The Stata Journal (2020)
20, Number 3, pp. 565–583 DOI: 10.1177/1536867X20953569

Local Whittle estimation of the long-memory parameter

Christopher F. Baum
Department of Economics
Boston College
Chestnut Hill, MA
baum@bc.edu

Stan Hurn
School of Economics and Finance
Queensland University of Technology
Brisbane, Australia
s.hurn@qut.edu.au

Kenneth Lindsay
Department of Mathematics
University of Glasgow
Glasgow, UK
kenneth.lindsay@glasgow.ac.uk

Abstract. In this article, we describe and implement the local Whittle and exact local Whittle estimators of the order of fractional integration of a time series.

Keywords: st0609, whittle, Whittle estimator, long memory, fractional integration

1 Introduction

Many time series exhibit too much long-range dependence to be classified as stationary or $I(0)$ processes but do not exhibit the infinite memory of a nonstationary process or $I(1)$ process. Time series that display long memory are known as fractionally integrated series, or $I(d)$, where d is no longer an integer and falls in the interval $(-1/2 < d < 1)$ but excludes 0. These processes have autocorrelations that decay more slowly than those of stationary processes, but the pattern of the decay differs from that of an integrated process (Granger and Joyeux 1980; Hosking 1981). Once the restriction that d takes only integer values is relaxed, the ARFIMA(p, d, q) class of model is introduced, where FI stands for "fractional integration". Fractionally integrated $I(d)$ time series have attracted the attention of empirical researchers because long memory provides a suitable description of the characteristics of economic and financial data and because it provides a useful extension to the $I(0)$ and $I(1)$ dichotomy.

The practical econometric problem posed by the concept of long memory is that of estimating the appropriate fractional difference parameter d from a long-memory process y_t. There are two broad approaches to the estimation of d. A classical time-series approach is to specify the full ARFIMA(p, d, q) model and estimate all the parameters, including d, by maximum likelihood (Sowell 1992). The Stata command `arfima` (see [TS] **arfima**) implements the full maximum-likelihood estimation of the ARFIMA(p, d, q) model, requiring specification of the orders of the AR(p) and MA(q) polynomials. A second approach to providing consistent and asymptotically normal estimates of the fractional difference parameter, d, without fully specifying the ARMA components of the model, involves shifting from the time domain to the frequency domain. Once an estimate of d is available, an ARMA model is then fit to the fractionally differenced series to obtain consistent estimators of the remaining model parameters. It is this latter approach with which this article is concerned.

A number of well-known estimators of d in the frequency domain do not require the specification of the full ARFIMA model. These include the estimators due to Geweke and Porter-Hudak (1983), Phillips (1999, 2007), and Robinson (1995b). These estimators are implemented in Stata using the community-contributed commands `gphudak`, `modlpr`, and `roblpr`, respectively. For a full discussion of the estimators and their implementation in Stata, see Baum and Wiggins (2000). These estimators of d are all essentially regression based. In this article, we implement the Whittle likelihood-based approach to estimating the fractional difference parameter d. Both the local Whittle and exact local Whittle estimators are provided in the new command `whittle`.

2 Whittle estimation

Maximum likelihood estimation of the parameter d of a fractionally integrated time series in the frequency domain is based on the approximation to a Gaussian likelihood introduced by Whittle (1951). See also Fox and Taqqu (1986), Whittle (1962), and Choudhuri, Ghosal, and Roy (2004). The popularity of the frequency domain approach stems from the fact that unlike the time domain estimator, the frequency domain maximum-likelihood estimator is invariant to the unknown mean of the process (Cheung and Diebold 1994).

2.1 The Whittle likelihood

Consider a sample of T observations of a stationary centered process $y_1, \ldots, y_T$ uniformly spaced in the time domain and a sequence of m frequencies,

$$\omega_j = \frac{2\pi j}{T} \qquad \text{for } j = 1, 2, \ldots, m$$

These frequencies, with $m \ll T$, represent a set of angular frequencies that are all multiples of the fundamental frequency $2\pi/T$, so called because it corresponds to a single oscillation with period T.[1] The discrete Fourier transform of y_t is given by

$$\widehat{c}(\omega_j) = \frac{1}{\sqrt{2\pi T}} \sum_{k=1}^{T} y_k \, e^{i\omega_j k} \tag{1}$$

The Whittle likelihood follows from the fact that the coefficients $\widehat{c}(\omega_j)$ of y_t are asymptotically independent Gaussian random variables with mean value zero and variance

1. In a sampling context, frequencies with $k > T/2$ cannot be identified because of aliasing. For a good discussion of aliasing, see Press et al. (1992).

given by the spectral density of the process at that frequency. Consequently, the likelihood function at frequency ω_j is

$$L_j = \frac{1}{\sqrt{2\pi f_y(\omega_j)}} \exp\left\{-\frac{I(\omega_j)}{2f_y(\omega_j)}\right\} \tag{2}$$

in which $I(\omega_j)$ is the sample periodogram given by

$$I(\omega_j) = |c(\omega_j)|^2$$

and $f_y(\omega_j)$ is the spectral density at ω_j,

$$f_y(\omega_j) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega_j k}$$

where γ_k denotes the autocovariance at lag k.

