Research Article

Rolf Larsson*

How Close Is a Fractional Process to a Random Walk with Drift?

Abstract: In this paper, we investigate how close a fractional process can be to a random walk with drift in terms of the sample path. Given the innovation sequence, we calculate the distance to the closest random walk with drift in the sum of squares sense. We also derive the expected distance between the processes under the assumption of white noise normal innovations. A local approximation formula for this distance is given in terms of the sample size, showing that it increases with the sample size more rapidly than the square of the number of observations. Two empirical examples illustrate the results.

Keywords: fractional process, random walk, distance

DOI 10.1515/jtse-2013-0032

1 Introduction

Long memory processes have been studied in physical sciences since the 1950s, see, for example, Hurst (1957). In the form of fractional processes, they were introduced in econometrics by Granger (1980), Granger and Joyeux (1980) and Hosking (1981), see further the survey article by Baille (1996). Of special interest to the topic of our paper is Granger and Joyeux (1980) who, among other things, numerically compare the fractional process to autoregressive and integrated moving average processes in terms of autocorrelations and partial correlations. In the present paper, we take another route and compare the fractional process to a random walk with drift in terms of the sample paths. This comparison is done both numerically and analytically. Our purpose is to find out if it is the case that data that seem to exhibit long memory may be modelled just as well by random walks with drift. For example, consider Figure 1, where we have simulated the fractional process $(1 - L)^{0.7} y_t = \varepsilon_t$ together with the random walk $y_t = -0.0476 + y_{t-1} + \varepsilon_t$

*Corresponding author: Rolf Larsson, Dept of Mathematics, Uppsala University, P.O.Box 480, SE-751 06 Uppsala, Sweden, E-mail: rolf.larsson@math.uu.se

Figure 1: Simulated sample paths of fractional process with parameter 0.7 (blue) and random walk with drift -0.0476 (red)

for the sample size $T = 100$. (This simulation, as well as all other numerical computations in the paper, was performed in Matlab 7.10.0.)

Naturally, a simulation like this is only for one set of seeds, but other sets of seeds seem to give rise to similar pictures, although with different drift parameters each time given the same fractional parameter. Thus, we may ask if, given a realization of $(1 - L)^d y_t = \varepsilon_t$, there is always a process $x_t = \alpha + x_{t-1} + \varepsilon_t$ that is close in some sense.

In a practical situation, it is observations of the process that are known, and the residuals may only be estimated given a model specification. Now, if two models are close in the above sense, this should mean that the estimated residuals arising from fitting the models are also close. So, the empirical counterpart of the theoretical problem studied here is to evaluate closeness of residuals when fitting fractional and random walk models to empirical data sets.

The motivation to study residuals is also that processes are often identified by comparing mean square sums of fitted residuals (MSR) or information criteria, the latter being based on MSR.

In the paper, we derive two propositions, where the first one quantifies the closeness between the fractional process and a random walk with drift in terms of their smallest possible sum of squares difference (SSD), given the innovations. The second proposition treats the innovations as independent normally distributed, and gives an asymptotic approximation of the expected SSD local to $d = 1$. The result of the latter proposition may be interpreted as counterevidence of our initial suspicion. In fact, the local approximation shows that the expected SSD diverges at a rate marginally larger than T^2 , when we let the sample size T tend to infinity.

We illustrate our findings with two empirical examples. The first one is on relatively short series obtained from Spanish opinion polls, while the second one is a long series containing tree ring data.

The next section gives our main results, while Section 3 discusses the two empirical examples. Section 4 offers some concluding remarks. The proof of proposition 4 is given in the appendix.

2 Main results

Given the fractional parameter d and the innovations $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)'$, the following proposition gives the minimum $\|\mathbf{x} - \mathbf{y}\|^2$ and corresponding drift parameter of the random walk.

Proposition 1 Assume that for $t = 1, 2, \dots, T$, we have realizations of

$$(1 - L)^d y_t = \varepsilon_t$$

$$x_t = \alpha + x_{t-1} + \varepsilon_t$$

where initial values are fixed to zero. Let $\mathbf{x} = (x_1, x_2, \dots, x_T)'$ and similarly for $\mathbf{y}$, $\mathbf{A} = \mathbf{I} - (\mathbf{I} - \mathbf{L})^{1-d}$ and $\mathbf{B} = (\mathbf{I} - \mathbf{L}')^{-1}(\mathbf{I} - \mathbf{L})^{-1}$. Then, given $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)'$ and d ,

$$\|\mathbf{x} - \mathbf{y}\|^2 = \alpha^2 \mathbf{1}' \mathbf{B} \mathbf{1} + 2\alpha \mathbf{1}' \mathbf{B} \mathbf{A} \varepsilon + \varepsilon' \mathbf{A}' \mathbf{B} \mathbf{A} \varepsilon \quad [1]$$

is minimized for

$$\alpha = -\frac{\mathbf{1}' \mathbf{B} \mathbf{A} \varepsilon}{\mathbf{1}' \mathbf{B} \mathbf{1}} \quad [2]$$

and the attained minimum is

$$m = \varepsilon' \mathbf{A}' \left(\mathbf{B} - \frac{\mathbf{B} \mathbf{1} \mathbf{1}' \mathbf{B}}{\mathbf{1}' \mathbf{B} \mathbf{1}} \right) \mathbf{A} \varepsilon \quad [3]$$

