

ESTIMATING THE LONG MEMORY PARAMETER IN NONSTATIONARY MODELS: FURTHER MONTE CARLO EVIDENCE

Margherita Gerolimetto

1. Introduction

In the last decades long memory time series models have been widely examined. Applications of these models can be found in several fields, *e.g.* hydrology, chemistry, economics and finance.

Several estimation techniques have been proposed in literature to detect the long memory phenomenon in both time and frequency domain (see for example Palma, 2007 for a review); they aim at estimating the long memory parameter d that incorporates the strength of the persistence. Most of the methods have been thought, in principle for the stationary case, *i.e.* their theoretical properties hold only when d is in $(-1/2, +1/2)$. Some recent simulation study have been carried out to compare the performance of the long memory parameter estimators in case of stationary models (among the others, Bouthahar *et al.*, 2007 and Tsay, 2009).

Here, we are interested in nonstationary long memory models, *i.e.* when the long memory parameter d no longer is in $(-1/2, +1/2)$, but it actually can be $\geq 1/2$. Broadly speaking, the issue of estimating nonstationary long memory has been addressed in two ways, either extending existing methods to estimate d to the case of nonstationarity (as in Velasco, 1999a and 1999b) or proposing new methods (resorting, for example, to wavelets as in Moulines *et al.*, 2008 and Boubaker and Péguin-Feissolle, 2013).

In this paper, we conduct a Monte Carlo experiment to show and compare the performance of a variety of estimators, traditionally conceived for stationary models, of the long memory parameter d in case of nonstationarity. On purpose, we did not focus on new-generation estimators, but did concentrate on traditional estimators, belonging to three group. Among *(i)* heuristic estimators, we consider the Higuchi method (1988), the aggregate variance method (1995) and Lo (1991) method. Among *(ii)* parametric estimators, we consider Whittle method (Fox and Taqqu, 1986) and among *(iii)* semiparametric methods, we consider the GPH method by Geweke and Porter-Hudak (1983) and its modified version by Smith (2005). All these methods have been employed on both the original time series and first difference of the series. This is done to include in the analysis an idea by

Hurvich and Ray (1995) who propose, in case of nonstationarity, to estimate d on the first difference of the series, i.e. on the series made stationary so that the estimators are expected to work again in the range of d where their properties are guaranteed.

Results of the Monte Carlo experiment show that the Whittle estimator has the best performance in case of nonstationarity, followed by the GPH. Moreover, the strategy of preliminarily differentiate the series helps improve the results.

The structure of the paper is as follows. In the second section we will briefly recall the most important characteristics of the long memory models. In the third section we present the estimators of the long memory parameter we will study. The fourth section is devoted to the Monte Carlo experiment and some conclusions.

2. Long memory models

Usually, a long memory model X_t can be characterized by a single memory parameter $d \in (-1/2, +1/2)$, called degree of the memory, which controls the shape of the spectrum near zero frequency and the hyperbolic rate decay of its autocorrelation function. More precisely the spectral density $f(\lambda)$ of the long memory model is approximated in the neighborhood of the zero frequency by

$$f(\lambda) \sim c\lambda^{-2d} \text{ as } \lambda \rightarrow 0^+, 0 < c < \infty \quad (1)$$

Thus $f(\lambda) \rightarrow \infty$ as $\lambda \rightarrow 0^+$. Under additional regularity assumptions of $f(\lambda)$, the autocorrelation function $\rho(k)$ of the long memory model has the following asymptotic behavior:

$$\rho(k) \sim ck^{2d-1} \text{ as } k \rightarrow \infty \quad (2)$$

these features characterize ARFIMA(p, d, q) models¹ (Granger and Joyeux, 1980)

$$\Phi(B)\Delta^d X_t = \Theta(B)\varepsilon_t \quad (3)$$

of which the fractional noise is a special case

$$\Delta^d X_t = \varepsilon_t \quad (4)$$

¹ ARFIMA models are a generalization of ARIMA, where d is not integer.

The properties of these models depend on the long memory parameter value d . More, specifically, an ARFIMA(p,d,q) model is stationary and invertible when $d \in (-1/2, +1/2)$, usually this interval is reduced to $(0, 1/2)$. When $d \geq 1/2$ the ARFIMA is nonstationary, although for $d \in [1/2, 1)$ it is mean-reverting, meaning that there is no long-run impact of an innovation on the value of the process. When $d \geq 1$ mean-reversion does not longer hold. Clearly, the case $d = 0$ and $d = 1$ (i.e. shot memory stationarity and unit root) are encompassed as particular cases of a more general parametrization.