Local Whittle estimation of the fractional differencing parameter, d, starts from the recognition that the behavior of the spectral density of y_t at low frequencies is defined by the condition

$$\lim_{\omega\to 0^+} \omega^{2d} f_y(\omega) = G \tag{3}$$

where G is a positive quantity that depends upon the parameter d. The process y_t has finite power provided $2d < 1$, and for this reason, d is taken as a measure of the long-term duration of the memory of process y_t. If $d \in [0, 0.5)$, the series is still covariance stationary, but the autocorrelations disappear more slowly than in the $I(0)$ case. In fact, they decay *hyperbolically* to zero by contrast with the faster, geometric decay of a stationary ARMA process. For $d \in [0.5, 1)$, the process is mean reverting, although it is not covariance stationary because there is no long-run impact of an innovation on future values of the process. Granger and Joyeux (1980) show that a process is in fact nonstationary for $d \geq 0.5$ because it possesses infinite variance. For $d \in (-0.5, 0]$, the process is said to exhibit intermediate memory (antipersistence), or long-range negative dependence.

When $f_y(\omega_j)$ is replaced by its asymptotic approximation $G\omega_j^{-2d}$ from (3), the negative of the log likelihood at ω_j satisfies

$$-\log L_j(G,d) = \frac{1}{2}\left\{\log 2\pi + \log G - 2d\log\omega_j + \frac{1}{G}\omega_j^{2d}|\widehat{c}(\omega_j)|^2\right\}$$

It is this expression that forms the basis of the local Whittle estimators of d.

2.2 The local Whittle estimator

The local Whittle estimator of d analyzed by Robinson (1995a), motivated by the approach of Künsch (1987) and based on the m lowest frequencies $\omega_1, \ldots, \omega_m$, is obtained by minimizing the negative log-likelihood function

$$-\log L(G,d) = \frac{m}{2}\left\{\log 2\pi + \log G - \frac{2d}{m}\sum_{j=1}^{m}\log\omega_j + \frac{1}{G}\frac{1}{m}\sum_{j=1}^{m}\omega^{2d}\,|\widehat{c}(\omega_j)|^2\right\} \tag{4}$$

Partial differentiation of (4) with respect to G demonstrates that the optimal value of G for any value of d is[2]

$$\widehat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}\;|\,\widehat{c}(\omega_j)\,|^2$$

Given this result, it is clear that the local Whittle estimator of d, $\widehat{d}$, minimizes

$$R(d) = \log\widehat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j \qquad \widehat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}I(\omega_j) \tag{5}$$

As demonstrated in the appendix, it is straightforward to prove that $R(d)$ is a convex function of d. Therefore, $R(d)$ will take its minimum value at an interior point of the interval $[d_0, d_1]$ provided $R'(d_0)\times R'(d_1) < 0$, and this stationary point will be unique.

Robinson (1995b) shows that the local Whittle estimator $\widehat{d}$ obtained by minimizing (2) is consistent if $d \in (1/2, 1)$ and asymptotically normally distributed for $d \in (1/2, 3/4)$, so that

$$\sqrt{m}\left(\widehat{d} - d\right) \xrightarrow{d} N\left(0, \frac{1}{4}\right)$$

2.3 The exact local Whittle estimator

An exact local Whittle estimator was introduced by Shimotsu and Phillips (2005). The exact local Whittle estimator is defined by the minimizer of the function[3]

$$R(d) = \log\widehat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j\,, \qquad \widehat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}I_{\Delta^d y}(\omega_j) \tag{6}$$

2. The expression for the Fourier coefficients in (1) can vary between authors by a constant factor. The value of the fractional differencing parameter d, however, is independent of this multiplier although the choice will affect the value of $\widehat{G}$.
3. The algorithm used here is that proposed by Shimotsu and Phillips (2005, 1893). The generalization that allows for an unknown initial value is not implemented.

where $I_{\Delta^d y}(\omega_k)$ is the periodogram of the fractionally differenced series $\Delta^d y_t$ and the difference operator is now defined by the binomial expansion

$$\Delta^d y_t = y_t - \frac{d}{1!} y_{t-1} + \frac{d(d-1)}{2!} y_{t-2} - \frac{d(d-1)(d-2)}{3!} y_{t-3} + \frac{d(d-1)(d-2)(d-3)}{4!} y_{t-4} \quad \cdots$$

The appearance of $I_{\Delta^d y}(\omega_j)$ in the function to be optimized requires that a fractional difference of y_t be computed every time d is altered. Exact local Whittle estimation is thus computationally more demanding than simple local Whittle estimation. Furthermore, it is no longer straightforward to demonstrate that the function in (6) is a convex function of d. Note, however, that the limiting properties of this estimator are the same as those of the local Whittle estimator, so that the asymptotic standard error of the estimate of d in both cases is

$$\text{se}\left(\widehat{d}\right) = \frac{1}{2\sqrt{m}}$$

3 Practical considerations

The local Whittle estimator is determined by the behavior of the spectrum of Y at low frequencies. It is common practice to consider the lowest frequencies $\omega_1 < \omega_2 < \cdots < \omega_m$. A common choice for m is $m = T^{2/3}$. In computing the periodogram $I(\omega_j)$ of Y by means of the discrete Fourier transform (DFT), one may be tempted to accelerate the computation using an implementation of the DFT commonly referred to as the fast Fourier transform. The most common implementation of this algorithm requires that T be a power of 2. If this is not the case, the original series is padded with zeros. The fast Fourier transform is not used here; instead, the simple DFT is used, and only the required m frequencies are computed to reduce computation time. The reason for this choice is that the process of padding does not allow comparison with other Stata commands that do not rely on padding, such as the `gphudak` log-periodogram regression.

