

Article

Reducing the Bias of the Smoothed Log Periodogram Regression for Financial High-Frequency Data

Erhard Reschenhofer * and Manveer K. Mangat

Department of Statistics and Operations Research, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria; manveer.mangat@univie.ac.at

* Correspondence: erhard.reschenhofer@univie.ac.at; Tel.: +43-1-4277-38646

Received: 8 March 2020; Accepted: 27 September 2020; Published: 10 October 2020



Abstract: For typical sample sizes occurring in economic and financial applications, the squared bias of estimators for the memory parameter is small relative to the variance. Smoothing is therefore a suitable way to improve the performance in terms of the mean squared error. However, in an analysis of financial high-frequency data, where the estimates are obtained separately for each day and then combined by averaging, the variance decreases with the sample size but the bias remains fixed. This paper proposes a method of smoothing that does not entail an increase in the bias. This method is based on the simultaneous examination of different partitions of the data. An extensive simulation study is carried out to compare it with conventional estimation methods. In this study, the new method outperforms its unsmoothed competitors with respect to the variance and its smoothed competitors with respect to the bias. Using the results of the simulation study for the proper interpretation of the empirical results obtained from a financial high-frequency dataset, we conclude that significant long-range dependencies are present only in the intraday volatility but not in the intraday returns. Finally, the robustness of these findings against daily and weekly periodic patterns is established.

Keywords: long-range dependence; log periodogram regression; smoothed periodogram; subsampling; intraday returns

JEL Classification: C13; C14; C22; C58

1. Introduction

After Mandelbrot (1971) had discussed the possibility that the strength of the statistical dependence of stock prices decreases very slowly, several researchers investigated this issue empirically. For example, Greene and Fielitz (1977) found indications of long-range dependence when they applied a technique called range over standard deviation (R/S) analysis (Hurst 1951; Mandelbrot and Wallis 1969; Mandelbrot 1972, 1975) to daily stock return series. This technique is based on the R/S statistic Q_n , which is defined as the range of all partial sums of a time series of length n from its mean divided by its standard deviation. For a large class of short-range dependent processes, Q_n/n^H converges to a non-degenerate random variable if $H = 0.5$. An analogous result with $H \neq 0.5$ holds for long-range dependent processes. The parameter H is called the Hurst coefficient and is used as a measure of long-range dependence. However, Lo (1991) pointed out that the results obtained with this technique may be misleading because of the sensitivity of Q_n to short-range dependence (see also Davis and Harte 1987; Hauser and Reschenhofer 1995) and proposed, therefore, a Newey and West (1987) type modification for the denominator of the R/S statistic, which is appropriate for general forms of short-range dependence. Contrary to the findings of Greene and Fielitz (1977) and others, he found no evidence of long-range dependence in daily and monthly index returns once the possible short-range dependence was properly taken care of. A disadvantage of Lo's (1991)

modified R/S analysis is its dependence on an important tuning parameter, namely the truncation lag q , which determines the number of included autocovariances. The general conditions that ensure the consistency of the Newey and West estimator provide little guidance in selecting q in finite samples. Additionally, Andrews's (1991) data-dependent rule for choosing q is based on asymptotic arguments.

Long-range dependence can not only be characterized by a Hurst coefficient $H \neq 0.5$ but also by a slowly decaying autocorrelation function ρ or a spectral density f that is steep in a small neighborhood of frequency zero, i.e.,

$$\rho(k)k^{1-2d} \rightarrow c \text{ as } k \rightarrow \infty, c > 0, d \neq 0, \quad (1)$$

and:

$$f(\omega) \sim c\omega^{-2d}, \omega \in (0, \varepsilon), c > 0, d \neq 0, \quad (2)$$

respectively. The parameter d is called a memory parameter (or fractional differencing parameter) and is related to H by $d = H - 0.5$. It can be estimated by replacing the unknown spectral density f in (2) by the periodogram (Geweke and Porter-Hudak 1983) or a more sophisticated estimate of f (Hassler 1993; Peiris and Court 1993; Reisen 1994), taking the log of both sides, and regressing the log estimate on a deterministic regressor. Robustness against short-range dependence can be achieved by using only the $K = n^\alpha$ lowest Fourier frequencies in the regression. A popular choice for the tuning parameter α is 0.5. For the purpose of testing, the asymptotic error variance is used. Applying the log periodogram regression method of Geweke and Porter-Hudak (1983) to the daily returns of the 30 components of the Dow Jones Industrials index and several indices, Barkoulas and Baum (1996) found no convincing evidence in favor of long-range dependence, which is not surprising in light of the finding of Mangat and Reschenhofer (2019) that the test based on the asymptotic error variance has very low power. Unfortunately, using the standard variance formula of the least squares estimator of the slope in a simple linear regression instead of the asymptotic error variance is also problematic because it leads to overrejecting the true null hypothesis (see Mangat and Reschenhofer 2019).

The negative results of Lo (1991) and Barkoulas and Baum (1996) are in line with the results obtained by Cheung and Lai (1995) with both modified R/S analysis and log periodogram regression for stock return data from eighteen countries and by Crato (1994) with fractionally integrated ARMA (ARFIMA) models (Granger and Joyeux 1980; Hosking 1981) for stock indices of the G-7 countries. Using not only the log periodogram regression with the asymptotic error variance but additionally also nonparametric techniques such as R/S analysis and modified R/S analysis as well as parametric techniques, Grau-Carles (2000) also found little evidence of long-range dependence in index returns but strong evidence of persistence in volatility measured as squared returns and absolute returns, respectively, which corroborates earlier findings of Crato and de Lima (1994) and Lobato and Savin (1998). In general, results obtained with ARFIMA models must be treated with caution. Firstly, the true model dimension is unknown in practice and reliable inference after automatic model selection is illusory. Secondly, Pötscher (2002) has shown that the problem of estimating the memory parameter d falls into the category of ill-posed estimation problems when the class of data generating processes is too rich. For example, Grau-Carles (2000) considered all ARFIMA(p, q) processes with $p \leq 3$ and $q \leq 3$, which is possibly an unnecessarily large class for return series.

While the bulk of empirical research focused on major capital markets, Barkoulas et al. (2000) examined an emerging capital market, namely the Greek stock market, with the log periodogram regression and obtained significant estimates of d in the range between 0.20 and 0.30 for values of the tuning parameter α between 0.5 and 0.6. However, their sample period is relatively short and the sampling frequency is weekly rather than daily. Even less confidence-inspiring are the positive results obtained by Henry (2002) with monthly data from several international stock markets. Clearly, methods that have been designed for large samples should not be applied to small and medium samples. Recently, small-sample tests for testing hypotheses about the memory parameter d have been proposed (Mangat and Reschenhofer 2019; Reschenhofer and Mangat 2020). When applied to asset returns, these tests produced negative results throughout. Cajueiro and Tabak (2004), Carbone et al. (2004),

Batten and Szilagyi (2007), Batten et al. (2008), Souza et al. (2008), Batten et al. (2013), and Auer (2016a, 2016b) observed time-variability of the Hurst exponent in stock returns, currency prices, and the prices of precious metals, respectively. These apparent changes were occasionally interpreted as indications of changing market efficiency or even used for the construction of trading strategies. Although it cannot be ruled out that some erratic estimator for the memory parameter d catches signals that are useful for trading purposes even when in fact there is no long-range dependence, there still seems to be a need for a more efficient estimator that actually allows to get some information about the true nature of the data generating process.