Remark 1 Via eq. [1], the solutions to $\|\mathbf{x} - \mathbf{y}\|^2 = 0$ are given by

$$\alpha_{1,2} = -\frac{\mathbf{1}' \mathbf{B} \mathbf{A} \varepsilon}{\mathbf{1}' \mathbf{B} \mathbf{1}} \pm \sqrt{-\frac{m}{\mathbf{1}' \mathbf{B} \mathbf{1}}}.$$

Here, because m is an attained value of $\|\mathbf{x} - \mathbf{y}\|^2$, we have $m \geq 0$ with equality only if $\mathbf{A} \varepsilon$ is proportional to $\mathbf{1}$. Hence, this is the only case when $\|\mathbf{x} - \mathbf{y}\|^2 = 0$. In other cases, there are no real solutions.

Remark 2 Note that by putting $\alpha = 0$ in eq. [1], we find that the corresponding distance between the fractional process and a random walk without drift is $\varepsilon' \mathbf{A}' \mathbf{B} \mathbf{A} \varepsilon$.

We note that in our numerical example, we get $\alpha = -0.0476$ using eq. [2].

Proof of proposition 1. We have

$$\mathbf{y} = (\mathbf{I} - \mathbf{L})^{-d} \varepsilon$$

and

$$\mathbf{x} = (\mathbf{I} - \mathbf{L})^{-1}(\alpha \mathbf{1} + \varepsilon),$$

so that

$$\|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y}),$$

with

$$\mathbf{x} - \mathbf{y} = (\mathbf{I} - \mathbf{L})^{-1}(\alpha \mathbf{1} + \mathbf{A} \varepsilon).$$

Furthermore,

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= (\alpha \mathbf{1} + \mathbf{A} \varepsilon)' \mathbf{B} (\alpha \mathbf{1} + \mathbf{A} \varepsilon) \\ &= \alpha^2 \mathbf{1}' \mathbf{B} \mathbf{1} + 2\alpha \mathbf{1}' \mathbf{B} \mathbf{A} \varepsilon + \varepsilon' \mathbf{A}' \mathbf{B} \mathbf{A} \varepsilon. \end{aligned}$$

To find the optimal α , we need to minimize $\|\mathbf{x} - \mathbf{y}\|^2$. Then, differentiating w.r.t. α , it is easily verified that a unique minimum is given by eq. [2], and that the attained minimum is as in eq. [3].

To quantify the attained minimum m , we may regard ε as a stochastic vector and calculate the expectation. For simplicity, assuming that ε is T -dimensional normal with covariance matrix $\sigma^2 \mathbf{I}$, we have

$$\begin{aligned} Em &= \sigma^2 \text{tr} \left\{ \mathbf{A}' \left(\mathbf{B} - \frac{\mathbf{B} \mathbf{1} \mathbf{1}' \mathbf{B}}{\mathbf{1}' \mathbf{B} \mathbf{1}} \right) \mathbf{A} \right\} \\ &= \sigma^2 \left\{ \text{tr}(\mathbf{A}' \mathbf{B} \mathbf{A}) - \frac{\mathbf{1}' \mathbf{B} \mathbf{A} \mathbf{A}' \mathbf{B} \mathbf{1}}{\mathbf{1}' \mathbf{B} \mathbf{1}} \right\}. \end{aligned} \quad [4]$$

This expectation may be calculated numerically given d , which enters the expression only through $\mathbf{A}$. For d values close to one, it is possible to derive a simple approximation. The idea is to write $\delta = 1 - d$ and note that for $k = 1, 2, 3, \dots$,

$$\binom{\delta}{k} = \frac{\delta(\delta-1) \cdots (\delta-k+1)}{k!} = \frac{(-1)^{k-1}}{k} \delta + O(\delta^2),$$

implying

$$\begin{aligned} \mathbf{A} &= \mathbf{I} - (\mathbf{I} - \mathbf{L})^\delta \\ &= \delta \mathbf{L} - \binom{\delta}{2} \mathbf{L}^2 + \cdots + (-1)^T \binom{\delta}{T-1} \mathbf{L}^{T-1} \\ &= \left(\mathbf{L} + \frac{1}{2} \mathbf{L}^2 + \frac{1}{3} \mathbf{L}^3 + \cdots + \frac{1}{T-1} \mathbf{L}^{T-1} \right) \delta + O(\delta^2). \end{aligned} \quad [5]$$

Inserting this expansion into eq. [4], further expansions and simplifications yield the following result.