Suppose that it must be established a priori that the function $R(d)$ has a minimum value in the interval $[d_0, d_1]$. Ordinarily, the quadratic rate of convergence of the Newton–Raphson algorithm would make this approach the method of choice, particularly because (as detailed in the appendix) analytic expressions for the first and second derivatives of $R(d)$ are available, at least for the local Whittle estimator. However, here we propose a golden section search that estimates d by systematically reducing the length of the interval containing the minimum of $R(d)$ from its original length $|d_0 - d_1|$ to a final length determined by an error tolerance, say, ε. Each evaluation of $R(d)$ allows the interval containing the minimum to be reduced by fraction $(\sqrt{5}-1)/2 \approx 0.6180$, which is the golden section ratio. When the minimum is contained within an interval of length 2ε, no further computations of $R(d)$ are performed, and $\widehat{d}$ is returned as the midpoint of the final interval of search. Unlike the Newton–Raphson algorithm, the golden section search is totally robust: the algorithm cannot trigger prematurely as can happen infrequently with a conventional convergence criterion. The default termination criterion is set at $\varepsilon = 5.0 \times 10^{-7}$ or six decimal places of rounding accuracy.

For completeness, we must consider the possibility that the minimum value of the negative log-likelihood function is achieved at the endpoints of the interval of search, that is, either $d = d_0$ or $d = d_1$. Suppose the latter is true without loss of generality. Then, in this case, the golden section search algorithm will select a sequence of values of d that increase monotonically toward d_1. Specifically, after n iterations, the search points will be $d_1 - Lr^n$ and $d_1 - Lr^{n+1}$, where $L = |d_0 - d_1|$ and $r = (\sqrt{5} - 1)/2$ is the golden ratio. The search will stop whenever the termination condition

$$\left|(d_1 - Lr^{n+1}) - (d_1 - Lr^n)\right| \leq 2\varepsilon$$

is satisfied, and the location of the minimum will be returned as $d = d_1 - Lr^n(1+r)/2 = d_1 - Lr^{n-1}/2$, the midpoint of the final interval of search. The termination condition may be further simplified to get $Lr^{n+2} \leq 2\varepsilon$, which in turn means that whenever the negative log-likelihood function has no minimum within the interval of search, the estimated value of d will lie within $\varepsilon/r^3 = (2 + \sqrt{5})\,\varepsilon < 5\,\varepsilon$ of an endpoint.

Consequently, an estimated value of d should be rejected whenever this value lies within $(2 + \sqrt{5})\,\varepsilon$ of an endpoint of the interval of search. The command automatically checks this condition and will issue an error message whenever the estimate of d lies within $5\,\varepsilon$ of an endpoint of the interval of search.

4 The whittle command

The command `whittle` calculates the local Whittle estimate of d, the order of fractional integration. The exact local Whittle estimate, derived by Shimotsu and Phillips (2005), can be computed as an option.

4.1 Syntax

Before using these commands, and as with other Stata time-series commands, one must `tsset` or `xtset` the data, so that the variable of interest is defined as a proper time series. The command syntax:

`whittle` *varname* [*if*] [*in*] [`,` `powers(`*numlist*`)` `detrend` `exact`]

Note that *varname* may not contain gaps. *varname* can contain time-series operators. The command can be applied to one unit of a panel. `whittle` supports the `by:` prefix, which can be used to operate on each time series in a panel.

4.2 Options

`powers(`*numlist*`)` specifies a list of one or more fractional values for the power of sample size T to be included in computing the local Whittle estimate. The default is `powers(0.65)`.

`detrend` specifies that a linear trend be removed from the *varname* before the local Whittle estimate is computed.

`exact` specifies that the exact local Whittle estimator of Shimotsu and Phillips (2005) be used rather than the local Whittle estimator.

4.3 Stored results

`whittle` stores the following in `r()`:

Macros	
`r(varname)`	variable name
`r(cmdname)`	command name
Matrices	
`r(whittle)`	$6 \times p$ array

The `r(whittle)` matrix contains p columns corresponding to the list of `powers()`. The 6 rows provide the number of observations, power, truncation lag (number of ordinates included), Whittle point estimate, estimated standard error, and asymptotic standard error.

5 Simulation experiments

To evaluate the performance of the local and exact local Whittle estimators, we undertake a small simulation exercise. The traditional approach to simulating long-memory processes is a two-step method suggested by Brockwell and Davis (1991) in which the autocorrelation function of a long-memory process is computed recursively and then used to generate the observations. For a given fractional difference parameter d, the autocorrelation function of a long-memory process may be constructed using the relation

$$\gamma_k = \frac{1}{k!} \prod_{j=0}^{k-1} (j - d) \tag{7}$$

Once the autocorrelation function is available, the synthetic data are computed using the recursion

$$y_t = -\sum_{k=1}^{\infty} \gamma_k y_{t-k} + \varepsilon_t \tag{8}$$

where ε_t ($t \in \mathrm{Z}$) are independent Gaussian deviates with mean value zero and variance $\sigma^2 = 1$ in this instance.

Each simulated time series generated using (7) and (8) was initialized with a draw from $N(0,1)$, and 10,000 observations were generated using $d = \{0.35, 0.65\}$. The first 8,000 observations were treated as the "burn in", and the remaining 2,000 observations were divided into samples of $T = 500$ and $T = 2000$ observations. For each sample size, the local Whittle and exact local Whittle estimators of d were used to estimate the fractional index based on three different sets of frequencies: $m = T^{0.55}$, $m = T^{0.65}$, and $m = T^{0.75}$. Each estimation was repeated 10,000 times to build an accurate picture of the distribution of the estimators. Figures 1 and 2 show the distributions of $\widehat{d}$ obtained by local Whittle estimation and exact local Whittle estimation, respectively. The left-hand column in each figure shows the results for $d = 0.35$, and the right-hand column shows the results for $d = 0.65$. In each cell, the histogram with no face color represents $T = 500$, and the gray-colored histogram represents $T = 2000$.