In general, there is always a trade-off between bias and variance. Estimators for the memory parameter d that are based on a smooth estimate of the spectral density have typically a smaller variance and a larger bias than those based on the periodogram (Chen et al. 1994; Reschenhofer et al. 2020), which is advantageous in situations where the squared bias is small relative to the variance. However, in the case of high-frequency financial data, there are usually gaps between the individual trading sessions, which make it necessary to estimate d separately for each trading session and compute the final estimate by averaging the individual estimates. Here, the variance decreases with the number of trading sessions but the bias remains fixed; hence, conventional smoothing methods, which achieve a reduction in the variance at the expense of an increase in the bias, are of no use. The goal of this paper is therefore to introduce a new method of smoothing that does not systematically have a negative impact on the bias. This method will be described in detail in the next section. Section 3 presents the results of an extensive simulation study, which compares the performance of various estimators for the memory parameter in terms of bias, variance, and root-mean-square error (RMSE). Using limit order book data obtained from Lobster, Section 4 searches for indications of long-range dependence both in the intraday volatility and in the intraday returns. Section 5 provides a conclusion.

2. Methods

2.1. Log Periodogram Regression

Fractionally integrated white noise satisfies the difference equation:

$$y_t = (1 - L)^{-d} u_t, \quad (3)$$

where L is the lag operator and u_t is white noise with mean zero and variance σ^2 (Adenstedt 1974). Its spectral density is given by:

$$f(\omega) = \frac{\sigma^2}{2\pi} |1 - e^{-i\omega}|^{-2d} = \frac{\sigma^2}{2^{1+2d} \pi} \left(\sin^2\left(\frac{\omega}{2}\right) \right)^{-d}. \quad (4)$$

The memory parameter d , which represents the degree of long memory if $d \neq 0$, can be estimated by regressing the log periodogram of the time series y_1, \dots, y_n on a deterministic regressor (Geweke and Porter-Hudak 1983). Indeed, we have:

$$L_j = \log I(\omega_j) = c + d x_j + v_j, \quad (5)$$

where:

$$I(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^n y_t e^{-i\omega t} \right|^2. \quad (6)$$

is the periodogram,

$$\omega_j = 2\pi j/n, \quad j = 1, \dots, K \leq m = \lfloor (n-1)/2 \rfloor, \quad (7)$$

are the first K Fourier frequencies between 0 and π ,

$$x_j = -2 \log\left(\sin\left(\frac{\omega_j}{2}\right)\right) \tag{8}$$

is a deterministic regressor,

$$c = \log\left(\sigma^2 / (2^{1+2d} \pi)\right) \tag{9}$$

is a constant, and

$$v_j = \log\left(I(\omega_j) / f(\omega_j)\right). \tag{10}$$

are random perturbations. Choosing $K \ll m$ rather than $K = m$ is advisable when it is suspected that not only long-term dependencies are present but also short-term dependencies, e.g., when the data come from an ARFIMA process:

$$y_t = \left(1 - \phi_1 L - \dots - \phi_p L^p\right)^{-1} (1 - L)^{-d} \left(1 + \theta_1 L + \dots + \theta_q L^q\right) u_t \tag{11}$$

(Granger and Joyeux 1980; Hosking 1981), where the parameter d takes care of the former dependencies and the parameters $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ take care of the latter. It is assumed that $d < 0.5$ (stationarity condition), $d > -0.5$ (invertibility condition), and all roots of the lag operator polynomials $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$ and $\Theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ lie outside the unit circle (causality condition and invertibility condition, respectively).

In the special case of $d = p = q = 0$ and Gaussianity, the ratios $I(\omega_j) / f(\omega_j)$ are independent and identically distributed (i.i.d.) standard exponential and v_1, \dots, v_m are, therefore, i.i.d. Gumbel with mean $-\gamma$ and variance $\pi^2/6$, where $\gamma = 0.57721 \dots$ is Euler's constant. The variance of the Geweke Porter-Hudak (GPH) estimator \hat{d}_{GPH} is then identical to the variance of the ordinary least squares (OLS) estimator for the slope in a simple regression model, i.e.,

$$var(\hat{d}_{GPH}) = \frac{\sigma_v^2}{S_{xx}} = \frac{\pi^2}{6S_{xx}}, \tag{12}$$

where:

$$S_{xx} = \sum_{t=1}^K (x_t - \bar{x})^2. \tag{13}$$

In a neighborhood of frequency zero:

$$\sin(\omega) \approx \omega, \tag{14}$$

Hence:

$$S_{xx} \approx 4 \sum_{t=1}^K (\log(t) - \overline{\log(t)})^2. \tag{15}$$

Furthermore:

$$\begin{aligned} \int_1^K \log^2(t) - \frac{1}{K} \left(\int_1^K \log(t)\right)^2 &= K \log^2(K) - 2K \log(K) + 2(K-1) \\ -\frac{1}{K} (K \log(K) - (K-1))^2 &= K + o(K). \end{aligned} \tag{16}$$

Indeed, we have:

$$s_{xx} = 4(K + o(K)) \tag{17}$$

If:

$$K \log(K) / n \rightarrow 0 \tag{18}$$

(see Hurvich and Beltrao 1993), hence, the variance formula (10) becomes:

$$var(\hat{d}_{GPH}) \approx \frac{\pi^2}{24K} \tag{19}$$

in line with the asymptotic result:

$$\sqrt{K}(\hat{d}_{GPH} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right) \quad (20)$$

derived by [Hurvich et al. \(1998\)](#) under the assumption that $K = o(n^{4/5})$ and $\log^2(n) = o(K)$ for a class of stationary Gaussian long-memory processes with spectral densities of the form:

$$f(\omega) = |1 - e^{-i\omega}|^{-2d} f^*(\omega), \quad (21)$$

which includes all stationary ARFIMA processes.

If $d \neq 0$, the ratios $I(\omega_j)/f(\omega_j)$ are neither independent nor identically distributed, not even asymptotically ([Robinson 1995](#)). The problem is the irregular behavior of the spectral density in the neighborhood of frequency zero, i.e., $f(\omega) \rightarrow \infty$ as $\omega \rightarrow 0$ if $d > 0$ and $f(\omega) \rightarrow 0$ as $\omega \rightarrow 0$ if $d < 0$. [Robinson \(1995\)](#), therefore, proposed to remove the lowest Fourier frequencies from the log periodogram regression. [Künsch \(1986\)](#) showed that in the case of ARFIMA processes, the ratios $I(\omega_j)/f(\omega_j)$, $j = H + 1, \dots, H + K$ are indeed asymptotically i.i.d. standard exponential provided that $(H + 1)/\sqrt{n} \rightarrow \infty$ and $(H + K)/n \rightarrow 0$. However, [Reisen et al. \(2001\)](#) and [Mangat and Reschenhofer \(2019\)](#) found that even the removal of only the first Fourier frequency already has a negative effect on the performance of the estimator \hat{d}_{GPH} .

