EC821: Time Series Econometrics
Spring 2003
Notes Section 10 Part 2

Christopher F Baum Department of Economics
Boston College baum@bc.edu

January 8, 2003

1. Fractionally integrated timeseries and ARFIMA modelling[1]

The model of an autoregressive fractionally integrated moving average process of a timeseries of order (p, d, q), denoted by ARFIMA (p, d, q), with mean μ, may be written using operator notation as

$$\Phi(L)(1-L)^d\,(y_t-\mu)=\Theta(L)\epsilon_t,\ \epsilon_t\sim i.i.d.(0,\sigma_\epsilon^2) \tag{1.1}$$

where L is the backward-shift operator, $\Phi(L)$ = 1 - $\phi_1 L$ - .. - $\phi_p L^p$, $\Theta(L)$ = 1 + $\vartheta_1 L$ + ... + $\vartheta_q L^q$, and $(1-L)^d$ is the fractional differencing operator defined by

$$(1-L)^d=\sum_{k=0}^{\infty}\frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)} \tag{1.2}$$

with $\Gamma(\cdot)$ denoting the gamma (generalized factorial) function. The parameter d is allowed to assume any real value. The arbitrary restriction of d to integer values gives rise to the standard autoregressive integrated moving average (ARIMA) model. The stochastic process y_t is both stationary and invertible if all roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle and $|d|<0.5$. The process

[1] This presentation of ARFIMA modelling draws heavily from Baum and Wiggins (2000).

is nonstationary for $d \geq 0.5$, as it possesses infinite variance, i.e. see Granger and Joyeux (1980).

Assuming that $d \in [0,\, 0.5)$, Hosking (1981) showed that the autocorrelation function, $\rho(\cdot)$, of an ARFIMA process is proportional to k^{2d-1} as $k \to \infty$. Consequently, the autocorrelations of the ARFIMA process decay hyperbolically to zero as $k \to \infty$ in contrast to the faster, geometric decay of a stationary ARMA process. For $d \in (0, 0.5)$, $\sum_{j=-n}^{n} |\rho(j)|$ diverges as $n \to \infty$, and the ARFIMA process is said to exhibit long memory, or long-range positive dependence. The process is said to exhibit intermediate memory (anti-persistence), or long-range negative dependence, for $d \in (-0.5, 0)$. The process exhibits short memory for $d = 0$, corresponding to stationary and invertible ARMA modeling. For $d \in [0.5,\, 1)$ the process is mean reverting, even though it is not covariance stationary, as there is no long-run impact of an innovation on future values of the process.

If a series exhibits long memory, it is neither stationary ($I(0)$) nor is it a unit root ($I(1)$) process; it is an $I(d)$ process, with d a real number. A series exhibiting long memory, or persistence, has an autocorrelation function that damps hyperbolically, more slowly than the geometric damping exhibited by "short memory" (ARMA) processes. Thus, it may be predictable at long horizons. An excellent survey of long memory models–which originated in hydrology, and have been widely applied in economics and finance–is given by Baillie (1996).

1.1. Approaches to estimation of the ARFIMA model

There are two approaches to the estimation of an ARFIMA (p, d, q) model: exact maximum likelihood estimation, as proposed by Sowell (1992), and semiparametric approaches. Sowell's approach requires specification of the p and q values, and estimation of the full ARFIMA model conditional on those choices. This involves all the attendant difficulties of choosing an appropriate ARMA specification, as well as a formidable computational task for each combination of p and q to be evaluated. We first describe semiparametric methods, in which we assume that the "short memory" or ARMA components of the timeseries are relatively unimportant, so that the long memory parameter d may be estimated without fully specifying the data generating process.

1.2. Semiparametric estimators for I(d) series

1.2.1. The Lo Modified Rescaled Range estimator[2]

`lomodrs` performs Lo's (1991) modified rescaled range (R/S, "range over standard deviation") test for long range dependence of a time series. The classical R/S statistic, devised by Hurst (1951) and Mandelbrot (1972), is the range of the partial sums of deviations of a timeseries from its mean, rescaled by its standard deviation. For a sample of n values $\{x_1, x_2, \ldots x_n\}$,

$$Q_n = \frac{1}{s_n}\left[Max_{1\leq k\leq n}\sum_{j=1}^{k}(x_j - \bar{x}_n) - Min_{1\leq k\leq n}\sum_{j=1}^{k}(x_j - \bar{x}_n)\right]$$

where s_n is the maximum likelihood estimator of the standard deviation of x. The first bracketed term is the maximum of the partial sums of the first k deviations of x_j from the full-sample mean, which is nonnegative. The second bracketed term is the corresponding minimum, which is nonpositive. The difference of these two quantities is thus nonnegative, so that $Q_n > 0$. Empirical studies have demonstrated that the R/S statistic has the ability to detect long-range dependence in the data.