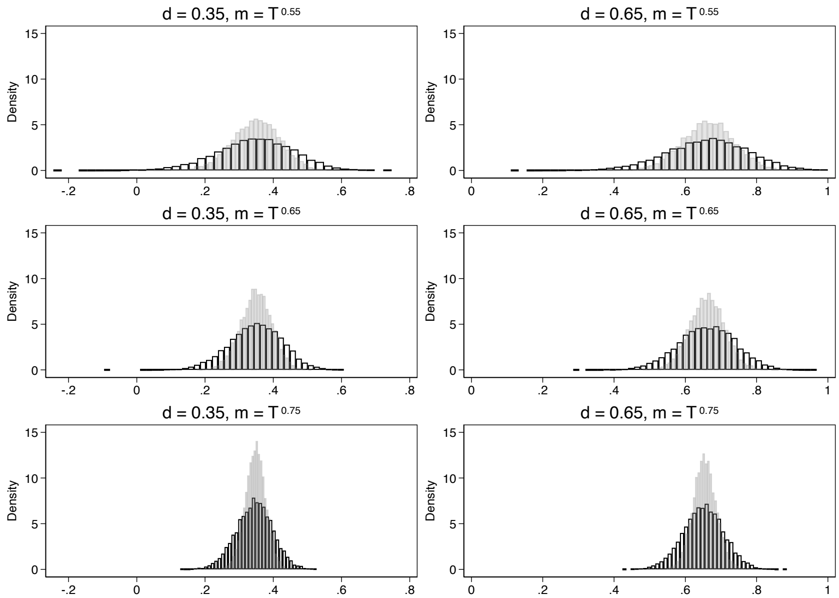

Figure 1. Distributions of estimated fractional difference parameter using the local `whittle` command

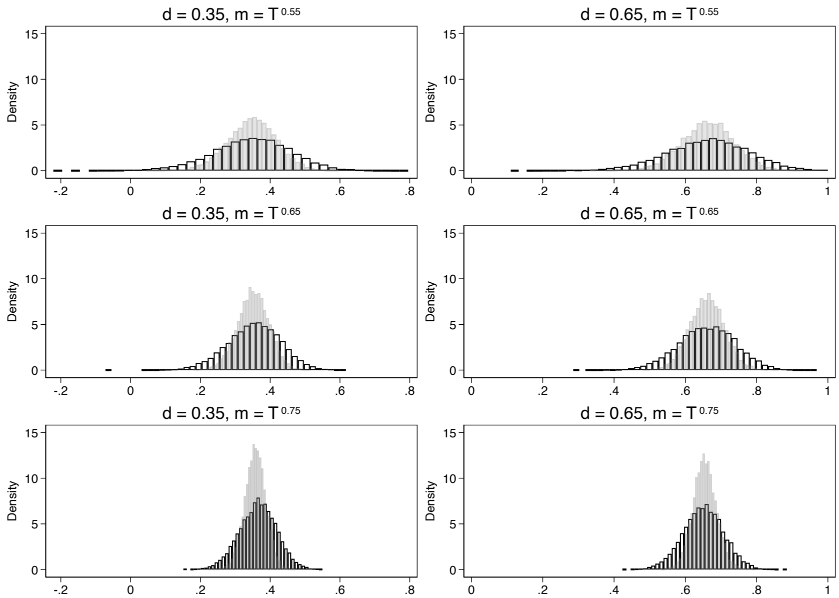

Figure 2. Distributions of estimated fractional difference parameter using the exact local `whittle` command

The results suggest that the two commands are behaving largely as expected. There is clear evidence of the asymptotic normality of the distribution of $\widehat{d}$. There is no clear evidence to suggest that the exact local Whittle estimator is superior to the local Whittle estimator. Despite the fact that the choice of $d = 0.65$ represents a nonstationary process and encroaches on the upper bound of 1, both estimators deal adequately with this situation. If anything, there may be a slight upward bias in $\widehat{d}$ when $d = 0.65$.

The intention of this simulation is not to provide evidence of the efficacy of various estimators of d or provide practical guidance concerning the choice of the estimator. The purpose is rather to demonstrate that the local and exact local Whittle estimators behave as expected. On the evidence in figures 1 and 2, however, there seems little to suggest that the extra computational burden of the exact local Whittle estimator is justified. Of course, this result could very well be different if the starting point is treated as unknown, as in the generalization of the exact local Whittle estimator in Shimotsu and Phillips (2005).

6 Empirical illustrations

6.1 The water level of the Nile

One of the best-known examples of long memory is the time series of the minimal water level of the Nile River for the years 622–1284 measured at the Roda Gauge near Cairo (Beran 1994). The data in `nile.dta` represent annual minimum levels of the Nile River for the period 622–1284.