3. Estimation techniques for ARFIMA processes

Now we briefly describe the methods we will consider in our Monte Carlo experiment to estimate the long memory parameters. For space reason we will not be able to go in details about the methods and refer to the original papers. It is possible to group these methods in three categories: heuristic, parametric and semiparametric methods.

Among heuristic methods, we consider: (a) Higuchi method (Higuchi, 1988) which measures the fractal dimension of a non-periodic and irregular time series; (b) the aggregate variance method (Fox and Taquq, 1995) that concentrates on the behavior of the variance of the sample mean and (c) the rescaled range (R/S) method, Lo (1991), which studies the behavior of the partial sums of deviation of the series from its sample mean.

As for parametric methods, we consider Whittle method (Fox and Taquq, 1986). Given the ARFIMA(p,d,q) in (3), the vector of parameters $\theta = (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ is estimated via the Whittle approximation of the log-likelihood by minimizing with respect to θ :

$$\hat{\sigma}^2(\theta) = \frac{1}{2} \sum_{j=1}^{T'} \frac{I(\lambda_j)}{f(\lambda_j)} \quad (5)$$

where T' is the integer part of $\frac{T-1}{2}$ and $I(\lambda_j)$ and $f(\lambda_j)$ are, respectively, the periodogram and the spectral density at the Fourier frequencies. The Whittle method has several theoretical and practical advantages. However, its disadvantage is in that the parametric form of the spectral density is assumed to be known a priori.

Among the semiparametric methods we consider the GPH estimator (Geweke and Porter-Hudak, 1983). The advantage in resorting to these methods is that there is no need to specify the entire model since the only necessary information is the

behavior of the spectral density near the origin. Given an I(1) process and ARFIMA(p,d,q) as in (5), its spectral density is:

$$f(\lambda) \sim c_f \left(4 \sin^2 \left(\frac{\lambda}{2} \right) \right)^{-d} \quad \lambda \rightarrow \infty \quad (6)$$

As the periodogram $I(\lambda)$ is an asymptotically unbiased estimator of $f(\lambda)$, it is possible to estimate d by running the OLS regression:

$$\log(I(\lambda_j)) = \log c_f + \beta \log \left(4 \sin^2 \left(\frac{\lambda}{2} \right) \right) + \varepsilon_j \quad (7)$$

The former is an asymptotic relationship that holds only in a neighborhood of the origin. This means that, considered over all the ordinates of the periodogram ($-\pi, +\pi$), it would produce highly biased estimates. Consequently, GPH is so actually calculated by running the least squares regression only for the m lowest frequencies. We also considered Smith's (2005) modified version of the GPH method that takes into account the approximation he derived of the bias.²

As anticipated, the 5 methods have been thought for the stationary setting. The theoretical properties no longer hold in case of nonstationary, or, in case they do, it is only for a limited interval.³ In case of nonstationarity the relative recent literature is rich of contributes along two directions. There are estimators that adapt existing methods in order to gain asymptotic properties also in case of nonstationarity; among these we can mention the tapered versions of the GPH or the Whittle method (Velasco and Robinson 1999a, 1999b). There are also brand-new methods, e.g. wavelet based estimators (McCoy and Walden, 1996; Moulined et al. 2008).

As for the brand-new methods, it should be stressed that often these methods are much more sophisticated (and complicated) than the existent. For this reason, in this paper we study, via Monte Carlo simulations, the effective performance of the traditional methods in case of non stationary long memory, also when they are employed on the first difference of the time series (now stationary) following Hurvich and Ray (1995).

² Actually, in our Monte Carlo experiment, the performance of the version of the GPH estimator modified by Smith (2005) is not particularly good.

³ For example the Whittle estimator is shown to possess asymptotic properties for $1/2 < d < 3/4$, included asymptotic normality. The same holds for the GPH.