2.2. Smoothing the Periodogram

An obvious possibility to further develop the estimator \hat{d}_{GPH} is to smooth the periodogram before it is used in the regression (5) ([Hassler 1993](#); [Peiris and Court 1993](#); [Reisen 1994](#)). In order to illustrate the effect of smoothing, we consider the simple case of $K/3$ non-overlapping averages:

$$(I(\omega_{j-1}) + I(\omega_j) + I(\omega_{j+1}))/3, \quad j = 2, 5, 8, \dots, K - 1. \quad (22)$$

In this case, the sample size is divided by three but at the same time the variance of the error term decreases approximately from:

$$\text{var}\left(\log\left(\frac{I(\omega_j)}{f(\omega_j)}\right)\right) \approx \frac{\pi^2}{6} \quad (23)$$

to the variance of the log chi-square distribution with 6 degrees of freedom because:

$$\text{var}\left(\log\left(\frac{I(\omega_{j-1}) + I(\omega_j) + I(\omega_{j+1}))}{3f(\omega_j)}\right)\right) \approx \text{var}\left(\log\left(\frac{2I(\omega_{j-1})}{f(\omega_{j-1})} + \frac{2I(\omega_j)}{f(\omega_j)} + \frac{2I(\omega_{j+1})}{f(\omega_{j+1})}\right)\right). \quad (24)$$

Noting that the mean (first cumulant) and the variance (second cumulant) of the log chi-square distribution with k degrees of freedom are given by:

$$\kappa_1 = \log(2) + \psi(k/2) \quad (25)$$

and:

$$\kappa_2 = \psi'(k/2), \quad (26)$$

respectively. we obtain for $k = 6$, $\kappa_1 = 1.615932$ and $\kappa_2 = 0.3949341$. Here, ψ is the digamma function and ψ' is its first derivative. Overall, the (approximate) variance of the least squares estimator of the memory parameter d decreases from

$$\frac{\pi^2}{6} \frac{1}{4K} = 1.644934 \frac{1}{4K} \quad (27)$$

to

$$\psi'(3) \frac{1}{4K/3} = 1.184802 \frac{1}{4K}, \tag{28}$$

where we have assumed that

$$\frac{1}{K/3} \sum_{t=2,5,\dots} (x_t - \bar{x})^2 \approx \frac{1}{K} \sum_{t=1}^K (x_t - \bar{x})^2 \approx 4. \tag{29}$$

The little practical relevance of asymptotic results such as (20) can be seen when the asymptotic values are confronted with the actual values obtained by simulations. In the simplest case of Gaussian white noise, we do not have to safeguard against short-range dependence and can therefore choose a value of α slightly below $4/5$. Choosing $\alpha = 0.7$ and $K \approx n^\alpha$, we obtain 0.00857 (27) and 0.00617 (28) vs. 0.01148 and 0.00885 (simulated) for $n = 250$ and $K = 48$, 0.00326 and 0.00235 vs. 0.00381 and 0.00282 for $n = 1000$ and $K = 126$, 0.00065 and 0.00047 vs. 0.00068 and 0.00050 for $n = 10,000$ and $K = 630$, and 0.00021 and 0.00015 vs. 0.00021 and 0.00015 for $n = 50,000$ and $K = 1947$. Obviously, huge sample sizes are required for good agreement. In the case of a nontrivial ARFIMA process, this problem will become even more serious because a smaller value of α must be chosen.

More sophisticated further developments of the estimator \hat{d}_{GPH} are obtained by using more than three periodogram ordinates, allowing for overlaps, and introducing weights, or, equivalently, by using a lag-window estimator of the form:

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{s=-m}^m w(s/m) \hat{\gamma}(s) e^{-i\omega_j s}, \quad j = 1, \dots, K, \tag{30}$$

where $\hat{\gamma}(s)$ denotes the sample autocovariance at lag s and the lag window w satisfies $w(0) = 1$, $|w(s)| \leq 1$, and $w(-s) = w(s)$ (see [Hassler 1993](#); [Peiris and Court 1993](#); [Reisen 1994](#)). A disadvantage of these estimation procedures is that they require the specification of a second tuning parameter, namely the length of the weighted averages in the former case and $m \leq n - 1$ in the latter case, in addition to K . Of course, suitable weights and a suitable lag window, respectively, must be chosen too. Carrying out an extensive simulation study to compare various frequency-domain estimators for d , [Reschenhofer et al. \(2020\)](#) found that too strong smoothing, e.g., caused by choosing a too small value for m , entails an extremely large bias. [Hunt et al. \(2003\)](#) derived an approximation for the bias and observed generally a good agreement between their approximation and the corresponding value obtained by simulations when $d > 0$. However, the practical relevance of this approximation is limited because of its dependence on characteristics of the data generating process, which are unknown in practice.

2.3. Using Subsamples

A simple method of smoothing without introducing a bias is to average estimates obtained from different subsamples. Assume, for example, that the final estimate \hat{d} is obtained by averaging over N preliminary estimates $\hat{d}_1, \dots, \hat{d}_N$ obtained from independent subsamples $y_{11}, \dots, y_{n1}, \dots, y_{1N}, \dots, y_{nN}$; then, the variance of \hat{d} vanishes as N increases while the bias remains unchanged. Of course, artificially splitting a long, homogeneous time series into non-overlapping subseries does not necessarily have a positive effect. For illustration, consider the simplest case where the time series y_1, \dots, y_n is split into two disjoint subseries $y_1, \dots, y_{n/2}$ and $y_{n/2+1}, \dots, y_n$ of equal length. To allow a fair comparison, the frequency range $(0, \omega_K]$, is kept constant, which implies that in the case of the two subseries the number of used Fourier frequencies is $K/2$. Under the simplistic and mostly unrealistic assumption that the two subseries are independent, the (approximate) variance of the mean of the two GPH estimators based on the two subseries is given by:

$$\frac{1}{4} \left(\frac{\pi^2}{6} \frac{1}{4K/2} + \frac{\pi^2}{6} \frac{1}{4K/2} \right) = \frac{\pi^2}{6} \frac{1}{4K} \tag{31}$$

which is, therefore, of the same size as that of the original estimator, which is based on the whole time series. However, there is still room for improvement. A reduction in the variance may be achieved by allowing for overlaps between the subseries, e.g., with a rolling estimation window or a combination of different partitions.