Proposition 2 *If the innovations ε_t are i.i.d. normal with variance σ^2 , the expected sums of squares distance between the fractional process with parameter d and the closest random walk with drift fulfills*

$$Em = \sigma^2 m_0 (1-d)^2 + O\{(1-d)^3\},$$

where

$$m_0 = \left\{ \frac{1}{10} H(T)^2 - \frac{43}{75} H(T) + \frac{1997}{1500} \right\} T^2 + o(T^2), \quad [6]$$

where $H(T) = \sum_{t=1}^T t^{-1}$ is the Harmonic number.

Proof. See the appendix.

Remark 3 *From Remark 3 and the proof of proposition 4, we may easily deduce the corresponding result for the random walk without drift, where we instead have*

$$m_0 = \left\{ \frac{1}{2} H(T)^2 - \frac{3}{2} H(T) + \frac{7}{4} \right\} T^2 + o(T^2).$$

For the sample size $T = 100$, we depict Em as well as the approximation in Figure 2. (The pattern is very similar regardless of sample size.)

The approximation seems to work well only very close to $d = 1$, but it gives a rough idea of the quantitative behavior of Em also for d values more distant from 1. Since Em diverges rather rapidly with T (quicker than T^2 local to $d = 1$, in fact the Harmonic number $H(T)$ behaves like $\log T$), it seems that fractional processes are more useful for modeling for very large sample sizes than for smaller ones, where they are harder to distinguish from random walks with drift.

Figure 2: Expected sum of squared distances between a fractional process and the closest random walk with drift. Numerical distance in black and local approximation in red

3 Empirical examples

This section discusses two empirical examples that have been analyzed previously in the literature. In these examples, we compare the fits of the fractional process and the random walk model with drift. The purpose is to illustrate the theoretical results of Section 2, in the sense that if the processes are “close,” they should fit the data about equally well, and if not, it is interesting to see which one that fits best. For means of comparison, we also include the random walk without drift.

3.1 Spanish opinion poll series

As our first empirical example, consider the Spanish Political Opinion Poll series of Dolado, Gonzalo and Mayoral (2003). The data are depicted in Figures 3 and 4. (For explanations of abbreviations, see Table 1 of Dolado, Gonzalo and Mayoral, 2003.)

These series, found at the webpage of the *Journal of Applied Econometrics*, have 87 quarterly observations, according to Dolado et al. ranging between 1978(3) and 1999(4) (although this would amount to only 86 observations). Dolado et al. estimate fractional models $(1 - L)^d y_t = \varepsilon_t$ for these data. We have estimated the d parameter up to two decimal points by trial and error minimization of the mean sum of squares of the filtered series (the MSR). (This practice gave slightly different estimates than those obtained by Dolado et al.) As for the random walk with drift, $\hat{\alpha}$ is simply the mean of the differences $x_t - x_{t-1}$ for $t \geq 2$. Then, we have compared to the corresponding MSRs for random walk models with and without drift.

Figure 3: Spanish opinion poll data for the parties AP-PP (blue), PSOE (magenta) and gPCE-IU (red)

Figure 4: Spanish opinion poll data for the parties CIU (blue) and PNV (black)

The results are given in Table 1. We find that the fractional model gives smaller MSR in all cases. Moreover, as expected, the difference to the MSR of the random walk model with drift is smaller the closer $\hat{d}$ is to 1. We should also remark that the Akaike information criterion (AIC) chooses the fractional model as the best model in all cases. The Bayesian information criterion (BIC) chooses the fractional model in the PSOE, CIU and PNV cases, whereas in the other cases, it chooses the random walk without drift. (We note that Dolado et al. use BIC to choose between different fractional specifications.) See Tables 2 and 3 for AIC and BIC respectively, where in each case, the smallest values of the criteria are in bold face. The definitions used were $T \log \hat{\alpha}^2 + 2p$ for AIC and $T \log \hat{\alpha}^2 + p \log T$ for BIC where $\hat{\alpha}^2$ is the MSR and p is the number of parameters.

