})^2\Gamma^2(1-d)} \sum_{h=0}^{n-1} \tilde{u}_h \left\{ 1 + \frac{2\theta_1}{1+\theta_1^2} \left(\frac{d(1-d)-h^2}{(1-d)^2-h^2} \right) \right\} \frac{(d)_h}{(1-d)_h}$$

when $q = 1$.

When $\{\varepsilon_t\}$ is ARFIMA$(p, d, 0)$, we have

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{\sigma^2\Gamma(1-2d)}{\sum_{t=1}^{n}(t-\bar{t})^2\Gamma^2(1-d)} \sum_{h=0}^{n-1} \tilde{u}_h \left\{ \sum_{j=1}^{p} \tilde{\zeta}_j \tilde{C}(d, p-h, \rho_j) \right\} \frac{(d)_{p-h}}{(1-d)_{p-h}},$$

which reduces to

$$\text{Var}_{\text{exact}}(\hat{\beta}_{\text{OLS}}) = \frac{\sigma^2\Gamma(1-2d)}{\sum_{t=1}^{n}(t-\bar{t})^2(1-\phi_1^2)\Gamma^2(1-d)} \sum_{h=0}^{n-1} \tilde{u}_h \tilde{C}(d, 1-h, \phi_1) \frac{(d)_{1-h}}{(1-d)_{1-h}}$$

when $p = 1$.

The equivalent sample size n_e for the models in this section proceeds as was done in (6) for an ARFIMA$(0, d, 0)$ series. We omit these computations.

Acknowledgments

The authors thank an anonymous referee for several valuable comments that substantially improved the paper.

References

Beran, J. (1994), *Statistics for Long-memory Processes* (Chapman and Hall, New York).

Brockwell, P.J. and Davis, R.A. (1991), *Time Series: Theory and Methods* (Springer, New York, 2nd ed.).

Doornik, J.A. and Ooms, M. (2003), Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models, *Comput. Statist. Data Anal.* **42**, 333–348.

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

Haslett, J. and Raftery, A. (1989), Space-time modelling with long-memory dependence: assessing Ireland's wind power resource, invited paper with discussion, *Appl. Statist.* **38**, 1–50.

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

Kleiber, C. (2001), Finite sample efficiency of OLS in linear regression models with long-memory disturbances, *Econom. Lett.* **72**, 131–136.

Koul, H.L. and Surgailis D. (2000), Asymptotic normality of the Whittle estimator in linear regression models with long memory errors, *Statist. Infer. Stoch. Proc.* **3**, 129–147.

Laurmann, J.A. and Gates, W.L. (1977), Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models, *J. Atmos. Sci.* **34**, 1187–1199.

Lee, J. and Lund, R.B. (2004), Revisiting simple linear regression with autocorrelated errors, *Biometrika* **91**, 240–245.

Lee, J. and Lund, R.B. (2007), Equivalent sample sizes in time series regressions, *J. Statist. Comput. Sim.*, to appear.

Robinson, P.M. and Hidalgo, F.J. (1997), Time series regression with long-range dependence, *Ann. Statist.* **25**, 77–104.

Sowell, F. (1992), Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *J. Economet.* **53**, 165–188.

Taqqu, M.S., Teverovsky, V. and Willinger, W. (1995), Estimators for long-range dependence: an empirical study, *Fractals* **3**, 785–798.

Taqqu, M.S. and Teverovsky, V. (1997), Robustness of Whittle-type estimators for time series with long-range dependence, *Stoch. Models* **13**, 723–757.

Thiébaux, H.J. and Zwiers, F.W. (1984), The interpretation and estimation of effective sample size, *J. Climate Appl. Meteor.* **23**, 800–811.

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.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Statistics & Probability Letters, published by Elsevier. Copyright restrictions may apply. doi: 10.1016/j.spl.2008.01.057