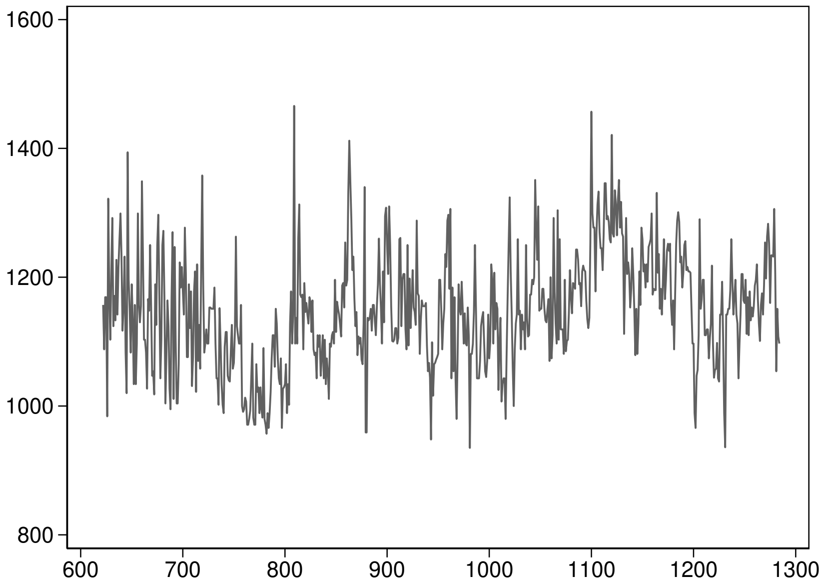

Figure 3. Annual minimum levels of the Nile River for the period 622–1284

From figure 3, it is apparent that there are long periods where the observations tend to stay at a high level, and, on the other hand, there are long periods with low levels. When we looking at short time periods, there seem to be cycles or a local trend. However, when we look at the whole series, there does not appear to be a persistent cycle.

Figure 4 plots the autocorrelations of the Nile data out to a horizon of 40 lags and compares it with the autocorrelations of simulated FI($d = 0.5$) series of similar length. The autocorrelations of the fractionally integrated data decay quickly at short horizons, but at longer horizons, the speed of decay slows down. The autocorrelations of the Nile data exhibit a pattern that is remarkably similar to that of the simulated data. This pattern is known as hyperbolic decay and is typical of a stochastic process with long memory.

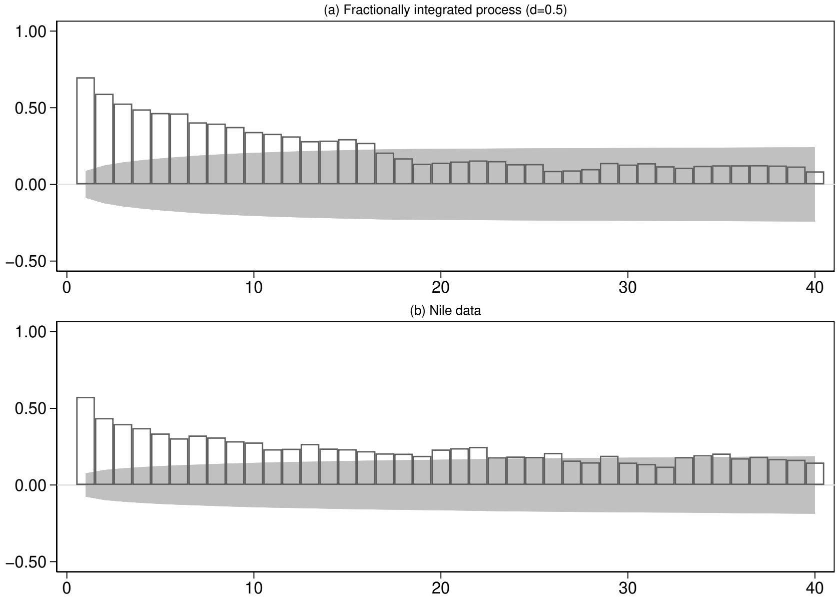

Figure 4. Autocorrelations of a simulated ARFIMA$(0, 0.5, 0)$ process (top panel) compared with the actual autocorrelations of the Nile data (bottom panel). The simulated dataset comprises 500 observations. The Nile data are annual data for the period 622–1286.

The fractional difference parameter d is now estimated using some existing community-contributed regression-based methods. The Geweke and Porter-Hudak (1983) estimator gives point estimates of d varying between 0.40 and 0.58 depending on the number of frequencies used in the estimation. Furthermore, standard t tests reject the null hypothesis of $d = 0$: this conclusion holds for tests based on both the estimated standard error of $\widehat{d}$ and the asymptotic standard error.