Table 1: Estimated parameters and MSR for Dolado et al. data

Party	$\hat{d}$	$\hat{\alpha}$	MSR, fractional	MSR, rw with drift	MSR, rw
AP-PP	0.86	$2.9 * 10^{-3}$	$4.695 * 10^{-4}$	$4.819 * 10^{-4}$	$4.905 * 10^{-4}$
PSOE	0.81	$2.5 * 10^{-3}$	$1.997 * 10^{-3}$	$2.108 * 10^{-3}$	$2.114 * 10^{-3}$
PCE-IU	0.83	$4.0 * 10^{-4}$	$1.371 * 10^{-4}$	$1.427 * 10^{-4}$	$1.429 * 10^{-4}$
CIU	0.67	$3.1 * 10^{-4}$	$3.228 * 10^{-5}$	$3.869 * 10^{-5}$	$3.878 * 10^{-5}$
PNV	0.66	$1.0 * 10^{-4}$	$1.080 * 10^{-5}$	$1.274 * 10^{-5}$	$1.275 * 10^{-5}$

Table 2: AIC for Dolado et al. data

Party	Fractional	rw with drift	rw
AP-PP	−664.8	−662.5	−663.0
PSOE	−538.8	−534.1	−535.9
PCE-IU	−771.8	−768.3	−770.2
CIU	−897.7	−881.9	−883.7
PNV	−992.9	−978.6	−980.5

Table 3: BIC for Dolado et al. data

Party	Fractional	rw with drift	rw
AP-PP	−662.3	−660.0	−663.0
PSOE	−536.3	−531.6	−535.9
PCE-IU	−769.4	−765.9	−770.2
CIU	−895.2	−879.5	−883.7
PNV	−990.5	−976.1	−980.5

We also compare the prediction performance of the models, by estimating them all but the 10 last observations and then evaluating the mean square prediction error (MSPE) when fitting the model to the last 10 observations. The parameters were estimated as above. For the fractional model, the predictions were generated by rewriting the model equation as

$$y_t = d\hat{y}_{t-1} - \binom{d}{2}\hat{y}_{t-2} + \dots,$$

where $\hat{y}_t$ equals the observation y_t if $t \leq T_0$ where T_0 is the number of observations in the model evaluation set (here 77), otherwise it equals the predicted value. The predictions for the random walk with drift are $\hat{y}_t = y_{T_0} + (t - T_0)\hat{\alpha}$ for $t > T_0$, whereas for the random walk without drift they are simply $\hat{y}_t = y_{T_0}$.

Table 4: Estimated parameters when fitting the models on the first 77 observations and MSPE for the last 10 observations for Dolado et al. data

Party	$\hat{d}$	$\hat{\alpha}$	MSPE, fractional	MSPE, rw with drift	MSPE, rw
AP-PP	0.85	$3.0 * 10^{-3}$	$1.0 * 10^{-3}$	$1.4 * 10^{-4}$	$5.3 * 10^{-4}$
PSOE	0.82	$3.1 * 10^{-3}$	$4.2 * 10^{-4}$	$1.5 * 10^{-3}$	$5.4 * 10^{-4}$
PCE-IU	0.82	$8.7 * 10^{-4}$	$5.1 * 10^{-4}$	$6.8 * 10^{-4}$	$4.4 * 10^{-4}$
CIU	0.69	$4.3 * 10^{-4}$	$3.3 * 10^{-5}$	$9.2 * 10^{-5}$	$5.4 * 10^{-5}$
PNV	0.68	$7.8 * 10^{-5}$	$7.2 * 10^{-6}$	$8.6 * 10^{-6}$	$1.0 * 10^{-5}$

The results, given in Table 4, are very much in line with the outcome of BIC in Table 3. The only difference is that random walk with drift, rather than without drift, is chosen as the best model for AP-PP.

3.2 Mount Campito tree ring data

Another example is the Mount Campito tree ring data of Baille (1996), accessed at Datamarket (2013), see Figure 5.

Figure 5: Mount Campito tree ring data

This data set consists of 5,405 yearly observations that extend from 3435BC up to AD1969. Here, we at first subtract the mean value from the data. Contrary to Baille (1996), who uses the MLE, as above we take as an estimate of d the value that minimizes the MSR. For reasons of computing time, we truncate the fractional filter at lag 40. As an estimate, we choose the d value that minimizes MSR, 0.46, which is slightly different from the estimate of Baille (1996), 0.449. (As in

the previous example, $\hat{\alpha}$ is the mean of the differences of the x_t process.) In this case, the fractional model is superior to the random walk, both in terms of low MSR and of the two studied information criteria. In light of Figure 2 and proposition 4, with such a large sample size and a d value far from 1, these results are not surprising. Moreover, we note that the MSRs of the random walk with and without drift are virtually identical. See Table 5 for results.

Table 5: Estimated parameters, MSR, AIC and BIC for the Mount Campito tree ring series

$\hat{d}$	0.46
$\hat{\alpha}$	0.0033
MSR, fractional	64.11
MSR, rw with drift	84.88
MSR, rw	84.88
AIC, fractional	22490
AIC, rw with drift	24,007
AIC, rw	24,005
BIC, fractional	22,496
BIC, rw with drift	24,014
BIC, rw	24,005

A comparison of the prediction performance, where we estimated the models on the first 5,375 observations, and then compared the MSPEs evaluated on the last 30 observations, is also performed. (Parameters are estimated in the usual manner.) The results are given in Table 6. As expected, the fractional model performs best. However, we must remark that when experimenting with different data truncations than at 5,375, we sometimes obtained that among the three models, the fractional model turned out to give the worst prediction fit. These results are not shown here, but may be obtained from the author at request.