4. Monte Carlo experiment

In this section we present the Monte Carlo experiment we conduct to show and compare the performance of the 5 estimators of the long memory parameter described in the previous sections: Higuchi, Aggregate Variance, Lo, GPH, GPH modified by Smith (GPH-S, hereafter) Whittle. The Data Generating Process (DGP, hereafter) we consider is the fractional noise, ARFIMA (0, d ,0), for various values of the long memory parameter.

In particular, we considered three scenarios. In the first we consider stationary DGPs and we simulate fractional noise with $d=0.1,0.2,0.3,0.4$. This scenario is included in the MC experiment with the role of benchmark, given that all the long memory parameter estimators should have a good performance in this case. In the second scenario, we generate time series with $d=0.6,0.7,0.8,0.9$, i.e. nonstationary but mean reverting long memory. In the third scenario, we study nonstationary and not mean-reverting long memory, done by generating time series data with $d=1.1,1.2,1.3,1.4$. Over all cases, the innovation is $\varepsilon_t \sim N(0,1)$, the sample size considered are $T=250,500,1000$ for 2000 Monte Carlo simulations. All series are generated with 200 additional values in order to obtain random starting values. The performance of the estimators is expressed in terms of mean squared error (MSE) across Monte Carlo simulations.⁴

The results of the experiment are reported in the tables 1-5. In Table 1 we present the MSE for the stationary case.

Table 1 – Stationary long memory: Monte Carlo MSE

| T | d | R/S | Aggr Var | Higuchi | GPH | GPH-S | Whittle |
|------|-----|--------|----------|---------|--------|--------|---------|
| 250 | 0.1 | 0.0117 | 0.009 | 0.014 | 0.0494 | 0.1672 | 0.0031 |
| | 0.2 | 0.0137 | 0.0126 | 0.0144 | 0.0526 | 0.1653 | 0.0034 |
| | 0.3 | 0.0162 | 0.0163 | 0.0173 | 0.0476 | 0.1617 | 0.0033 |
| | 0.4 | 0.0207 | 0.0214 | 0.0141 | 0.0452 | 0.1804 | 0.0034 |
| 500 | 0.1 | 0.0085 | 0.0073 | 0.0116 | 0.0296 | 0.0869 | 0.0013 |
| | 0.2 | 0.0109 | 0.0087 | 0.0132 | 0.0296 | 0.0902 | 0.0016 |
| | 0.3 | 0.0131 | 0.0118 | 0.0153 | 0.029 | 0.091 | 0.0015 |
| | 0.4 | 0.0158 | 0.0174 | 0.0147 | 0.0303 | 0.09 | 0.0015 |
| 1000 | 0.1 | 0.0069 | 0.0057 | 0.012 | 0.0193 | 0.0193 | 0.0007 |
| | 0.2 | 0.0092 | 0.0076 | 0.0135 | 0.0196 | 0.0196 | 0.0007 |
| | 0.3 | 0.011 | 0.0099 | 0.0141 | 0.0188 | 0.0188 | 0.0007 |
| | 0.4 | 0.0149 | 0.0158 | 0.0137 | 0.0194 | 0.0194 | 0.0007 |

⁴ For the GPH the estimation has been conducted setting m equal to the square root of the sample size, as suggested in the original article by Geweke and Porter-Hudak (1983).

In Table 2 we present the MSE results for the nonstationary mean reverting case, both for the original series (upper panel) and the first difference of the series (lower panel).

Table 2 – *Nonstationary (mean reverting) long memory: Monte Carlo MSE (original series upper panel, first differenced series lower panel)*