		Haslett-Raftery					Whittle				
		n					n				
d		50	100	200	500	1000	50	100	200	500	1000
0.05	Bias	-.0278	-.0231	-.0184	-.0127	-.0078	-.0212	-.0194	-.0164	-.0119	-.0074
	SD	.0499	.0450	.0397	.0316	.0247	.0597	.0487	.0409	.0319	.0247
0.10	Bias	-.0614	-.0481	-.0334	-.0168	-.0089	-.0517	-.0427	-.0305	-.0158	-.0086
	SD	.0656	.0625	.0530	.0377	.0259	.0769	.0663	.0542	.0378	.0259
0.15	Bias	-.0920	-.0651	-.0393	-.0174	-.0087	-.0790	-.0577	-.0359	-.0163	-.0082
	SD	.0806	.0755	.0603	.0375	.0260	.0935	.0797	.0612	.0377	.0262
0.20	Bias	-.1162	-.0749	-.0408	-.0183	-.0096	-.0995	-.0657	-.0366	-.0169	-.0089
	SD	.0955	.0848	.0618	.0375	.0256	.1095	.0894	.0631	.0379	.0260
0.25	Bias	-.1368	-.0810	-.0423	-.0177	-.0100	-.1168	-.0696	-.0372	-.0157	-.0089
	SD	.1073	.0907	.0625	.0368	.0258	.1224	.0960	.0639	.0374	.0260
0.30	Bias	-.1511	-.0863	-.0429	-.0185	-.0101	-.1259	-.0726	-.0363	-.0156	-.0085
	SD	.1169	.0920	.0631	.0372	.0256	.1340	.0987	.0658	.0380	.0260
0.35	Bias	-.1663	-.0909	-.0469	-.0202	-.0106	-.1365	-.0735	-.0383	-.0161	-.0081
	SD	.1223	.0916	.0605	.0366	.0255	.1409	.1004	.0645	.0379	.0260
0.40	Bias	-.1781	-.0963	-.0491	-.0218	-.0116	-.1437	-.0736	-.0363	-.0155	-.0078
	SD	.1241	.0872	.0586	.0355	.0255	.1442	.0985	.0649	.0379	.0265
0.45	Bias	-.1932	-.1049	-.0577	-.0267	-.0139	-.1541	-.0769	-.0397	-.0160	-.0070
	SD	.1242	.0814	.0539	.0329	.0233	.1444	.0939	.0619	.0375	.0258

Table 1: Simulated biases and standard deviations (SD) of estimated d

		Haslett-Raftery					Whittle				
		n					n				
d		50	100	200	500	1000	50	100	200	500	1000
0.05	CAL	.9511	.9461	.9409	.9432	.9425	.9504	.9453	.9407	.9431	.9425
	ASY	.9408	.9397	.9383	.9418	.9420	.9396	.9391	.9384	.9417	.9420
	WLS	.9442	.9414	.9383	.9427	.9415	.9436	.9411	.9388	.9425	.9419
0.10	CAL	.9498	.9396	.9423	.9432	.9443	.9460	.9372	.9428	.9426	.9441
	ASY	.9369	.9320	.9397	.9416	.9435	.9339	.9304	.9394	.9414	.9433
	WLS	.9408	.9334	.9415	.9419	.9432	.9371	.9321	.9416	.9414	.9429
0.15	CAL	.9436	.9380	.9368	.9444	.9426	.9403	.9357	.9363	.9439	.9424
	ASY	.9309	.9295	.9317	.9423	.9416	.9260	.9282	.9321	.9419	.9410
	WLS	.9334	.9322	.9332	.9411	.9430	.9303	.9301	.9337	.9411	.9425
0.20	CAL	.9409	.9323	.9375	.9446	.9459	.9357	.9305	.9366	.9451	.9458
	ASY	.9269	.9224	.9328	.9419	.9432	.9212	.9215	.9319	.9422	.9431
	WLS	.9281	.9241	.9334	.9412	.9427	.9233	.9232	.9330	.9412	.9429
0.25	CAL	.9365	.9283	.9346	.9427	.9500	.9296	.9267	.9337	.9429	.9503
	ASY	.9170	.9189	.9263	.9386	.9472	.9105	.9173	.9256	.9378	.9473
	WLS	.9233	.9200	.9277	.9361	.9446	.9157	.9181	.9269	.9360	.9449
0.30	CAL	.9334	.9382	.9416	.9464	.9509	.9246	.9346	.9399	.9462	.9509
	ASY	.9134	.9250	.9321	.9387	.9461	.9073	.9199	.9312	.9378	.9459
	WLS	.9166	.9248	.9324	.9392	.9449	.9109	.9220	.9307	.9389	.9450
0.35	CAL	.9335	.9357	.9419	.9520	.9529	.9273	.9316	.9402	.9508	.9526
	ASY	.9148	.9226	.9298	.9416	.9446	.9064	.9156	.9266	.9415	.9445
	WLS	.9192	.9226	.9308	.9411	.9464	.9135	.9181	.9290	.9412	.9467
0.40	CAL	.9372	.9424	.9462	.9525	.9544	.9296	.9352	.9422	.9514	.9543
	ASY	.9177	.9256	.9335	.9417	.9427	.9077	.9192	.9300	.9409	.9427
	WLS	.9206	.9270	.9321	.9395	.9455	.9130	.9219	.9289	.9389	.9450
0.45	CAL	.9375	.9406	.9514	.9562	.9552	.9274	.9335	.9477	.9544	.9539
	ASY	.9216	.9279	.9385	.9462	.9451	.9087	.9209	.9347	.9447	.9442
	WLS	.9233	.9269	.9387	.9454	.9445	.9126	.9209	.9353	.9435	.9434

Table 2: Empirical coverage probabilities of the calibrated OLS interval (CAL), Yajima's asymptotic OLS interval (ASY), and the feasible WLS interval (WLS) when $1-\alpha = 0.95$.