Like many other estimators of long-range dependence, though, the R/S statistic has been shown to be excessively sensitive to "short-range dependence," or short memory, features of the data. Lo (1991) shows that a sizable $AR(1)$ component in the data generating process will seriously bias the R/S statistic. He modifies the R/S statistic to account for the effect of short-range dependence by applying a "Newey-West" correction (using a Bartlett window) to derive a consistent estimate of the long-range variance of the timeseries. For `maxlag`> 0, the denominator of the statistic is computed as the Newey-West estimate of the long run variance of the series. If `maxlag` is set to zero, the test performed is the classical Hurst-Mandelbrot rescaled-range statistic. Critical values for the test are taken from Lo, 1991, Table II.

Inference from the modified R/S test for long range dependence is complementary to that derived from that of other tests for long memory, or fractional integration in a timeseries, such as `kpss`, `gphudak`, `modlpr` and `roblpr`.

[2] This discussion is drawn from Baum and Room (2000).

1.2.2. The Geweke–Porter-Hudak log periodogram regression estimator

`gphudak` performs the Geweke and Porter-Hudak (GPH, 1983) semiparametric log periodogram regression, often described as the "GPH test," for long memory (fractional integration) in a timeseries. The GPH method uses nonparametric methods–a spectral regression estimator–to evaluate d without explicit specification of the "short memory" (ARMA) parameters of the series. The series is usually differenced so that the resulting d estimate will fall in the [-0.5, 0.5] interval.

Geweke and Porter-Hudak (1983) proposed a semiparametric procedure to obtain an estimate of the memory parameter d of a fractionally integrated process X_t in a model of the form

$$(1-L)^d X_t = \epsilon_t, \tag{1.3}$$

where ϵ_t is stationary with zero mean and continuous spectral density $f_\epsilon(\lambda) > 0$. The estimate $\hat{d}$ is obtained from the application of ordinary least squares to

$$\log(I_x(\lambda_s)) = \hat{c} - \hat{d} \log \left|1 - e^{i\lambda_s}\right|^2 + residual \tag{1.4}$$

computed over the fundamental frequencies $\left\{\lambda_s = \frac{2\pi s}{n}, s = 1, ..., m, m < n\right\}$. We define $\omega_x(\lambda_s) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t e^{it\lambda_s}$ as the discrete Fourier transform (dft) of the timeseries X_t, $I_x(\lambda_s) = \omega_x(\lambda_s)\omega_x(\lambda_s)^*$ as the periodogram, and $x_s = \log\left|1 - e^{i\lambda_s}\right|$. Ordinary least squares on (1.4) yields

$$\hat{d} = 0.5 \frac{\sum_{s=1}^{m} x_s \log I_x(\lambda_s)}{\sum_{s=1}^{m} x_s^2}. \tag{1.5}$$

Various authors have proposed methods for the choice of m, the number of Fourier frequencies included in the regression. The regression slope estimate is an estimate of the slope of the series'p ower spectrum in the vicinity of the zero frequency; if too few ordinates are included, the slope is calculated from a small sample. If too many are included, medium and high-frequency components of the spectrum will contaminate the estimate. A choice of $\sqrt{T}$, or `power` $= 0.5$ is often employed. To evaluate the robustness of the GPH estimate, a range of power values (from 0.40–0.75) is commonly calculated as well. Two estimates of the d coefficient's standard error are commonly employed: the regression standard error, giving rise to a standard t-test, and an asymptotic standard error, based upon the theoretical variance of the log periodogram of $\frac{\pi^2}{6}$. The statistic based upon that standard error has a standard normal distribution under the null.