Table 6: Estimated parameters when fitting the models on the first 5,375 observations and MSPE for the last 30 observations for the Mount Campito tree ring series

$\hat{d}$	0.46
$\hat{\alpha}$	0.0044
MSPE, fractional	116.6
MSPE, rw with drift	150.5
MSPE, rw	149.4

4 Concluding remarks

In the present paper, we have investigated how close the trajectory of a fractional process can be to a random walk with drift. The initial suspicion, based on simulations, was that they can indeed be very close, so that the presence of fractional processes in the econometrician's toolbox may come under question. However, our theoretical results prove otherwise: local to $d = 1$, the sums of squares distance between the trajectories of the fractional process and the closest random walk with drift increases with T more rapidly than T^2 , where T is the sample size. The empirical examples confirm our theoretical findings, although one caution should be that the random walk without drift may be picked as the best model by information criteria.

Most probably, the results in our paper could be generalized to more complicated settings, for example more general fractional autoregressive models. However, we expect that this would not change the essence of our findings.

5 Appendix: Proof of proposition 2

Noting that

$$\mathbf{A}'\mathbf{B}\mathbf{A} = (\mathbf{C}\mathbf{A})'(\mathbf{C}\mathbf{A}),$$

where

$$\mathbf{C} = (\mathbf{I} - \mathbf{L})^{-1} = \mathbf{I} + \mathbf{L} + \mathbf{L}^2 + \cdots + \mathbf{L}^{T-1},$$

and via (5), we have

$$\mathbf{C}\mathbf{A} = \mathbf{D}\delta + O(\delta^2), \quad [7]$$

with (observe that $\mathbf{L}^t = 0$ for all $t \geq T$)

$$\begin{aligned} \mathbf{D} &= (\mathbf{I} + \mathbf{L} + \mathbf{L}^2 + \cdots + \mathbf{L}^{T-1}) \\ &\quad \left(\mathbf{L} + \frac{1}{2}\mathbf{L}^2 + \frac{1}{3}\mathbf{L}^3 + \cdots + \frac{1}{T-1}\mathbf{L}^{T-1} \right) \\ &= \sum_{t=1}^{T-1} \sum_{j=1}^t \frac{1}{j} \mathbf{L}^t. \end{aligned}$$

Hence,

$$\text{tr}(\mathbf{A}'\mathbf{B}\mathbf{A}) = c_1\delta^2 + O(\delta^3),$$

where $c_1 = \text{tr}(\mathbf{D}'\mathbf{D})$, and because $\text{tr}(\mathbf{L}^{n'}\mathbf{L}^m) = T - n$ if $n = m$ and 0 otherwise, with $H = H(T - 1)$,

$$\begin{aligned}
 c_1 &= \sum_{t=1}^{T-1} \left(\sum_{j=1}^t \frac{1}{j} \right)^2 (T-t) = \sum_{j=1}^{T-1} \frac{1}{j} \sum_{k=1}^{T-1} \frac{1}{k} \sum_{t=j \vee k}^{T-1} (T-t) \\
 &= \frac{1}{2} \sum_{j=1}^{T-1} \frac{1}{j} \sum_{k=1}^{T-1} \frac{1}{k} (T-j \vee k)(T-j \vee k + 1) \\
 &= \frac{T(T+1)}{2} \left(\sum_{j=1}^{T-1} \frac{1}{j} \right)^2 - \frac{2T+1}{2} \sum_{j=1}^{T-1} \frac{1}{j} \sum_{k=1}^{T-1} \frac{j \vee k}{k} + \frac{1}{2} \sum_{j=1}^{T-1} \frac{1}{j} \sum_{k=1}^{T-1} \frac{(j \vee k)^2}{k} \\
 &= \frac{T(T+1)}{2} H^2 - \frac{2T+1}{2} \sum_{j=1}^{T-1} \left(\sum_{k=1}^j \frac{1}{k} + \frac{T-1-j}{j} \right) + \frac{1}{2} \sum_{j=1}^{T-1} \left(j \sum_{k=1}^j \frac{1}{k} + \frac{1}{j} \sum_{k=j+1}^{T-1} k \right) \\
 &= \frac{T(T+1)}{2} H^2 - \frac{2T+1}{2} \sum_{k=1}^{T-1} \frac{T-k}{k} - \frac{(2T+1)(T-1)}{2} (H-1) \\
 &\quad + \frac{1}{2} \sum_{k=1}^{T-1} \frac{1}{k} \left(k + 2 \sum_{j=k+1}^{T-1} j \right) \\
 &= \frac{T(T+1)}{2} H^2 - \frac{2T+1}{2} (TH - T + 1) - \frac{(2T+1)(T-1)}{2} (H-1) \\
 &\quad + \frac{T-1}{2} + \frac{1}{2} \sum_{k=1}^{T-1} \frac{1}{k} (T-1-k)(T+k) \\
 &= \frac{T(T+1)}{2} H^2 - \frac{2T+1}{2} (TH - T + 1) - \frac{(2T+1)(T-1)}{2} (H-1) \\
 &\quad + \frac{T-1}{2} + \frac{1}{2} \left(T^2 H - TH - \frac{T(T+1)}{2} + 1 \right) \\
 &= \frac{T^2}{2} \left(H^2 - 3H + \frac{7}{2} \right) + o(T^2)
 \end{aligned}$$