At first glance, the idea of improving an OLS estimator by averaging the OLS estimators obtained from the whole sample and the first and second halves, respectively, seems to be at odds with the Gauß-Markov theorem because the combined estimator is still linear. However, the crucial point here is that only the observations are partitioned and not the log periodogram, which is used as dependent variable in the regression and is obtained from the observations through nonlinear transformations. For illustration, consider an estimator of the form:

$$\tilde{d}_2 = (1 - 2\lambda)\hat{d}_1 + \lambda\hat{d}_{21} + \lambda\hat{d}_{22}, \tag{32}$$

where $\hat{d}_1, \hat{d}_{21}, \hat{d}_{22}$ are the OLS estimators for d based on the log periodograms L^1, L^{21}, L^{22} of the whole sample and the first and second halves, respectively. In the special case of Gaussian white noise with variance 2π , the constant c in the regression (3) vanishes, and we may, therefore, use the simpler estimators:

$$\check{d}_1 = \frac{\sum_{j=1}^K \check{x}_j L_j^1}{\sum_{j=1}^K \check{x}_j^2} \approx \frac{1}{4K} \sum_{j=1}^K \check{x}_j L_j^1, \tag{33}$$

and:

$$\check{d}_{2s} = \frac{\sum_{j=1}^{K/2} \check{x}_{2j} L_j^{2s}}{\sum_{j=1}^{K/2} \check{x}_{2j}^2} \approx \frac{1}{2K} \sum_{j=1}^{K/2} \check{x}_{2j} L_j^{2s}, \quad s = 1, 2, \tag{34}$$

where $\check{x}_j = x_j - \bar{x}$. For the variances of the simplistic estimators \check{d}_1 and:

$$\check{d}_2 = (1 - 2\lambda)\check{d}_1 + \lambda\check{d}_{21} + \lambda\check{d}_{22}, \tag{35}$$

we obtain approximately:

$$\text{var}(\check{d}_1) \approx \left(\frac{1}{4K}\right)^2 \sum_{j=1}^K \check{x}_j^2 \frac{\pi^2}{6} \approx \frac{\pi^2}{24K} \tag{36}$$

and:

$$\begin{aligned} \text{var}(\check{d}_2) &\approx \frac{\pi^2}{24K} ((1 - 2\lambda)^2 + 4\lambda^2) + 4\lambda(1 - 2\lambda)\text{cov}(\check{d}_1, \check{d}_{21}) \\ &\approx \frac{\pi^2}{24K} ((1 - 2\lambda)^2 + 4\lambda^2 + 4\lambda(1 - 2\lambda)(\rho_0 + \rho_1)) \\ &\approx 0.69 \frac{\pi^2}{24K}, \text{ if } \lambda = \frac{1}{4}, \end{aligned} \tag{37}$$

respectively, where we have used that $\text{cov}(\check{d}_1, \check{d}_{21}) = \text{cov}(\check{d}_1, \check{d}_{22})$ and $\text{cov}(\check{d}_{21}, \check{d}_{22}) = 0$ as well as the rough approximations:

$$\sum_{j=1}^{\frac{K}{2}} \check{x}_{2j}^2 \approx \sum_{j=1}^{\frac{K}{2}} \check{x}_{2j}\check{x}_{2j-1} \approx \sum_{j=1}^{\frac{K}{2}-1} \check{x}_{2j}\check{x}_{2j+1} \approx 2K, \tag{38}$$

$$\text{cor}(L_j^1, L_k^{2s}) \approx \begin{cases} \rho_0 = 0.35, & \text{if } 2k = j, \\ \rho_1 = 0.13, & \text{if } |2k - j| = 1, \\ 0, & \text{else} \end{cases} \tag{39}$$

(see Table 1), and:

$$\begin{aligned}
 cov(\check{d}_1, \check{d}_{21}) &\approx \frac{1}{8K^2} cov\left(\sum_{j=1}^{\frac{K}{2}} x_{2j} L_{2j}^1 + \sum_{j=1}^{\frac{K}{2}} x_{2j-1} L_{2j-1}^1, \sum_{k=1}^{\frac{K}{2}} x_{2k} L_k^{21}\right) \\
 &\approx \frac{1}{8K^2} \left(\sum_{j=1}^{\frac{K}{2}} \sum_{k=1}^{\frac{K}{2}} x_{2j} x_{2k} cov(L_{2j}^1, L_{2k}^{21}) + \sum_{j=1}^{\frac{K}{2}} \sum_{k=1}^{\frac{K}{2}} x_{2j-1} x_{2k} cov(L_{2j-1}^1, L_{2k}^{21}) \right) \\
 &\approx \frac{1}{8K^2} \left(\rho_0 \frac{\pi^2}{6} \sum_{j=1}^{\frac{K}{2}} x_{2j}^2 + \rho_1 \frac{\pi^2}{6} \sum_{j=1}^{\frac{K}{2}} x_{2j} x_{2j-1} \right) \\
 &\approx \frac{\pi^2}{24K} (\rho_0 + \rho_1)
 \end{aligned}
 \tag{40}$$

For a further reduction of the variance, we may consider more general estimators of the form:

$$\tilde{d}_k = \frac{1}{k} \left(\hat{d}_1 + \sum_{j=2}^k \frac{1}{j} (\hat{d}_{j1} + \dots + \hat{d}_{jj}) \right),
 \tag{41}$$

which are based on k partitions. The next section examines whether this possible reduction actually materializes and whether it is accompanied by an increase in the bias. All computations are carried out with the free statistical software R (R Core Team 2018).

Table 1. Sample correlations between L_j , $j = 1, \dots, 20$, and L_k^1 , $k = 1, \dots, 10$, obtained from 10,000,000 realizations of Gaussian white noise ($n = 400$).

	1	2	3	4	5	6	7	8	9	10
1	0.1475	0.0186	0.0072	0.0044	0.0027	0.0014	0.0013	0.001	0.0008	0.0005
2	0.3541	0.0002	-0.0001	-0.0004	0	0.0003	-0.0003	0.0002	0.0002	0.0005
3	0.1364	0.133	0.0154	0.006	0.0032	0.0025	0.0009	0.001	0.0007	0.0003
4	-0.0001	0.3541	-0.0001	-0.0002	0.0002	0.0008	-0.0005	-0.0002	-0.0004	-0.0003
5	0.0164	0.1316	0.1307	0.0144	0.005	0.0027	0.0019	0.0016	0.0008	0.0008
6	-0.0001	-0.0003	0.354	0.0002	0.0002	-0.0004	0.0004	0.0001	-0.0005	0.0005
7	0.007	0.0147	0.1311	0.1308	0.014	0.0043	0.0025	0.0021	0.0013	0.0011
8	0	0.0001	0.0004	0.3541	0.0004	0.0001	-0.0001	-0.0002	-0.0001	-0.0002
9	0.0035	0.0054	0.0143	0.1302	0.1302	0.0139	0.0051	0.003	0.0016	0.0009
10	-0.0003	0	-0.0001	0.0004	0.3539	-0.0003	0.0003	0.0001	-0.0005	0.0003
11	0.0023	0.0033	0.0047	0.0138	0.1301	0.13	0.0133	0.0054	0.0025	0.0014
12	-0.0004	-0.0001	-0.0004	-0.0001	0.0003	0.3542	0.0001	-0.0001	0.0002	0
13	0.0013	0.002	0.0032	0.0053	0.0137	0.1305	0.1309	0.0147	0.004	0.003
14	-0.0004	0.0001	0.0003	0.0004	0.0008	0.0002	0.3544	-0.0002	0.0005	-0.0002
15	0.0011	0.0016	0.002	0.0025	0.0059	0.014	0.1304	0.1297	0.0141	0.0055
16	-0.0006	0.0001	-0.0004	0	0.0002	-0.0001	-0.0001	0.354	0.0002	0.0002
17	0.0011	0.0009	0.0009	0.0021	0.0025	0.0049	0.0138	0.1305	0.1304	0.0137
18	0.0003	-0.0002	0	-0.0001	-0.0006	-0.0004	-0.0002	-0.0004	0.3541	-0.0001
19	0.0008	0.0005	0.0011	0.0015	0.0019	0.0026	0.0046	0.0138	0.1306	0.1302
20	-0.0001	0.0005	0.0001	0.0002	0.0008	0.0001	0.0007	-0.0003	-0.0005	0.3541