Phillips' (1999, 2007) modified Geweke and Porter-Hudak (1983) estimator indicates clearly that the value of d is different from either 0 or 1, implying that the Nile series is neither $I(0)$ nor $I(1)$, reflecting long-memory behavior. Robinson's (1995b) estimator of the series also clearly rejects the null hypothesis of $d = 0$, with $\widehat{d}$ varying between 0.38 and 0.58. All three of the semiparametric estimators yield similar results, and while the hypothesis of $d = 0$ is soundly rejected, none of the estimation methods can reject the hypothesis that $d > 0.5$. This is a unanimous result that has important implications for the characterization of the series according to Hosking's (1981) taxonomy.

```
. use nile

. tsset year
        time variable:  year, 622 to 1284
                delta:  1 unit

. gphudak nile, powers(0.5(0.05)0.7)

GPH estimate of fractional differencing parameter
-------------------------------------------------------------------------------
                                                      Asy.
Power   Ords    Est d   StdErr  t(H0: d=0)  P>|t|    StdErr  z(H0: d=0)  P>|z|
-------------------------------------------------------------------------------
   .5     26  .503829    .1451     3.4730   0.002      .157     3.2088   0.001
  .55     36  .575852    .1397     4.1227   0.000     .1276     4.5133   0.000
   .6     50   .53672     .118     4.5486   0.000     .1045     5.1353   0.000
  .65     69  .449863    .1004     4.4787   0.000    .08666     5.1911   0.000
   .7     95  .396243   .07975     4.9686   0.000    .07249     5.4661   0.000
-------------------------------------------------------------------------------

. modlpr nile, powers(0.5(0.05)0.7)

Modified LPR estimate of fractional differencing parameter for nile
-------------------------------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|    z(H0: d=1)    P>|z|
-------------------------------------------------------------------------------
   .5     25   .5311095   .1487055     3.5716     0.001       -3.6559     0.000
  .55     35    .612678   .1507924     4.0631     0.000       -3.5732     0.000
   .6     49   .5745827   .1304911     4.4032     0.000       -4.6438     0.000
  .65     68   .4693571   .1081714     4.3390     0.000       -6.8236     0.000
   .7     94   .4020765   .0847216     4.7459     0.000       -9.0399     0.000
-------------------------------------------------------------------------------

. roblpr nile, powers(0.5(0.05)0.7)

Robinson estimates of fractional differencing parameter for nile
-------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|
-------------------------------------------------------
   .5     25   .5034895   .1449611     3.4733     0.002
  .55     35   .5751942   .1394794     4.1239     0.000
   .6     49   .5354816   .1176952     4.5497     0.000
  .65     69   .4551935   .0987049     4.6117     0.000
   .7     95   .3846952   .0785415     4.8980     0.000
-------------------------------------------------------

. whittle nile
------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
------------------------------------------------------------
   663  0.65    68    .409044    .06212    .06063
------------------------------------------------------------

. whittle nile, detrend
------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
------------------------------------------------------------
   663  0.65    68    .393717    .06541    .06063
------------------------------------------------------------
```

```
. whittle nile, exact
-------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
-------------------------------------------------------------
   663  0.65    68    .407459    .06243    .06063
-------------------------------------------------------------
. whittle nile, detrend exact
-------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
-------------------------------------------------------------
   663  0.65    68    .397066    .06582    .06063
-------------------------------------------------------------
```

The local Whittle and exact local Whittle estimators produce a lower estimate of d. The range for the estimate is from 0.394 to 0.409. The estimated standard errors of $\widehat{d}$ are very similar to the asymptotic standard errors. Although these results would not reject the hypothesis of $d = 0.5$ and hence long-memory behavior, the evidence in favor of nonstationarity, namely, $d > 0.5$, is much weaker than in the regression-based approaches presented earlier. These results do not appear to be sensitive to the choice of truncation lag.

```
. whittle nile, powers(0.5(0.05)0.7)
-------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
-------------------------------------------------------------
   663  0.50    25    .466848     .1139         .1
   663  0.55    35    .469123    .09495    .08452
   663  0.60    49    .459277    .07914    .07143
   663  0.65    68    .409044    .06212    .06063
   663  0.70    94    .385763    .05091    .05157
-------------------------------------------------------------
```

6.2 Global sea level

The thermal expansion of oceans and melting of land-based ice implies that global warming is very likely contributing to the rise in sea level observed during the 20th century. Ventosa-Santaulària, Heres, and Martínez-Hernández (2014) argue strongly that mean sea-level data exhibit long memory and build a model for sea level and global temperature based on fractional cointegration.

Figure 5 plots monthly time-series data from January 1880 to December 2009 from `sealevel.dta`. The plot indicates the presence of a trend in global sea levels, so the issue of identifying the long-memory parameter must account for this. Calling the command to estimate the fractional difference parameter using local or exact local Whittle estimation, with small numbers of frequencies, results in a fractional difference parameter close to 1. The value of -999 returned as the standard error indicates that this estimate of d should be ignored.

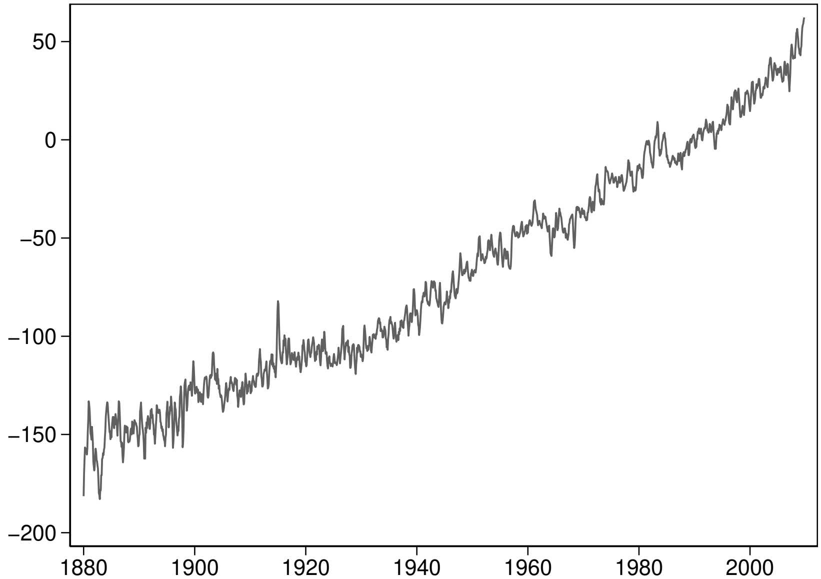

Figure 5. Monthly mean global sea level for January 1880–December 2009

```
. use sealevel, clear
. format datevec %tmCCYY
. whittle Sea, powers(0.5 0.65)
------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
------------------------------------------------------------
Unreliable Whittle estimate - value too close to upper boundary
  1558  0.50    39          1      -999    .08006
  1558  0.65   118    .859211    .04384    .04603
------------------------------------------------------------
. whittle Sea, exact powers(0.5 0.65)
------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
------------------------------------------------------------
Unreliable eWhittle estimate - value too close to upper boundary
  1558  0.50    39          1      -999    .08006
  1558  0.65   118    .801612    .03494    .04603
------------------------------------------------------------
```