[8]

Similarly,

$$\mathbf{1}'\mathbf{B1} = (\mathbf{C1})'(\mathbf{C1}),$$

where

$$\mathbf{C}\mathbf{1} = (\mathbf{I} + \mathbf{L} + \mathbf{L}^2 + \cdots + \mathbf{L}^{T-1})\mathbf{1} = \tau,$$

with $\tau = (1, 2, \dots, T)'$, implying that

$$\mathbf{1}'\mathbf{B}\mathbf{1} = \sum_{t=1}^T t^2 = \frac{1}{3}T^3 + O(T^2). \quad [9]$$

Finally, to approximate $\mathbf{1}'\mathbf{BAA}'\mathbf{B}\mathbf{1}$, from eq. [7] we find

$$\mathbf{1}'\mathbf{BAA}'\mathbf{B}\mathbf{1} = \tau'\mathbf{CA}(\mathbf{CA})'\tau = c_2\delta^2 + O(\delta^3),$$

where $c_2 = \tau'\mathbf{D}\mathbf{D}'\tau$. Here, since

$$\tau'\mathbf{L}^k = (k+1, \dots, T, 0, \dots, 0),$$

it follows that

$$\begin{aligned} c_2 &= \sum_{t=1}^T \left\{ \sum_{j=1}^{T-t} \left(\sum_{k=1}^j \frac{1}{k} \right) (t+j) \right\}^2 \\ &= \sum_{t=1}^T \sum_{j=1}^{T-t} \sum_{l=1}^{T-t} \sum_{k=1}^j \sum_{m=1}^l \frac{1}{km} (t+j)(t+l) \\ &= \sum_{j=1}^{T-1} \sum_{l=1}^{T-1} \sum_{k=1}^j \sum_{m=1}^l \frac{1}{km} \sum_{t=1}^{T-(j \vee l)} (t+j)(t+l) \\ &= \sum_{j=1}^{T-1} \sum_{k=1}^j \sum_{m=1}^j \frac{1}{km} \sum_{t=1}^{T-j} (t+j)^2 + 2 \sum_{j=1}^{T-1} \sum_{l=1}^{j-1} \sum_{k=1}^j \sum_{m=1}^l \frac{1}{km} \sum_{t=1}^{T-j} (t+j)(t+l), \end{aligned} \quad [10]$$

where the first sum is

$$\begin{aligned} &\sum_{j=1}^{T-1} \sum_{k=1}^j \sum_{m=1}^j \frac{1}{km} \sum_{t=1}^{T-j} (t+j)^2 \\ &= \sum_{j=1}^{T-1} \sum_{k=1}^j \sum_{m=1}^j \frac{1}{km} \left(\frac{1}{3}T^3 + \frac{1}{2}T^2 + \frac{1}{6}T - \frac{1}{3}j^3 - \frac{1}{2}j^2 - \frac{1}{6}j \right) \\ &= \left(\frac{1}{3}T^3 + \frac{1}{2}T^2 + \frac{1}{6}T \right) \sum_{k=1}^{T-1} \sum_{m=1}^{T-1} \frac{1}{km} (T - k \vee m) - \sum_{k=1}^{T-1} \sum_{m=1}^{T-1} \frac{1}{km} \sum_{j=k \vee m}^{T-1} \left(\frac{1}{3}j^3 + \frac{1}{2}j^2 + \frac{1}{6}j \right) \\ &= o(T^5) \end{aligned} \quad [11]$$

and the second sum is

$$\begin{aligned}
 & \sum_{j=1}^{T-1} \sum_{l=1}^{j-1} \sum_{k=1}^j \sum_{m=1}^l \frac{1}{km} \sum_{t=1}^{T-j} (t+j)(t+l) \\
 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{1}{km} \left(\frac{1}{3} T^3 - \frac{1}{2} T^2 j + \frac{1}{2} T^2 l - \frac{1}{2} j^2 l + \frac{1}{6} j^3 \right) + o(T^5) \\
 &= \frac{1}{3} T^3 \gamma_1 - \frac{1}{2} T^2 \gamma_2 + \frac{1}{2} T^2 \gamma_3 - \frac{1}{2} \gamma_4 + \frac{1}{6} \gamma_5 + o(T^5).
 \end{aligned} \tag{12}$$