| T | d | R/S | Aggr Var | Higuchi | GPH | GPH-S | Whittle |
|--------------------------|-----|--------|----------|---------|--------|--------|---------|
| Original series | | | | | | | |
| 250 | 0.6 | 0.0393 | 0.0529 | 0.0323 | 0.0511 | 0.1696 | 0.0035 |
| | 0.7 | 0.0567 | 0.0846 | 0.0645 | 0.0474 | 0.1385 | 0.0036 |
| | 0.8 | 0.0872 | 0.13 | 0.1155 | 0.0464 | 0.1461 | 0.0037 |
| | 0.9 | 0.125 | 0.1965 | 0.1858 | 0.0454 | 0.1465 | 0.0033 |
| 500 | 0.6 | 0.0325 | 0.047 | 0.0352 | 0.0308 | 0.0931 | 0.0016 |
| | 0.7 | 0.0509 | 0.0791 | 0.0656 | 0.0316 | 0.0943 | 0.0019 |
| | 0.8 | 0.0838 | 0.1278 | 0.1192 | 0.031 | 0.0903 | 0.0019 |
| | 0.9 | 0.1267 | 0.1899 | 0.1861 | 0.028 | 0.0805 | 0.0017 |
| 1000 | 0.6 | 0.0314 | 0.0444 | 0.0339 | 0.0202 | 0.0202 | 0.0008 |
| | 0.7 | 0.0517 | 0.0767 | 0.0667 | 0.0188 | 0.0188 | 0.0009 |
| | 0.8 | 0.0837 | 0.1237 | 0.116 | 0.0204 | 0.0204 | 0.0012 |
| | 0.9 | 0.1323 | 0.1888 | 0.1863 | 0.0184 | 0.0184 | 0.0011 |
| First differenced series | | | | | | | |
| 250 | 0.6 | 0.0434 | 0.0102 | 0.0084 | 0.0488 | 0.1672 | 0.0027 |
| | 0.7 | 0.0275 | 0.0073 | 0.0065 | 0.05 | 0.1561 | 0.0035 |
| | 0.8 | 0.0184 | 0.0064 | 0.0074 | 0.0464 | 0.1678 | 0.0031 |
| | 0.9 | 0.0126 | 0.0076 | 0.01 | 0.0517 | 0.1736 | 0.0037 |
| 500 | 0.6 | 0.0337 | 0.0067 | 0.0062 | 0.0313 | 0.0948 | 0.0014 |
| | 0.7 | 0.0201 | 0.0044 | 0.005 | 0.0295 | 0.0894 | 0.0014 |
| | 0.8 | 0.0123 | 0.0042 | 0.0063 | 0.0314 | 0.0915 | 0.0013 |
| | 0.9 | 0.0085 | 0.0048 | 0.0076 | 0.0288 | 0.0892 | 0.0016 |
| 1000 | 0.6 | 0.0259 | 0.0046 | 0.0046 | 0.0199 | 0.0199 | 0.0007 |
| | 0.7 | 0.0147 | 0.003 | 0.0044 | 0.0187 | 0.0187 | 0.0007 |
| | 0.8 | 0.0088 | 0.0033 | 0.0056 | 0.0175 | 0.0175 | 0.0007 |
| | 0.9 | 0.0065 | 0.004 | 0.0073 | 0.018 | 0.018 | 0.0007 |

In Table 3 we present the MSE results for the nonstationary not mean-reverting case, both for the original series (upper panel) and the first difference of the series (lower panel).

Table 3 – *Nonstationary (nont mean reverting) long memory: Monte Carlo MSE (original series upper panel, first differenced series lower panel)*

| T | d | R/S | Aggr Var | Higuchi | GPH | GPH-S | Whittle |
|-----------------|---|-----|----------|---------|-----|-------|---------|
| Original series | | | | | | | |