3. Simulations

In this section, we compare the new estimator \tilde{d}_k (41) for $k = 2, 3, 5, 10$ with Geweke and Porter-Hudak’s (1983) estimator \hat{d}_{GPH} , which is based on the log periodogram regression (5), and the estimators \hat{d}_{sm} and \hat{d}_{smP}^β , which are obtained by replacing the periodogram ordinates in (5) by simple moving averages of neighboring periodogram ordinates and lag-window estimates of the form (30) with truncation lags $m = \lceil n^\beta \rceil$, $\beta = 0.5, 0.7, 0.9, 1$, respectively. In the latter case, the Parzen window is used, which is given by:

$$w(z) = \begin{cases} 1 - 6z^2 + 6|z|^3, & |z| < \frac{1}{2}, \\ 2(1 - |z|)^3, & \frac{1}{2} \leq |z| \leq 1. \end{cases}
 \tag{42}$$

With a view to the later application of the estimators to 1-min intraday returns in Section 4, the sample size $n = 390$ is chosen for our simulation study because there are 390 min in a regular

trading session for U.S. stocks, which starts at 9:30 a.m. and ends at 4:00 p.m. The number K of Fourier frequencies included in the log periodogram regression is defined by setting $K = 20 \approx [n^\alpha]$, $\alpha = 0.5$. For $k = 2$, the first $K/k = 10$ Fourier frequencies of the two disjoint subseries of length $n/k = 195$ are given by $\omega_2, \omega_4, \dots, \omega_K$, and for $k = 10$, the first $K/k = 2$ Fourier frequencies of the 10 disjoint subseries of length $n/k = 39$ are given by ω_{10}, ω_K . Clearly, we cannot go beyond $k = 10$ because at least two frequencies are required to carry out the log periodogram regression. Additionally, using frequencies outside the interval $(0, \omega_K]$ is not an option because this would amount to an unfair advantage, particularly when there are no short-term dependencies which have to be taken into account.

With the help of the R-package ‘fracdiff’, 10,000 realizations of length $n = 390$ of ARFIMA(1, d ,0) processes with standard normal innovations and parameter values $d = -0.25, -0.1, 0, 0.1, 0.25$ and $\phi_1 = -0.25, -0.1, 0, 0.1, 0.25$, respectively, are generated using a burn-in period of 10,000. For each realization, the estimators $\hat{d}_{GPH}, \hat{d}_{sm}, \hat{d}_{smP}^\beta, \beta = 0.5, 0.7, 0.9, 1, \tilde{d}_k, k = 2, 3, 5, 10$, are employed for the estimation of the memory parameter d . The competing estimators are compared with respect to bias (Table 2), variance (Table 3), and RMSE (Table 4). Table 3 shows that \tilde{d}_2 has indeed a smaller variance than $\hat{d}_{GPH} = \tilde{d}_1$. The variance keeps decreasing as the number of partitions increases from two to 10. Table 2 shows that this improvement does in general not come at the cost of a greater bias. In contrast, the reduction in the variance achieved in the case of the estimator \hat{d}_{smP}^β by increasing the degree of smoothing from $\beta = 0.9$ to $\beta = 0.5$ is for $d \neq 0$ accompanied by a dramatic increase in the bias. Overall, in terms of the RSME, the best results are obtained with $\hat{d}_{smP}^{0.5}$ for small values of d and with $\hat{d}_{smP}^{0.7}$ for larger value of d . However, this is only relevant in the standard case where only a single time series is available. When a large number of time series are examined simultaneously (as in the empirical study of Section 4), the bias is the decisive factor and the new estimators \tilde{d}_k are therefore more appropriate than the conventional estimators \hat{d}_{smP}^β .

Since values of β such as 0.5, 0.7, or 0.9 are usually chosen to minimize the MSE for a single sample, we may suspect that the estimator \hat{d}_{smP}^β becomes more competitive in the case of multiple samples when the averaging is taken into account. This can be done by further reducing the degree of smoothing. Unfortunately, there is a limit to what can be achieved by increasing the value of β . Table 2 shows that large biases are still obtained with the maximum possible value of β , i.e., $\beta = 1$. This is due to the fact that global smoothing inevitably causes local distortions and cutting off higher-order sample autocovariances is not the only source of smoothing. Downweighting the sample autocovariances with the Parzen window also has a strong smoothing effect, even when all sample autocovariances are used.

Table 2. Bias of the estimators \hat{d}_{GPH} (log periodogram regression), \hat{d}_{sm} (simple smoothing), \hat{d}_{smP}^β , $\beta = 1, 0.9, 0.7, 0.5$ (smoothing with Parzen window and truncation lag $m = [n^\beta]$), and $\tilde{d}_k, k = 2, 3, 5, 10$ (k partitions) obtained from 10,000 realizations (length: $n = 390$, number of used Fourier frequencies: $K = 20$) of Gaussian ARFIMA(1, d ,0) processes with $d = -0.25, -0.1, 0, 0.1, 0.25$ and $\phi_1 = -0.25, -0.1, 0, 0.1, 0.25$.

d	ϕ_1	\hat{d}_{GPH}	\hat{d}_{sm}	\hat{d}_{smP}^1	$\hat{d}_{smP}^{0.9}$	$\hat{d}_{smP}^{0.7}$	$\hat{d}_{smP}^{0.5}$	\tilde{d}_2	\tilde{d}_3	\tilde{d}_5	\tilde{d}_{10}
-0.25	-0.25	0.0074	-0.0001	-0.0073	-0.0099	0.0345	0.1609	0.0087	0.0084	0.0098	0.0107
	-0.1	0.0050	0.0002	-0.0083	-0.0107	0.0345	0.1625	0.0080	0.0084	0.0087	0.0092
	0	0.0042	-0.0031	-0.0098	-0.0124	0.0337	0.1641	0.0065	0.0065	0.0076	0.0086
	0.1	0.0097	0.0036	-0.0049	-0.0073	0.0380	0.1664	0.0126	0.0120	0.0128	0.0140
	0.25	0.0151	0.0110	0.0006	-0.002	0.0436	0.1717	0.0165	0.0179	0.0201	0.0216
-0.1	-0.25	0.0002	-0.0029	-0.0211	-0.0280	-0.008	0.0570	0.0008	0.0016	0.0006	0.0002
	-0.1	0.0015	-0.0028	-0.0212	-0.0286	-0.0085	0.0578	-0.0001	0.0005	0.0001	-0.0001
	0	0.0039	0.0017	-0.0184	-0.0251	-0.0053	0.0601	0.0038	0.0052	0.0060	0.0057
	0.1	0.0014	0.0007	-0.0197	-0.0263	-0.0056	0.0612	0.0024	0.0028	0.0039	0.0037
	0.25	0.0055	0.0059	-0.0148	-0.0215	-0.0003	0.0666	0.0086	0.0099	0.0093	0.0101