Here,

$$\begin{aligned}
 \gamma_1 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{1}{km} \\
 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \frac{1}{km} \sum_{j=k}^{T-1} 1\{j > m\} (j-m) \\
 &= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \sum_{j=k}^{T-1} (j-m) + \sum_{m=k}^{T-2} \frac{1}{km} \sum_{j=m+1}^{T-1} (j-m) \right\} \\
 &= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(-Tm + km + \frac{1}{2} T^2 - \frac{1}{2} k^2 \right) + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{2} T^2 - Tm + \frac{1}{2} m^2 \right) \right\} \\
 &\quad + o(T^2),
 \end{aligned}$$

where

$$\sum_{k=1}^{T-1} \sum_{m=1}^k \frac{m}{k} = \sum_{m=1}^{T-1} \sum_{k=m}^{T-1} \frac{m}{k} = \sum_{k=1}^{T-1} \sum_{m=k}^{T-1} \frac{k}{m},$$

and so,

$$\begin{aligned}
 \gamma_1 &= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(-Tm + km + \frac{1}{2} T^2 \right) + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{2} T^2 - Tm \right) \right\} + o(T^2) \\
 &= \sum_{k=1}^{T-1} \frac{1}{k} \left(\frac{1}{2} T^2 \sum_{m=1}^{T-2} \frac{1}{m} - T^2 \right) + \sum_{k=1}^{T-1} k + o(T^2) \\
 &= \frac{1}{2} T^2 H^2 - T^2 H + \frac{1}{2} T^2 + o(T^2).
 \end{aligned} \tag{13}$$

Similarly,

$$\begin{aligned}
 \gamma_2 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{j}{km} \\
 &= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \sum_{j=k}^{T-1} j(j-m) + \sum_{m=k}^{T-2} \frac{1}{km} \sum_{j=m+1}^{T-1} j(j-m) \right\} \\
 &= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(-\frac{1}{2}T^2m + \frac{1}{2}k^2m + \frac{1}{3}T^3 - \frac{1}{3}k^3 \right) \right. \\
 &\quad \left. + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{3}T^3 - \frac{1}{2}T^2m + \frac{1}{6}m^3 \right) \right\} + o(T^3) \\
 &= \sum_{k=1}^{T-1} \frac{1}{k} \left(\frac{1}{3}T^3 \sum_{m=1}^{T-2} \frac{1}{m} - \frac{1}{2}T^3 \right) + \frac{1}{2} \sum_{k=1}^{T-1} k^2 - \frac{1}{6} \sum_{k=1}^{T-1} k^2 \sum_{m=1}^{k-1} \frac{1}{m} + o(T^3) \\
 &= \frac{1}{3}T^3H^2 - \frac{5}{9}T^3H + \frac{2}{9}T^3 + o(T^2),
 \end{aligned} \tag{14}$$

$$\begin{aligned}
 \gamma_3 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{l}{km} \\
 &= \frac{1}{2} \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \sum_{j=k}^{T-1} (j^2 - m^2) + \sum_{m=k}^{T-2} \frac{1}{km} \sum_{j=m+1}^{T-1} (j^2 - m^2) \right\} + o(T^3) \\
 &= \frac{1}{2} \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(\frac{1}{3}T^3 - Tm^2 + km^2 - \frac{1}{3}k^3 \right) \right. \\
 &\quad \left. + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{3}T^3 - Tm^2 + \frac{2}{3}m^3 \right) \right\} + o(T^3) \\
 &= \frac{1}{2} \sum_{k=1}^{T-1} \frac{1}{k} \left(\frac{1}{3}T^3 \sum_{m=1}^{T-2} \frac{1}{m} - \frac{1}{2}T^3 \right) + \frac{1}{4} \sum_{k=1}^{T-1} k^2 + \frac{1}{6} \sum_{k=1}^{T-1} k^2 \sum_{m=1}^{k-1} \frac{1}{m} + o(T^3) \\
 &= \frac{1}{6}T^3H^2 - \frac{7}{36}T^3H + \frac{1}{36}T^3,
 \end{aligned} \tag{15}$$

$$\begin{aligned}
\gamma_4 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{j^2 l}{km} \\
&= \frac{1}{2} \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \sum_{j=k}^{T-1} j^2 (j^2 - m^2) + \sum_{m=k}^{T-2} \frac{1}{km} \sum_{j=m+1}^{T-1} j^2 (j^2 - m^2) \right\} + o(T^5) \\
&= \frac{1}{2} \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(\frac{1}{5} T^5 - \frac{1}{3} T^3 m^2 + \frac{1}{3} k^3 m^2 - \frac{1}{5} k^5 \right) \right. \\
&\quad \left. + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{5} T^5 - \frac{1}{3} T^3 m^2 + \frac{2}{15} m^5 \right) \right\} + o(T^5) \\
&= \frac{1}{2} \sum_{k=1}^{T-1} \frac{1}{k} \left(\frac{1}{5} T^5 \sum_{m=1}^{T-2} \frac{1}{m} - \frac{1}{6} T^5 \right) + \frac{1}{12} \sum_{k=1}^{T-1} k^4 - \frac{1}{30} \sum_{k=1}^{T-1} k^4 \sum_{m=1}^{k-1} \frac{1}{m} + o(T^5) \\
&= \frac{1}{10} T^5 H^2 - \frac{9}{100} T^5 H + \frac{9}{100} o(T^5),
\end{aligned} \tag{16}$$