This result is not an artifact of the estimation procedure. Using regression-based approaches on the same data shows that the commands are trying to push the difference parameter toward the boundary at 1.

```
. gphudak Sea, powers(0.5(0.05)0.7)
GPH estimate of fractional differencing parameter
------------------------------------------------------------------------------
                                                          Asy.
Power   Ords    Est d   StdErr  t(H0: d=0)  P>|t|    StdErr  z(H0: d=0)  P>|z|
------------------------------------------------------------------------------
   .5     40   1.0218   .04089     24.9904  0.000     .1194      8.5575  0.000
  .55     58  .987866   .03885     25.4297  0.000    .09552     10.3416  0.000
   .6     83  .973678   .03605     27.0089  0.000    .07768     12.5346  0.000
  .65    119   .90265   .03487     25.8852  0.000    .06353     14.2076  0.000
   .7    172  .928184   .03527     26.3151  0.000    .05206     17.8305  0.000
------------------------------------------------------------------------------
. roblpr Sea, powers(0.5(0.05)0.7)
Robinson estimates of fractional differencing parameter for Sea
--------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|
--------------------------------------------------------
   .5     39   1.021511   .0408652    24.9971     0.000
  .55     57   .9872563   .0388347    25.4220     0.000
   .6     83   .9685342   .0357526    27.0899     0.000
  .65    119   .9009264   .0345313    26.0901     0.000
   .7    171   .9237293   .0350966    26.3196     0.000
--------------------------------------------------------
```

It is apparent that not accounting for a deterministic trend in the data causes the estimation to fail. Operating on the detrended data yields the following results.

```
. whittle Sea, detrend powers(0.5 0.65)
-------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
-------------------------------------------------------------
  1558  0.50    39    .551102    .08104    .08006
  1558  0.65   118    .454184    .04165    .04603
-------------------------------------------------------------
. whittle Sea, detrend exact powers(0.5 0.65)
-------------------------------------------------------------
    N   Power  Trunc   Est d      StdErr    Asy.StdErr
-------------------------------------------------------------
  1558  0.50    39    .524068    .07937    .08006
  1558  0.65   118    .486036    .04394    .04603
-------------------------------------------------------------
```

The estimate of the fractional difference parameter is very similar to that returned by the `modlpr` command, which accounts for a deterministic trend by default.

```
. modlpr Sea, powers(0.5(0.05)0.7)
Modified LPR estimate of fractional differencing parameter for Sea
------------------------------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|    z(H0: d=1)    P>|z|
------------------------------------------------------------------------------
   .5     39   .6135026   .1324401     4.6323     0.000       -3.7639    0.000
  .55     57    .587982   .1036759     5.6713     0.000       -4.8508    0.000
   .6     82   .5449442   .0819157     6.6525     0.000       -6.4258    0.000
  .65    118   .5245408   .0674101     7.7813     0.000       -8.0540    0.000
   .7    171   .4766134   .0526453     9.0533     0.000      -10.6728    0.000
------------------------------------------------------------------------------
```

7 Conclusion

The command `whittle` was introduced, which computes the local Whittle and exact local Whittle estimates of the fractional difference parameter d in time series assumed to exhibit long memory. The command always removes the mean from the time series under investigation and allows for detrending using a deterministic trend and for controlling the number of frequencies to be used in constructing the Whittle log-likelihood function. The command returns an estimate of the standard error of the estimated degree of fractional integration. Because the problem is a one-dimensional optimization and the solution is confined to the range -0.5 to 1.0, a golden section search is used to find the estimate. Golden section is a robust search algorithm that is guaranteed to find the optimum provided that the search area brackets the optimum and the function is convex in the search domain.

A small simulation study demonstrated that the command performs broadly as expected, and the empirical examples highlighted some of the practical problems that are often encountered when estimating the fractional difference parameter. Note that there is no attempt to compare the various available estimators, because this would represent an endeavor worthy of an entire article in its own right.