1.2.3. The Phillips Modified GPH log periodogram regression estimator

`modlpr` computes a modified form of the GPH estimate of the long memory parameter, d, of a timeseries, proposed by Phillips (1999a, 1999b). Phillips (1999a) points out that the prior literature on this semiparametric approach does not address the case of $d = 1$, or a unit root, in (1.3), despite the broad interest in determining whether a series exhibits unit-root behavior or long memory behavior, and his work showing that the $\hat{d}$ estimate of (1.5) is inconsistent when $d > 1$, with $\hat{d}$ exhibiting asymptotic bias toward unity. This weakness of the GPH estimator is solved by Phillips' Modified Log Periodogram Regression estimator, in which the dependent variable is modified to reflect the distribution of d under the null hypothesis that $d = 1$. The estimator gives rise to a test statistic for $d = 1$ which is a standard normal variate under the null. Phillips suggests that deterministic trends should be removed from the series before application of the estimator. Accordingly, the routine will automatically remove a linear trend from the series. This may be suppressed with the `notrend` option. The comments above regarding `power` apply equally to `modlpr`.

Phillips' (1999b) modification of the GPH estimator is based on an exact representation of the dft in the unit root case. The modification expresses

$$\omega_x(\lambda_s) = \frac{\omega_u(\lambda_s)}{1 - e^{i\lambda_s}} - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{X_n}{\sqrt{2\pi n}}$$

and the modified dft as

$$\upsilon_x(\lambda_s) = \omega_x(\lambda_s) + \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{X_n}{\sqrt{2\pi n}}$$

with associated periodogram ordinates $I_v(\lambda_s) = \upsilon_x(\lambda_s)\, \upsilon_x(\lambda_s)^*$ (1999b, p.9). He notes that both $\upsilon_x(\lambda_s)$ and, thus, $I_v(\lambda_s)$ are observable functions of the data. The log-periodogram regression is now the regression of $\log I_v(\lambda_s)$ on $a_s = \log\left|1 - e^{i\lambda_s}\right|$. Defining $\bar{a} = m^{-1}\sum_{s=1}^{m} a_s$ and $x_s = a_s - \bar{a}$, the modified estimate of the long-memory parameter becomes

$$\tilde{d} = 0.5 \frac{\sum_{s=1}^{m} x_s \log I_\nu(\lambda_s)}{\sum_{s=1}^{m} x_s^2}. \tag{1.6}$$

Phillips proves that, with appropriate assumptions on the distribution of ϵ_t, the distribution of $\tilde{d}$ follows

$$\sqrt{m}\left(\tilde{d}-d\right) \rightarrow_d N\left(0, \frac{\pi^2}{24}\right), \tag{1.7}$$

so that $\tilde{d}$ has the same limiting distribution at $d = 1$ as does the GPH estimator in the stationary case so that $\tilde{d}$ is consistent for values of d around unity. A semiparametric test statistic for a unit root against a fractional alternative is then based upon the statistic (1999a, p.10):

$$z_d = \frac{\sqrt{m}\left(\tilde{d}-1\right)}{\pi/\sqrt{24}} \tag{1.8}$$

with critical values from the standard normal distribution. This test is consistent against both $d < 1$ and $d > 1$ fractional alternatives.

1.2.4. Robinson's Log Periodogram Regression estimator

`roblpr` computes the Robinson (1995) multivariate semiparametric estimate of the long memory (fractional integration) parameters, $d(g)$, of a set of G timeseries, $y(g)$, $g = 1, G$ with $G \geq 1$. When applied to a set of timeseries, the $d(g)$ parameter for each series is estimated from a single log-periodogram regression which allows the intercept and slope to differ for each series. One of the innovations of Robinson's estimator is that it is not restricted to using a small fraction of the ordinates of the empirical periodogram of the series: that is, the reasonable values of `power` need not exclude a sizable fraction of the original sample size. The estimator also allows for the removal of one or more initial ordinates, and for the averaging of the periodogram over adjacent frequencies. The rationales for using non-default values of either of these options are presented in Robinson (1995).

Robinson (1995) proposes an alternative log-periodogram regression estimator which he claims provides "modestly superior asymptotic efficiency to $\bar{d}(0)$" ($\bar{d}(0)$ being the Geweke and Porter-Hudak estimator) (1995, p.1052). Robinson's formulation of the log-periodogram regression also allows for the formulation of a multivariate model, providing justification for tests that different time series share a common differencing parameter. Normality of the underlying time series is assumed, but Robinson claims that other conditions underlying his derivation are milder than those conjectured by GPH.