Table 2. Cont.

d	ϕ_1	\hat{d}_{GPH}	\hat{d}_{sm}	\hat{d}_{smP}^1	$\hat{d}_{smP}^{0.9}$	$\hat{d}_{smP}^{0.7}$	$\hat{d}_{smP}^{0.5}$	\tilde{d}_2	\tilde{d}_3	\tilde{d}_5	\tilde{d}_{10}
0	-0.25	-0.0043	-0.0035	-0.0282	-0.0376	-0.0321	-0.0107	-0.0038	-0.0039	-0.0048	-0.0049
	-0.1	-0.0011	0.0006	-0.0258	-0.0353	-0.0299	-0.0096	-0.0004	-0.0007	-0.0004	-0.0010
	0	-0.0011	-0.0001	-0.0265	-0.0361	-0.0305	-0.0087	-0.0016	-0.0004	-0.0006	-0.0006
	0.1	-0.0001	0.0009	-0.0235	-0.0333	-0.0278	-0.0063	0.0016	0.0025	0.0019	0.0025
	0.25	0.0040	0.0064	-0.0214	-0.0309	-0.0250	-0.0022	0.0033	0.0060	0.0053	0.0073
0.1	-0.25	0.0009	0.0057	-0.0274	-0.039	-0.0475	-0.0762	0.0009	-0.0003	0.0008	-0.0001
	-0.1	0.0016	0.0056	-0.0277	-0.0396	-0.0478	-0.0754	-0.0003	0.0002	-0.0007	-0.0006
	0	-0.0005	0.0043	-0.0277	-0.0396	-0.0479	-0.0745	-0.0012	-0.0012	-0.0012	-0.0010
	0.1	0.0029	0.0059	-0.0250	-0.0374	-0.0458	-0.0727	0.0020	0.0028	0.0038	0.0034
	0.25	0.0097	0.0149	-0.0186	-0.0305	-0.0392	-0.0685	0.0088	0.0096	0.0114	0.0115
0.25	-0.25	0.0006	0.0102	-0.0314	-0.0451	-0.0690	-0.1748	0.0021	0.0018	0.0009	0.0006
	-0.1	0.0016	0.0112	-0.0314	-0.0453	-0.0689	-0.1744	0.0006	0.0011	0.0014	0.0010
	0	0.0044	0.0140	-0.0281	-0.0420	-0.0656	-0.1730	0.0032	0.0037	0.0040	0.0039
	0.1	0.0049	0.0162	-0.0269	-0.0408	-0.0649	-0.1718	0.0049	0.0065	0.0061	0.0060
	0.25	0.0079	0.0229	-0.0228	-0.0364	-0.0600	-0.1682	0.0105	0.0120	0.0130	0.0137

Table 3. Variance of the estimators \hat{d}_{GPH} (log periodogram regression), \hat{d}_{sm} (simple smoothing), \hat{d}_{smP}^β , $\beta = 1, 0.9, 0.7, 0.5$ (smoothing with Parzen window and truncation lag $m = \lfloor n^\beta \rfloor$), and \tilde{d}_k , $k = 2, 3, 5, 10$ (k partitions) obtained from 10,000 realizations (length: $n = 390$, number of used Fourier frequencies: $K = 20$) of Gaussian ARFIMA(1, d ,0) processes with $d = -0.25, -0.1, 0, 0.1, 0.25$ and $\phi_1 = -0.25, -0.1, 0, 0.1, 0.25$.

d	ϕ_1	\hat{d}_{GPH}	\hat{d}_{sm}	\hat{d}_{smP}^1	$\hat{d}_{smP}^{0.9}$	$\hat{d}_{smP}^{0.7}$	$\hat{d}_{smP}^{0.5}$	\tilde{d}_2	\tilde{d}_3	\tilde{d}_5	\tilde{d}_{10}
-0.25	-0.25	0.0330	0.0328	0.0201	0.018	0.0106	0.0011	0.0287	0.0259	0.0254	0.0238
	-0.1	0.0334	0.0339	0.0207	0.0185	0.0110	0.0012	0.0297	0.0266	0.0261	0.0245
	0	0.0342	0.0337	0.0209	0.0185	0.0108	0.0011	0.0296	0.0267	0.0262	0.0248
	0.1	0.0327	0.0330	0.0202	0.0180	0.0107	0.0011	0.0287	0.0262	0.0257	0.0240
	0.25	0.0323	0.0325	0.0199	0.0178	0.0106	0.0011	0.0287	0.0260	0.0258	0.0242
-0.1	-0.25	0.0333	0.0327	0.0211	0.0187	0.0114	0.0011	0.0295	0.0268	0.0264	0.0250
	-0.1	0.0332	0.0317	0.0209	0.0186	0.0114	0.0011	0.0291	0.0264	0.0260	0.0250
	0	0.0334	0.0330	0.0212	0.0189	0.0115	0.0012	0.0298	0.0271	0.0267	0.0251
	0.1	0.0330	0.0315	0.0208	0.0185	0.0112	0.0011	0.0289	0.0262	0.0258	0.0246
	0.25	0.0328	0.0320	0.0209	0.0185	0.0112	0.0011	0.0291	0.0266	0.0263	0.0248
0	-0.25	0.0333	0.0322	0.0212	0.0191	0.0120	0.0012	0.0296	0.0268	0.0263	0.0250
	-0.1	0.0328	0.0320	0.0212	0.0191	0.0120	0.0012	0.0293	0.0268	0.0261	0.0252
	0	0.0335	0.0319	0.0214	0.0192	0.0119	0.0012	0.0297	0.0271	0.0266	0.0254
	0.1	0.0338	0.0323	0.0217	0.0195	0.0122	0.0012	0.0299	0.0271	0.0270	0.0260
	0.25	0.0332	0.0324	0.0213	0.0192	0.0120	0.0012	0.0300	0.0273	0.0269	0.0255
0.1	-0.25	0.0332	0.0327	0.0218	0.0198	0.0130	0.0012	0.0299	0.0274	0.0271	0.0260
	-0.1	0.0327	0.0321	0.0218	0.0199	0.0130	0.0012	0.0294	0.0269	0.0262	0.0252
	0	0.0328	0.0317	0.0214	0.0194	0.0127	0.0012	0.0293	0.0264	0.0263	0.0250
	0.1	0.0331	0.0321	0.0215	0.0195	0.0129	0.0012	0.0295	0.0269	0.0267	0.0256
	0.25	0.0326	0.0321	0.0217	0.0197	0.0130	0.0012	0.0293	0.0268	0.0263	0.0254
0.25	-0.25	0.0333	0.0315	0.0220	0.0202	0.0145	0.0013	0.0300	0.0271	0.0271	0.0260
	-0.1	0.0327	0.0323	0.0222	0.0205	0.0148	0.0013	0.0302	0.0278	0.0275	0.0265
	0	0.0328	0.0312	0.0219	0.0202	0.0146	0.0012	0.0297	0.0268	0.0264	0.0255
	0.1	0.0333	0.0325	0.0226	0.0207	0.0147	0.0013	0.0301	0.0274	0.0274	0.0262
	0.25	0.0339	0.0319	0.0226	0.0208	0.0150	0.0012	0.0302	0.0275	0.0272	0.0261