		Haslett-Raftery					Whittle				
		n					n				
d		50	100	200	500	1000	50	100	200	500	1000
0.05	CAL	.0449	.0161	.0059	.0015	.0006	.0452	.0162	.0059	.0015	.0006
	ASY	.0427	.0157	.0058	.0015	.0006	.0429	.0157	.0058	.0015	.0006
	WLS	.0433	.0158	.0058	.0015	.0006	.0435	.0158	.0058	.0015	.0006
0.10	CAL	.0508	.0190	.0071	.0020	.0007	.0511	.0191	.0071	.0020	.0007
	ASY	.0478	.0184	.0070	.0019	.0007	.0480	.0184	.0070	.0019	.0007
	WLS	.0484	.0184	.0070	.0019	.0007	.0485	.0185	.0070	.0019	.0007
0.15	CAL	.0585	.0228	.0087	.0025	.0010	.0587	.0229	.0088	.0025	.0010
	ASY	.0544	.0218	.0085	.0024	.0010	.0546	.0219	.0085	.0024	.0010
	WLS	.0547	.0217	.0085	.0024	.0009	.0549	.0218	.0085	.0024	.0009
0.20	CAL	.0673	.0270	.0107	.0032	.0013	.0673	.0271	.0107	.0032	.0013
	ASY	.0619	.0255	.0103	.0031	.0013	.0619	.0256	.0103	.0031	.0013
	WLS	.0620	.0253	.0102	.0031	.0012	.0619	.0254	.0102	.0031	.0012
0.25	CAL	.0766	.0321	.0130	.0040	.0017	.0754	.0322	.0131	.0040	.0017
	ASY	.0699	.0301	.0125	.0039	.0016	.0690	.0301	.0125	.0039	.0016
	WLS	.0696	.0296	.0122	.0039	.0016	.0687	.0297	.0123	.0039	.0016
0.30	CAL	.0853	.0381	.0160	.0052	.0022	.0832	.0379	.0161	.0052	.0022
	ASY	.0776	.0354	.0152	.0050	.0022	.0759	.0353	.0152	.0050	.0022
	WLS	.0769	.0346	.0148	.0049	.0021	.0751	.0345	.0148	.0049	.0021
0.35	CAL	.0931	.0441	.0198	.0067	.0030	.0898	.0433	.0198	.0067	.0030
	ASY	.0848	.0409	.0186	.0064	.0029	.0819	.0402	.0186	.0064	.0029
	WLS	.0836	.0397	.0180	.0062	.0028	.0808	.0391	.0180	.0062	.0028
0.40	CAL	.0994	.0495	.0237	.0086	.0040	.0961	.0481	.0233	.0086	.0040
	ASY	.0907	.0460	.0223	.0082	.0038	.0876	.0447	.0219	.0082	.0038
	WLS	.0891	.0446	.0214	.0078	.0036	.0862	.0433	.0211	.0078	.0036
0.45	CAL	.1040	.0534	.0273	.0108	.0052	.1004	.0516	.0266	.0106	.0052
	ASY	.0952	.0500	.0259	.0103	.0050	.0918	.0481	.0251	.0102	.0050
	WLS	.0933	.0482	.0248	.0098	.0048	.0901	.0465	.0241	.0097	.0048

Table 3: Average lengths of the calibrated OLS interval (CAL), the Yajima's asymptotic OLS interval (ASY), and the feasible WLS interval (WLS).

	Haslett-Raftery					Whittle				
	n					n				
d	50	100	200	500	1000	50	100	200	500	1000
0.05	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000
0.10	.0000	.0000	.0000	.0000	.0000	.0005	.0000	.0000	.0000	.0000
0.15	.0001	.0000	.0000	.0000	.0000	.0023	.0001	.0000	.0000	.0000
0.20	.0019	.0001	.0000	.0000	.0000	.0099	.0004	.0000	.0000	.0000
0.25	.0144	.0006	.0000	.0000	.0000	.0363	.0034	.0000	.0000	.0000
0.30	.0507	.0035	.0000	.0000	.0000	.0827	.1280	.0003	.0000	.0000
0.35	.1273	.0319	.0015	.0000	.0000	.1686	.0605	.0085	.0000	.0000
0.40	.2408	.1190	.0231	.0004	.0000	.2637	.1598	.0547	.0034	.0000
0.45	.3988	.3058	.1861	.0402	.0033	.4072	.3341	.2391	.0885	.0242

Table 4: Proportion of bias-corrected $\hat{d} > 0.5$