| | | | | | | | |
|--------------------------|-----|--------|--------|--------|--------|--------|--------|
| 250 | 1.1 | 0.2613 | 0.3867 | 0.3807 | 0.0396 | 0.1284 | 0.0066 |
| | 1.2 | 0.332 | 0.5103 | 0.5095 | 0.0487 | 0.1295 | 0.0242 |
| | 1.3 | 0.448 | 0.6576 | 0.6525 | 0.0821 | 0.1314 | 0.0682 |
| | 1.4 | 0.5142 | 0.8236 | 0.8207 | 0.1398 | 0.177 | 0.1353 |
| 500 | 1.1 | 0.2663 | 0.3778 | 0.3841 | 0.0223 | 0.0607 | 0.0051 |
| | 1.2 | 0.3766 | 0.5039 | 0.5077 | 0.0377 | 0.0745 | 0.0236 |
| | 1.3 | 0.4989 | 0.6494 | 0.6571 | 0.071 | 0.0912 | 0.0651 |
| | 1.4 | 0.5861 | 0.8147 | 0.8217 | 0.1348 | 0.1446 | 0.1402 |
| 1000 | 1.1 | 0.2994 | 0.3789 | 0.384 | 0.0193 | 0.0193 | 0.0046 |
| | 1.2 | 0.3815 | 0.5053 | 0.5116 | 0.0309 | 0.0309 | 0.0236 |
| | 1.3 | 0.5208 | 0.6503 | 0.6564 | 0.0681 | 0.0681 | 0.0683 |
| | 1.4 | 0.6095 | 0.8162 | 0.8213 | 0.1352 | 0.1352 | 0.1413 |
| First differenced series | | | | | | | |
| 250 | 1.1 | 0.0129 | 0.009 | 0.0144 | 0.0488 | 0.1752 | 0.0033 |
| | 1.2 | 0.0143 | 0.011 | 0.0148 | 0.0453 | 0.1727 | 0.0031 |
| | 1.3 | 0.0157 | 0.0141 | 0.0156 | 0.0466 | 0.1672 | 0.0034 |
| | 1.4 | 0.0208 | 0.0214 | 0.0188 | 0.046 | 0.1596 | 0.0033 |
| 500 | 1.1 | 0.0089 | 0.007 | 0.0112 | 0.0279 | 0.0952 | 0.0015 |
| | 1.2 | 0.0106 | 0.0077 | 0.0136 | 0.0278 | 0.0897 | 0.0014 |
| | 1.3 | 0.0135 | 0.0111 | 0.0146 | 0.0301 | 0.0941 | 0.0014 |
| | 1.4 | 0.0182 | 0.0178 | 0.0142 | 0.0309 | 0.0933 | 0.0015 |
| 1000 | 1.1 | 0.0076 | 0.0061 | 0.0119 | 0.0183 | 0.0183 | 0.0007 |
| | 1.2 | 0.0093 | 0.0074 | 0.0126 | 0.0194 | 0.0194 | 0.0007 |
| | 1.3 | 0.0118 | 0.0105 | 0.0144 | 0.019 | 0.019 | 0.0007 |
| | 1.4 | 0.0144 | 0.0156 | 0.0127 | 0.0188 | 0.0188 | 0.0007 |

From Table 1 (stationarity case) we can observe that while d is far from the nonstationarity area, almost all estimation methods have a low level of MSE, also at relatively small sample sizes. It is in particular when d gets close to the bound $\frac{1}{2}$ that it is possible to appreciate the better performance of the Whittle method, followed by the GPH and Higuchi methods, as the other methods worsen their performance visibly.

In Table 2, upper panel, we observe for the majority of methods the process of worsening of the MSE performance with the increase of d continues. Only for Whittle and GPH estimators the performance is steadily good, more precisely not only they are the methods with the best performance, but also their MSE level stays approximately at the same level as in Table 1. This means that the two methods do not suffer excessively from the lack nonstationarity (probably because mean-reversion still holds). In general, things improve when all 5 methods are applied to the first difference of the time series (lower panel of Table 2). However, we note that for Whittle method in particular, there seems to be no relevant difference from

upper panel and lower panel, leading us to believe that for this methods differencing is not necessary.

In Table 3 we study the nonstationary and non mean-reverting scenario. In this case, for all methods this is a rather difficult task because we are very far from the area where the theoretical properties hold. Taking the first difference (lower panel) leads to quite better results, especially if Whittle and GPH are adopted. So in this case, taking the first difference seems to be really a reasonable option, that leads to good MSE performance (in line with the stationary case magnitude order), especially if Whittle and GPH are used.

In Table 4 and 5 we present for the nonstationary (respectively mean reverting and non-meanreverting) case, the ratio of the MSE of the estimate on the original series and on its first difference. These Tables help emphasize the effective improvement in adopting the first difference and under which conditions this happens.

Table 4 – *Nonstationary (mean reverting) long memory: ratio of Monte Carlo MSE of the estimate on the original series and on first differenced series*