and

$$\begin{aligned}
\gamma_5 &= \sum_{k=1}^{T-1} \sum_{m=1}^{T-2} \sum_{j=k}^{T-1} \sum_{l=m}^{j-1} \frac{j^3}{km} \\
&= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \sum_{j=k}^{T-1} j^3 (j - m) + \sum_{m=k}^{T-2} \frac{1}{km} \sum_{j=m+1}^{T-1} j^3 (j - m) \right\} \\
&= \sum_{k=1}^{T-1} \left\{ \sum_{m=1}^{k-1} \frac{1}{km} \left(\frac{1}{5} T^5 - \frac{1}{4} T^4 m + \frac{1}{4} k^4 m - \frac{1}{5} k^5 \right) \right. \\
&\quad \left. + \sum_{m=k}^{T-2} \frac{1}{km} \left(\frac{1}{5} T^5 - \frac{1}{4} T^4 m + \frac{1}{20} m^5 \right) \right\} + o(T^5) \\
&= \sum_{k=1}^{T-1} \frac{1}{k} \left(\frac{1}{5} T^5 \sum_{m=1}^{T-2} \frac{1}{m} - \frac{1}{4} T^5 \right) + \frac{1}{4} \sum_{k=1}^{T-1} k^4 - \frac{3}{20} \sum_{k=1}^{T-1} k^4 \sum_{m=1}^{k-1} \frac{1}{m} + o(T^5) \\
&= \frac{1}{5} T^5 H^2 - \frac{7}{25} T^5 H + \frac{7}{125} T^5 + o(T^5),
\end{aligned} \tag{17}$$

Inserting eqs [13]–[17] into eq. [12], and then further into eq. [10] (together with eq. [11]), we get the result via eqs [3], [8] and [9].

6 Notation and symbols

We define $(1 - L)^d$ by binomial expansion through

$$(1 - L)^d = 1 - dL + \binom{d}{2}L^2 - \dots,$$

where for $k = 0, 1, 2, \dots$,

$$\binom{d}{k} = \frac{d(d-1) \cdots (d-k+1)}{k!}.$$

Similarly,

$$(\mathbf{I} - \mathbf{L})^d = \mathbf{I} - d\mathbf{L} + \binom{d}{2}\mathbf{L}^2 + \dots + (-1)^{T-1} \binom{d}{T-1} \mathbf{L}^{T-1},$$

where $\mathbf{I}$ and $\mathbf{L}$ are the T -dimensional identity and lag matrices. (Observe that $\mathbf{L}^n = \mathbf{0}$ for $n \geq T$.) We also define $\mathbf{1}$ as the T -dimensional vector of ones. Moreover, the squared norm of a vector $\mathbf{x} = (x_1, x_2, \dots, x_T)'$ is defined as

$$\|\mathbf{x}\|^2 = \sum_{t=1}^T x_t^2.$$

Hence, the sum of squares distance between $\mathbf{x}$ and $\mathbf{y} = (y_1, y_2, \dots, y_T)'$ is denoted by $\|\mathbf{x} - \mathbf{y}\|^2$.

We denote the minimum of the real numbers a and b by $a \wedge b$ and the maximum by $a \vee b$. Finally, $1\{a > b\} = 1$ if $a > b$ and 0 otherwise.

Acknowledgements: I am grateful to the editor and to the anonymous referee for their careful reading and constructive suggestions, which helped to improve the paper.

References

- Baile, R. T. 1996. "Long Memory Processes and Fractional Integration in Econometrics." *Journal of Econometrics* 73:5–59.
- Datamarket. 2013. <http://datamarket.com/data/set/22ye/campito-mnt-tree-ring-data-in-01-mm-n5405-form-3435bc-to-1969ad#!display=line&ds=22ye>
- Dolado, J. J., J. Gonzalo, and M. Laura. 2003. "Long-Range Dependence in Spanish Opinion Poll Series." *Journal of Applied Econometrics* 18:137–55.
- Granger, C. W. J. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14:227–38.

- 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.
- Hosking, J. R. M. 1981. "Fractional Differencing." *Biometrika* 68:165–76.
- Hurst, H. E. 1957. "A Suggested Statistical Model of Some Time Series That Occur in Nature." *Nature* 180:494.