8 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 20-3
. net install st0609      (to install program files, if available)
. net get st0609          (to install ancillary files, if available)
```

9 References

Baum, C. F., and V. Wiggins. 2000. sts16: Tests for long memory in a time series. *Stata Technical Bulletin* 57: 39–44. Reprinted in *Stata Technical Bulletin Reprints*. Vol. 10, pp. 362–368. College Station, TX: Stata Press.

Beran, J. 1994. *Statistics for Long-Memory Processes*. New York: Chapman & Hall/CRC.

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

Cheung, Y.-W., and F. X. Diebold. 1994. On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. *Journal of Econometrics* 62: 301–316. https://doi.org/10.1016/0304-4076(94)90026-4.

Choudhuri, N., S. Ghosal, and A. Roy. 2004. Contiguity of the Whittle measure for a Gaussian time series. *Biometrika* 91: 211–218. https://doi.org/10.1093/biomet/91.1.211.

Fox, R., and M. S. Taqqu. 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* 14: 517–532. http://doi.org/10.1214/aos/1176349936.

Geweke, J., and S. Porter-Hudak. 1983. The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4: 221–238. http://doi.org/10.1111/j.1467-9892.1983.tb00371.x.

Granger, C. W. J., and R. Joyeux. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1: 15–29. http://doi.org/10.1111/j.1467-9892.1980.tb00297.x.

Hosking, J. R. M. 1981. Fractional differencing. *Biometrika* 68: 165–176. https://doi.org/10.1093/biomet/68.1.165.

Künsch, H. 1987. Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernouilli Society*, 64–74. Utrecht: VNU Science Press.

Phillips, P. C. B. 1999. Discrete fourier transforms of fractional processes. Discussion Paper No. 1243, Cowles Foundation. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=216308.

———. 2007. Unit root log periodogram regression. *Journal of Econometrics* 138: 104–124. https://doi.org/10.1016/j.jeconom.2006.05.017.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. Cambridge: Cambridge University Press.

Robinson, P. M. 1995a. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23: 1630–1661. https://doi.org/10.1214/aos/1176324317.

———. 1995b. Log-periodogram regression of time series with long range dependence. *Annals of Statistics* 23: 1048–1072. https://doi.org/10.1214/aos/1176324636.

Shimotsu, K., and P. C. B. Phillips. 2005. Exact local Whittle estimation of fractional integration. *Annals of Statistics* 33: 1890–1933. https://doi.org/10.1214/009053605000000309.

Sowell, F. 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53: 165–188. https://doi.org/10.1016/0304-4076(92)90084-5.

Ventosa-Santaulària, D., D. R. Heres, and L. C. Martínez-Hernández. 2014. Long-memory and the sea level–temperature relationship: A fractional cointegration approach. *PLOS ONE* 9: e113439. https://doi.org/10.1371/journal.pone.0113439.

Whittle, P. 1951. *Hypothesis Testing in Time Series Analysis*. Uppsala: Almqvist & Wiksells.

———. 1962. Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute* 39: 105–129.

About the authors

Christopher F. Baum is Professor of Economics and Social Work at Boston College and DIW Research Fellow at the German Institute for Economic Research. He is the author of two Stata Press books and maintains the Statistical Software Components archive of community-contributed software.

Stan Hurn is Professor of Econometrics in the School of Economics and Finance at Queensland University of Technology and Director of the National Centre for Econometric Research in Australia.

Kenneth Lindsay is Honorary Senior Research Fellow in the School of Mathematics and Statistics at the University of Glasgow.

A Appendix: Convexity of the local Whittle log-likelihood function

Formal differentiation of $R(d)$ in (5) with respect to d yields

$$\frac{dR}{dd} = \frac{\widehat{G}'(d)}{\widehat{G}(d)} - \frac{2}{m}\sum_{k=1}^{m}\omega_k\,, \qquad \frac{d^2R}{dd^2} = \frac{\widehat{G}''(d)\,\widehat{G}(d) - \left\{\widehat{G}'(d)\right\}^2}{\widehat{G}(d)^2}$$

where

$$\widehat{G}'(d) = \frac{2}{m}\sum_{k=1}^{m}(\log\omega_k)\omega_k^{2d}I(\omega_k)\,, \qquad \widehat{G}''(d) = \frac{4}{m}\sum_{k=1}^{m}(\log\omega_k)^2\omega_k^{2d}I(\omega_k)$$

Consider

$$\begin{aligned}
&\widehat{G}''(d)\,\widehat{G}(d) - \left\{\widehat{G}'(d)\right\}^2 \\
&= \frac{4}{m^2}\left\{\sum_{j=1}^{m}\omega_j^{2d}I(\omega_j)\sum_{k=1}^{m}(\log\omega_k)^2\omega_k^{2d}I(\omega_k) - \sum_{j=1}^{m}(\log\omega_j)\omega_j^{2d}I(\omega_j)\right. \\
&\qquad\left.\sum_{k=1}^{m}(\log\omega_k)\omega_k^{2d}I(\omega_k)\right\} \\
&= \frac{4}{m^2}\sum_{j,k=1}^{m}\omega_j^{2d}\omega_k^{2d}I(\omega_k)I(\omega_j)\left\{(\log\omega_k)^2 - (\log\omega_j)(\log\omega_k)\right\}
\end{aligned} \qquad (9)$$

Reversing the order of summation in (9) gives

$$\widehat{G}''(d)\,\widehat{G}(d) - \left\{\widehat{G}'(d)\right\}^2 = \frac{4}{m^2}\sum_{k,j=1}^{m} \omega_k^{2d}\omega_j^{2d} I(\omega_j) I(\omega_k) \left\{(\log\omega_j)^2 - (\log\omega_k)(\log\omega_j)\right\} \quad (10)$$

Adding (9) and (10) gives

$$\widehat{G}''(d)\,\widehat{G}(d) - \left\{\widehat{G}'(d)\right\}^2 = \frac{2}{m^2}\sum_{k,j=1}^{m} (\omega_k\omega_j)^{2d}\, I(\omega_j) I(\omega_k) \left(\log\omega_j - \log\omega_k\right)^2 > 0$$

As a by-product of these computations, the local Whittle estimate, $\widehat{d}$, which minimizes $R(d)$, has an estimated standard error given by

$$\widehat{\mathrm{se}}\left(\widehat{d}\right) = \frac{\sum_{k=1}^{m} \omega_k^{2\widehat{d}} I(\omega_k)}{2\sqrt{\sum_{k=1}^{m} \omega_k^{2\widehat{d}} I(\omega_k) \sum_{j=k+1}^{m} \omega_j^{2\widehat{d}} I(\omega_j) \left(\log\omega_j - \log\omega_k\right)^2}}$$