We present here Robinson's multivariate formulation, which applies to a single time series as well. Let X_t represent a $G-$dimensional vector with g^{th} element $X_{gt}, g = 1, ..., G$. Assume that X_t has a spectral density matrix $\int_{-\pi}^{\pi} e^{ij\lambda} f(\lambda) \, d\lambda$, with (g, h) element denoted as $f_{gh}(\lambda)$. The g^{th} diagonal element, $f_{gg}(\lambda)$, is the power spectral density of X_{gt}. For $0 < C_g < \infty$ and $-\frac{1}{2} < d_g < \frac{1}{2}$, assume that $f_{gg}(\lambda) \sim C_g \lambda^{-2d_g}$ as $\lambda \to 0+$ for $g = 1, ..., G$. The periodogram of X_{gt} is then denoted as

$$I_g(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} X_{gt} e^{it\lambda} \right|^2, g = 1, ...G \tag{1.9}$$

Without averaging the periodogram over adjacent frequencies nor omission of l initial frequencies from the regression, we may define $Y_{gk} = \log I_g(\lambda_k)$. The least squares estimates of $c = (c_1, ...c_G)'$ and $d = (d_1, ..., d_G)'$ are given by

$$\begin{bmatrix} \tilde{c} \\ \tilde{d} \end{bmatrix} = vec \left\{ Y'Z(Z'Z)^{-1} \right\}, \tag{1.10}$$

where $Z = (Z_1, ...Z_m)'$, $Z_k = (1, -2\log\lambda_k)'$, $Y = (Y_1, ...Y_G)$, and $Y_g = (Y_{g,1}, ...Y_{g,m})'$ for m periodogram ordinates. Standard errors for $\tilde{d}_g$ and for a test of the restriction that two or more of the d_g are equal may be derived from the estimated covariance matrix of the least squares coefficients. The standard errors for the estimated parameters are derived from a pooled estimate of the variance in the multivariate case, so that their interval estimates differ from those of their univariate counterparts. Modifications to this derivation when the frequency-averaging (`j`) or omission of initial frequencies (`l`) options are selected may be found in Robinson (1995).

1.3. Maximum likelihood estimators of ARFIMA models

The theory and implementation of Sowell's exact maximum likelihood estimator of the $ARFIMA(p, d, q)$ model using Ox is described in Doornik and Ooms (1999).

1.4. Applications

Examples of the application of the lomodrs and classical rescaled range estimators:

Data from Terence Mills' *Econometric Analysis of Financial Time Series* on returns from the annual S&P 500 index of stock prices, 1871-1997, are analyzed.

```
. use http://fmwww.bc.edu/ec-p/data/Mills2d/sp500a.dta
. lomodrs sp500ar

Lo Modified R/S test for sp500ar

Critical values for H0: sp500ar is not long-range dependent

90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]

Test statistic: .780838 (1 lags via Andrews criterion) N = 124

. lomodrs sp500ar, max(0)

Hurst-Mandelbrot Classical R/S test for sp500ar

Critical values for H0: sp500ar is not long-range dependent

90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]

Test statistic: .799079 N = 124

. lomodrs sp500ar if tin(1946,)

Lo Modified R/S test for sp500ar

Critical values for H0: sp500ar is not long-range dependent
```

```
90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]

Test statistic:  1.08705  (0 lags via Andrews criterion)  N = 50
```

For the full sample, the null of stationarity may be rejected at 95% using either the Lo modified R/S statistic or the classic Hurst-Mandelbrot statistic. For the postwar data, the null may not be rejected at any level of significance. Long-range dependence, if present in this series, seems to be contributed by pre-World War II behavior of the stock price series.

Examples of `gphudak`, `modlpr`, and `roblpr` estimators:

Data from Terence Mills' *Econometric Analysis of Financial Time Series* on UK FTA All Share stock returns (ftaret) and dividends (ftadiv) are analyzed.

```
. use http://fmwww.bc.edu/ec-p/data/Mills2d/fta.dta

. tsset
        time variable:  month, 1965m1 to 1995m12

. gphudak ftaret,power(0.5 0.6 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|
----------------------------------------------------------------------------
  .50     20  -.00204  .160313    -0.0127   0.990   .187454    -0.0109   0.991
  .60     35  .228244  .145891     1.5645   0.128   .130206     1.7529   0.080
  .70     64  .141861  .089922     1.5776   0.120   .091267     1.5544   0.120
----------------------------------------------------------------------------


. modlpr ftaret, power(0.5 0.55:0.8)

Modified LPR estimate of fractional differencing parameter
```

```
------------------------------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|    z(H0: d=1)    P>|z|
------------------------------------------------------------------------------
  .50     19   .0231191    .139872     0.1653     0.870      -6.6401     0.000
  .55     25   .2519889   .1629533     1.5464     0.135      -5.8322     0.000
  .60     34   .2450011   .1359888     1.8016     0.080      -6.8650     0.000
  .65     46   .1024504   .1071614     0.9560     0.344      -9.4928     0.000
  .70     63   .1601207   .0854082     1.8748     0.065     -10.3954     0.000
  .75     84   .1749659     .08113     2.1566     0.034     -11.7915     0.000
  .80    113   .0969439   .0676039     1.4340     0.154     -14.9696     0.000
------------------------------------------------------------------------------

. roblpr ftaret

Robinson estimates of fractional differencing parameter
-------------------------------------------------------
Power   Ords      Est d    Std Err  t(H0: d=0)    P>|t|
-------------------------------------------------------
  .90    205   .1253645   .0446745     2.8062     0.005
-------------------------------------------------------

. roblpr ftap ftadiv

Robinson estimates of fractional differencing parameters
Power =  .90                           Ords        =    205
-------------------------------------------------------------
Variable         |     Est d       Std Err      t       P>|t|
-----------------+-------------------------------------------
ftap             |  .8698092      .0163302  53.2640     0.000
ftadiv           |  .8717427      .0163302  53.3824     0.000
-------------------------------------------------------------
Test for equality of d coefficients:   F(1,406) =  .00701   Prob > F = 0.9333

. constraint define 1 ftap=ftadiv

. roblpr ftap ftadiv ftaret, c(1)
```

```
Robinson estimates of fractional differencing parameters
Power =  .90                                  Ords         =    205
-------------------------------------------------------------
Variable          |     Est d       Std Err      t       P>|t|
------------------+------------------------------------------
ftap              |  .8707759      .0205143  42.4473     0.000
ftadiv            |  .8707759      .0205143  42.4473     0.000
ftaret            |  .1253645      .0290116   4.3212     0.000
-------------------------------------------------------------
Test for equality of d coefficients:   F(1,610) =  440.11   Prob > F = 0.0000
```

The GPH test, applied to the stock returns series, generates estimates of the long memory parameter that cannot reject the null at the ten percent level using the t-test. Phillips' modified LPR, applied to this series, finds that $d = 1$ can be rejected for all powers tested, while $d = 0$ (stationarity) may be rejected at the ten percent level for powers 0.6, 0.7, and 0.75. Robinson's estimate for the returns series alone is quite precise. Robinson's multivariate test, applied to the price and dividends series, finds that each series has $d > 0$. The test that they share the same d cannot be rejected. Accordingly, the test is applied to all three series subject to the constraint that price and dividends series have a common d, yielding a more precise estimate of the difference in d parameters between those series and the stock returns series.

References

[1] Andrews, D., 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59, 817-858.

[2] Baillie, R. 1996. Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics*, 73, 5-59.

[3] Doornik, Jurgen A. and Marius Ooms. 1999. A package for estimating, forecasting and simulating Arfima Models: Arfima package 1.0 for Ox. Available from the course homepage.

[4] Baum, Christopher F and Tairi Room, 2000. The modified rescaled range test for long memory. Help file for Stata module `lomodrs`, available from SSC-IDEAS at `http://ideas.repec.org`.

[5] Baum, Christopher F and Vince Wiggins, 2000. Tests for long memory in a timeseries. Stata Technical Bulletin 57. Available from the course home page.

[6] Geweke, J. and Porter-Hudak, S. 1983. The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, 221-238.

[7] 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, 15-39.

[8] Hosking, J. R. M. 1981. Fractional Differencing, *Biometrika*, 68, 165-176.

[9] Hurst, H., 1951. Long Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770-799.

[10] Lo, Andrew W., 1991. Long-Term Memory in Stock Market Prices. *Econometrica*, 59, 1991, 1279-1313.

[11] Mandelbrot, B., 1972. Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis. *Annals of Economic and Social Measurement*, 1, 259-290.

[12] Phillips, Peter C.B. 1999a. Discrete Fourier Transforms of Fractional Processes, Unpublished working paper No. 1243, Cowles Foundation for Research in Economics, Yale University. `http://cowles.econ.yale.edu/P/cd/d12a/d1243.pdf`

[13] Phillips, Peter C.B. 1999b. Unit Root Log Periodogram Regression, Unpublished working paper No. 1244, Cowles Foundation for Research in Economics, Yale University. `http://cowles.econ.yale.edu/P/cd/d12a/d1244.pdf`

[14] Robinson, P.M. 1995. Log-Periodogram Regression of Time Series with Long Range Dependence. *Annals of Statistics*, 23:3, 1048-1072.

[15] Sowell, F. 1992. Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models, *Journal of Econometrics*, 53, 165-188.