Table 4. RMSE of the estimators \hat{d}_{GPH} (log periodogram regression), \hat{d}_{sm} (simple smoothing), \hat{d}_{smP}^β , $\beta = 1, 0.9, 0.7, 0.5$ (smoothing with Parzen window and truncation lag $m = \lceil n^\beta \rceil$), and \tilde{d}_k , $k = 2, 3, 5, 10$ (k partitions) obtained from 10,000 realizations (length: $n = 390$, number of used Fourier frequencies: $K = 20$) of Gaussian ARFIMA(1, d ,0) processes with $d = -0.25, -0.1, 0, 0.1, 0.25$ and $\phi_1 = -0.25, -0.1, 0, 0.1, 0.25$.

d	ϕ_1	\hat{d}_{GPH}	\hat{d}_{sm}	\hat{d}_{smP}^1	$\hat{d}_{smP}^{0.9}$	$\hat{d}_{smP}^{0.7}$	$\hat{d}_{smP}^{0.5}$	\tilde{d}_2	\tilde{d}_3	\tilde{d}_5	\tilde{d}_{10}
-0.25	-0.25	0.1818	0.1811	0.1421	0.1344	0.1084	0.1643	0.1697	0.1612	0.1595	0.1545
	-0.1	0.1827	0.1840	0.1442	0.1365	0.1104	0.1661	0.1724	0.1634	0.1618	0.1567
	0	0.1851	0.1837	0.1449	0.1368	0.1092	0.1674	0.1721	0.1635	0.1621	0.1578
	0.1	0.1812	0.1816	0.1423	0.1343	0.1103	0.1698	0.1700	0.1623	0.1607	0.1555
	0.25	0.1803	0.1807	0.1412	0.1335	0.1119	0.1751	0.1701	0.1621	0.1619	0.1571
-0.1	-0.25	0.1825	0.1808	0.1466	0.1396	0.1070	0.0663	0.1717	0.1636	0.1624	0.1581
	-0.1	0.1823	0.1782	0.1460	0.1394	0.1072	0.0669	0.1705	0.1625	0.1611	0.1580
	0	0.1829	0.1816	0.1467	0.1398	0.1072	0.0691	0.1727	0.1647	0.1636	0.1585
	0.1	0.1817	0.1775	0.1454	0.1386	0.1061	0.0698	0.1699	0.1618	0.1607	0.1569
	0.25	0.1811	0.1789	0.1451	0.1378	0.1059	0.0745	0.1707	0.1634	0.1625	0.1578
0	-0.25	0.1826	0.1796	0.1481	0.1431	0.1142	0.0360	0.1721	0.1639	0.1624	0.1583
	-0.1	0.1812	0.1790	0.1479	0.1426	0.1137	0.0359	0.1713	0.1638	0.1615	0.1588
	0	0.1831	0.1785	0.1486	0.1433	0.1132	0.0351	0.1723	0.1646	0.1630	0.1593
	0.1	0.1837	0.1796	0.1491	0.1435	0.1139	0.0351	0.1729	0.1647	0.1645	0.1611
	0.25	0.1824	0.1801	0.1475	0.1418	0.1123	0.0345	0.1731	0.1653	0.1640	0.1599
0.1	-0.25	0.1822	0.1810	0.1502	0.146	0.1237	0.0837	0.1728	0.1657	0.1646	0.1612
	-0.1	0.181	0.1793	0.1502	0.1464	0.1237	0.0831	0.1715	0.1639	0.1617	0.1588
	0	0.181	0.1781	0.1490	0.1448	0.1226	0.0820	0.1711	0.1624	0.1622	0.1582
	0.1	0.1819	0.1792	0.1489	0.1446	0.1223	0.0805	0.1717	0.1641	0.1633	0.1599
	0.25	0.1808	0.1799	0.1485	0.1437	0.1206	0.0768	0.1713	0.1640	0.1626	0.1596
0.25	-0.25	0.1824	0.1778	0.1517	0.1493	0.1390	0.1784	0.1733	0.1648	0.1647	0.1612
	-0.1	0.1809	0.1800	0.1522	0.1502	0.1398	0.1780	0.1738	0.1666	0.1657	0.1629
	0	0.1810	0.1772	0.1505	0.1483	0.1375	0.1765	0.1723	0.1636	0.1626	0.1598
	0.1	0.1824	0.1809	0.1526	0.1495	0.1377	0.1754	0.1737	0.1657	0.1657	0.1621
	0.25	0.1842	0.1799	0.1522	0.1487	0.1363	0.1718	0.1740	0.1663	0.1654	0.1623

4. Empirical Results

In this section, we employ the estimators discussed in the previous sections for the search of possible long-range dependencies in intraday returns and absolute intraday returns. For this purpose, the limit order book data from 27 June 2007 to 30 April 2019 (2980 trading days) of the iShares Core S&P 500 ETF (IVV) are downloaded from Lobster (<https://lobsterdata.com>). In the process of data cleaning, 27 early-closure days (the day before Independence Day, the day after Thanksgiving, and Christmas Eve) are removed as well as 9 January 2019 because of a large number of missing values. For each of the remaining days, the first mid-quotes (midpoints of the best bid and ask quotes) in each minute and the last mid-quote in the last minute are computed and subsequently used to obtain 1-min log returns. Finally, another three days are omitted because of extreme returns, namely 19 September 2008, 6 May 2010, and 24 August 2015, which leaves 2949 days for our analysis. Estimates are computed for each day, divided by the number of days, and plotted cumulatively; hence, the last values correspond to the averages of the estimates. The validity of these values is reinforced by the striking linearity of the curves. This linearity also implies that the possible long-range dependence is not changing over time; hence, there appears to be no such thing as fractal dynamics. Figure 1a suggests that d is close to zero in case of the 1-min log returns. The large negative values obtained with $\hat{d}_{smP}^{0.9}$ and $\hat{d}_{smP}^{0.7}$ as well as the comparatively inconspicuous values obtained with $\hat{d}_{smP}^{0.5}$ can be explained with the help of the results of our simulation study. According to Table 2, they are indicative for $d = 0$. In contrast, there is strong evidence of long-range dependence in the volatility (see Figure 1b). Most estimators suggest

that the memory parameter d is approximately in the range between 0.3 and 0.4. Only the estimator $\hat{d}_{smP}^{0.5}$, which is severely downward biased in case of positive d (see Table 2), favors a smaller value.