| T | d | R/S | Aggr Var | Higuchi | GPH | GPH-S | Whittle |
|------|-----|--------|----------|----------|----------|---------|---------|
| 250 | 0.6 | 0.547 | 2.975 | 2.2985 | 0.9141 | 0.8133 | 0.1221 |
| | 0.7 | 0.986 | 6.9778 | 6.5715 | 1.9012 | 2.0766 | 0.1789 |
| | 0.8 | 1.6752 | 22.0144 | 18.624 | 3.7096 | 3.0392 | 0.5752 |
| | 0.9 | 3.0927 | 54.9223 | 64.572 | 1.6946 | 1.557 | 0.0138 |
| 500 | 0.6 | 0.5446 | 3.2609 | 2.7961 | 0.963 | 0.7935 | 0.5118 |
| | 0.7 | 1.0532 | 9.1372 | 9.017 | 2.3609 | 2.6762 | 1.0504 |
| | 0.8 | 1.9897 | 41.7805 | 60.1748 | 6.8971 | 4.3158 | 1.1782 |
| | 0.9 | 3.7907 | 68.415 | 57.1477 | 3.2824 | 2.158 | 0.7604 |
| 1000 | 0.6 | 0.5768 | 3.8248 | 3.3794 | 1.1626 | 1.1626 | 1.3785 |
| | 0.7 | 1.1811 | 10.903 | 11.7163 | 1.787 | 1.787 | 3.8535 |
| | 0.8 | 2.3524 | 56.6822 | 577.6202 | 3.3903 | 3.3903 | 2.4796 |
| | 0.9 | 4.5605 | 58.3148 | 50.1638 | 18.44891 | 18.4489 | 2.7687 |

When the figures in the Tables are smaller than 1, this means that the MSE coming from the estimate on the first differenced time series is larger than that on the original series. On the contrary, the larger the figures are with respect to 1, the more recommended is to estimate d on the first difference of the time series.

As expected, in Table 4, regarding nonstationary mean-reverting time series, the figures have an oscillatory behavior around 1, especially for the Whittle method, thus confirming what emerged from Table 2, i.e. if $\frac{1}{2} < d < 1$ the effects of nonstationarity are non that severe and, consequently, is not so relevant (sometime

not even visible) the improvement in the performance obtained by adopting the strategy of taking the first difference of the series before estimating d .

Table 5 – *Nonstationary (not mean reverting) long memory: ratio of Monte Carlo MSE of the estimate on the original series over first differenced series*

| T | d | R/S | Aggr Var | Higuchi | GPH | GPH-S | Whittle |
|------|-----|----------|----------|---------|----------|----------|----------|
| 250 | 1.1 | 16.79 | 16.4638 | 26.3091 | 41.9561 | 10.7115 | 3.2634 |
| | 1.2 | 93.086 | 12.4855 | 25.3358 | 13.9063 | 4.5804 | 8.2807 |
| | 1.3 | 16.3558 | 9.3688 | 23.0359 | 22.4873 | 18.5583 | 13.2683 |
| | 1.4 | 9.8334 | 7.3744 | 18.1095 | 35.0908 | 267.265 | 25.7747 |
| 500 | 1.1 | 21.9624 | 17.3914 | 36.1323 | 34.2809 | 7.09977 | 5.2897 |
| | 1.2 | 253.9619 | 14.8114 | 26.0578 | 44.379 | 7.83361 | 17.8265 |
| | 1.3 | 23.1643 | 10.5538 | 25.0266 | 53.8314 | 414.2553 | 27.0515 |
| | 1.4 | 11.1353 | 8.12 | 24.4066 | 59.2308 | 31.3384 | 45.5639 |
| 1000 | 1.1 | 90.7188 | 18.8324 | 24.8224 | 14.458 | 14.458 | 11.8739 |
| | 1.2 | 86.4642 | 14.6381 | 30.2226 | 32.4125 | 32.4125 | 27.2745 |
| | 1.3 | 20.3471 | 10.7909 | 26.324 | 144.1566 | 144.1566 | 44.5244 |
| | 1.4 | 13.7794 | 8.5732 | 28.0049 | 25.8235 | 25.8235 | 205.3707 |

In Table 5, instead, all figures are systematically larger than 1. This is because for all considered estimation methods (even for the Whittle), the performance hugely improves in case the first difference is preliminarily taken. Once more, this is in line with the previous results, in particular those shown in Table 3. The effects of nonstationarity are very severe and, consequently, it is significant the improvement in the performance obtained by adopting the strategy of taking the first difference of the series before estimating d .

5. Conclusions

To conclude, in this work we present a Monte Carlo study to show and compare the performance of some traditional and well-known estimator of the long memory parameter in case of nonstationary fractional noise models. We are aware that in the literature recently has been proposed a variety of methods for estimating the long memory parameter in the nonstationary case, yet we are interested in how the traditional methods perform in case the first difference of the series is taken and in this work we intend to fill this literature gap.

The simulation results show that, among the traditional methods the Whittle estimate (followed by the GPH) is the best performing in terms of Monte Carlo MSE and this holds also when stationarity no longer holds, in particular if mean-

reversion is preserved. Indeed, if the nonstationary time series is mean reverting the performance of the Whittle estimator is comparable with the stationary case and there seem to be no special need to preliminarily take the first difference. Instead, when the nonstationarity is so strong that mean-reversion is lost and all methods perform badly, working with first difference of the time series (in particular estimating with Whittle method) is recommended.

To sum up, we conclude that in several cases it could be that there is no need to resort to sophisticated (and difficult to implement) methods for estimating nonstationary long memory. It may happen that taking the first difference of the series and then proceed with the traditional estimators, especially Whittle estimator is a good enough strategy to obtain reliable estimates of the long memory parameter in the nonstationary hypotheses.

Future research on this topic is in order with the aim of extending the simulation experiment so that also new-generation method, such as wavelets methods will be included.

Riferimenti bibliografici

- BOUBAKER H., PÉGUIN-FEISSOLLE A. 2013: Estimating the long memory parameter in nonstationary processes using wavelets, *Computational Economics*, Vol. 42, pp. 291-306.
- BOUTHAHAR M., MARIMOUTOU V., NOUIRA L. 2007: Estimation methods of the long memory parameter: Monte Carlo analysis and application, *Journal of Applied Statistics*, Vol. 34, pp. 261-301.
- HURVICH C.M., RAY B.K. 1995: Estimation of the long memory parameter for nonstationary or non invertible fractionally integrated processes, *Journal of Time Series Analysis*, Vol.16, pp. 17-41.
- FOX R., TAQQU M. 1986. Large sample properties of parameter estimates for strongly dependent stationary gaussian time series, *The Annals of Statistics*, Vol. 14, pp. 517-532.
- GEWEKE J., PORTER-HUDAK S. 1983. The estimation and application of long-memory time series models, *Journal of Time Series Analysis*, Vol.4, pp. 221-237.
- GRANGER C.W.J., JOYEUX, R. 1980. An introduction to long-range time series models and fractional differencing, *Journal of Time Series Analysis*, Vol.1, pp. 15-30.
- HIGUCHI T. 1988. Approach to an irregular time series on the basis of the fractal theory, *Physica D*, Vol. 31, pp. 277-283
- LO A.W. 1991: Long-term memory in stock market prices, *Econometrica*, Vol. 59, pp. 1279-313.

- McCOY E.J., WALDEN A.T. 1996. Wavelet analysis and synthesis of stationary long memory processes, *Journal of Computational and Graphical Statistics*, Vol. 5, pp. 26-56.
- MOULINES E., ROUEFF F., TAQQU M.S. 2008. A wavelet whittle estimator of the long memory parameter gaussian time series, *The Annals of Statistics*, Vol. 36, pp. 1925-1956.
- PALMA W. 2007. *Long memory time series: theory and methods*. New York: Wiley.
- SMITH A. 2005. Level shifts and the illusion of long memory in economic time series, *Journal of Business and Economic Statistics*, Vol. 23, pp. 231-335.
- TAQQU M., TEVEROVSKY V., WILLINGER W. 1995. Estimators for long range dependence: an empirical study, *Fractals*, Vol. 3, pp. 785-798.
- TSAY W.J. 2009: Estimating long memory time series cross-section-data parameter, *Electoral Studies*, Vol. 28, pp. 129-140.
- VELASCO C. 1999a. Nonstationary logperiodogram regression, *Journal of Econometrics*, Vol. 91, pp. 325-371.
- VELASCO C. 1999b. Gaussian semiparametric estimation of nonstationary time series, *Journal of Time Series Analysis*, Vol. 20, pp. 87-127.

SUMMARY**Estimating the Long Memory Parameter in Nonstationary Models: Further Monte Carlo Evidence**

In this work we perform a Monte Carlo experiment to show and compare the performance of a variety of estimators of the long memory parameter d in case of nonstationary processes. Both parametric and semiparametric estimators are considered. Moreover they have been employed both on the original time series and on the first difference of the series. Results show that the Whittle estimator is the best performing and the strategy of preliminarily differentiate the series is worthy, but not for all the estimators.

Margherita GEROLIMETTO, Professore Associato,
margherita.gerolimetto@unive.it