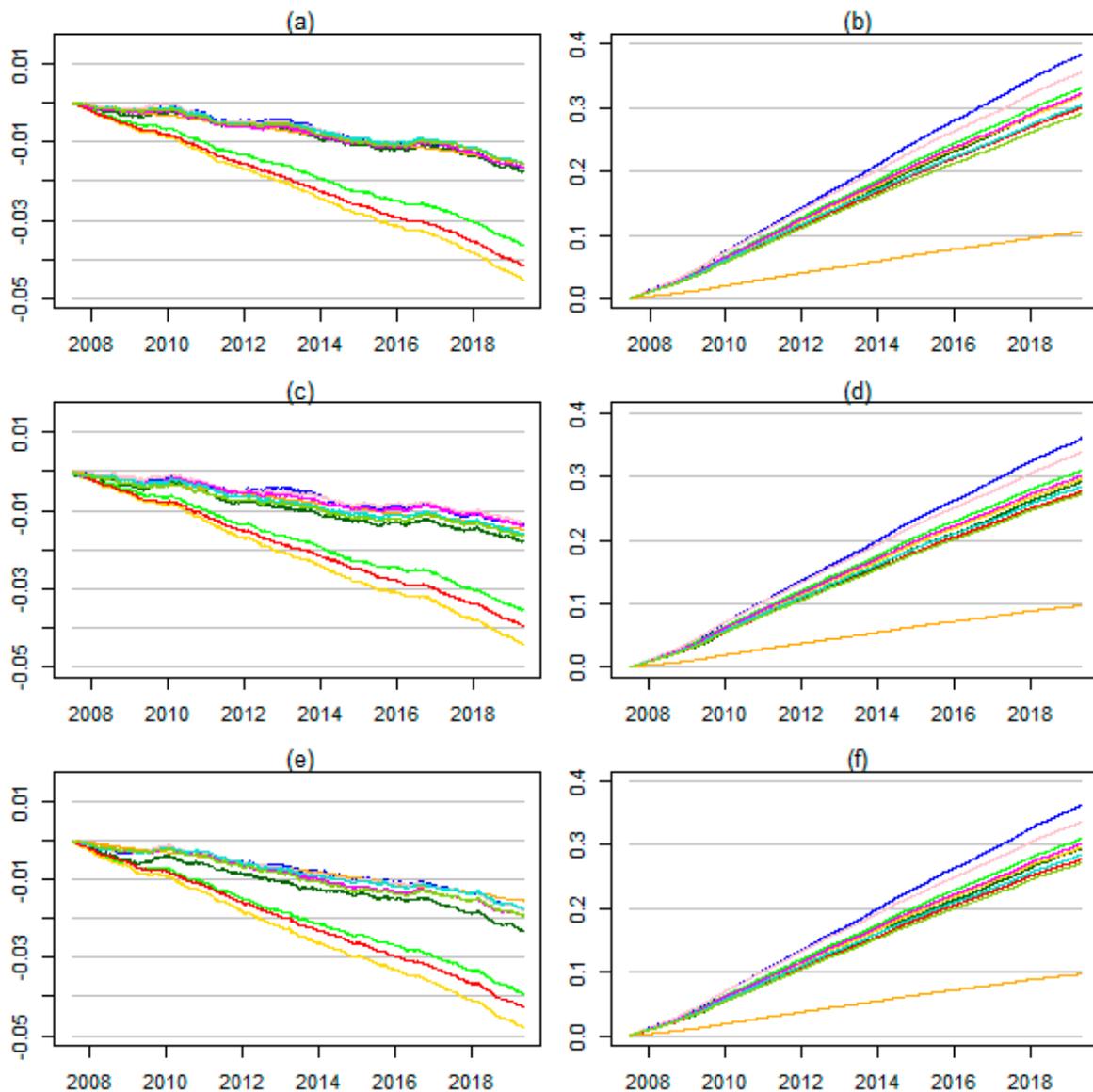


Figure 1. Cumulative plots of the estimates obtained by applying \hat{d}_{GPH} (blue), \hat{d}_{sm} (darkgreen), \hat{d}_{smP}^1 (green), $\hat{d}_{smP}^{0.9}$ (gold), $\hat{d}_{smP}^{0.7}$ (red), $\hat{d}_{smP}^{0.5}$ (orange), \hat{d}_2 (pink), \hat{d}_3 (magenta), \hat{d}_5 (turquoise), \hat{d}_{10} (yellowgreen) to the (a) 1-min intraday log returns $r_t(s)$, $s = 1, \dots, 390$, (b) absolute 1-min intraday log returns $|r_t(s)|$, (c) $r_t(s) - r_{t-1}(s)$, (d) $|r_t(s) - r_{t-1}(s)|$, (e) $r_t(s) - r_{t-5}(s)$, (f) $|r_t(s) - r_{t-5}(s)|$.

Visual significance of the differences between certain estimates can be ascertained just by observing the large differences between the slopes of the corresponding lines in Figure 1 and noting the striking stability of these lines over time. However, we still might want to augment our visual analysis with a formal statistical test. A simple way to accomplish that is to calculate the difference between two estimates separately for each trading day and compare the number of positive differences to the number of negative differences (sign test). Not surprisingly, the resulting p-values are infinitesimal. For example, even in the case of the two neighboring lines corresponding to \hat{d}_{GPH} and \hat{d}_2 in Figure 1b, the p-value is less than 2.2×10^{-16} . It is still less than 9.7×10^{-8} when we omit most of the trading days and use only Wednesdays in order to ensure approximate independence of the subsamples. Note that there are $4 \times 390 = 1560$ 1-min returns between the last 1-min return of some Wednesday

and the first 1-min return of the next Wednesday plus five overnight breaks and a whole weekend. Even for a relatively large value of the memory parameter such as $d = 0.3$, the autocorrelation of an ARFIMA(0, d ,0) process at lag $j = 1561$ is quite small, i.e.,

$$\rho(j) = \frac{\Gamma(j+d)\Gamma(1-d)}{\Gamma(j-d+1)\Gamma(d)} = \prod_{h=1}^j \frac{h-1+d}{h-d} \approx 0.023. \quad (43)$$

Finally, in order to check the robustness of our findings against daily and weekly periodic patterns, we repeat the graphical analyses with suitably transformed data. Replacing the 1-min log returns $r_t(s)$, $s = 1, \dots, 390$, by the daily differences $r_t(s) - r_{t-1}(s)$ and the weekly differences $r_t(s) - r_{t-5}(s)$, respectively, ensures that any daily or weekly periodic patterns are erased while long-range dependencies remain unaffected. Figure 1c,e are very similar to Figure 1a, which shows that the insights gained from Figure 1a are genuine. Analogously, comparing Figure 1d,f with Figure 1b, we see that the same is true for the absolute returns

5. Discussion

In this paper, we have introduced a new estimator for the memory parameter d , which is based on running a log periodogram regression repeatedly for different partitions of the data. In contrast to conventional smoothing methods, which manage to achieve a reduction in the variance at the expense of an increase in the bias, our approach does not systematically have a negative impact on the bias, which makes it particularly useful for applications where the bias is the decisive factor. For example, intraday returns are usually only available during trading hours and estimation must therefore be carried out separately for each trading day. When the individual estimates are eventually combined by averaging, the variance decreases as the sample size increases, but the bias does not change. The results of an extensive simulation study confirm the good performance of the new estimator. It outperforms all of its competitors when both bias and variance are taken into account, but the bias is weighted more heavily.

The importance of results obtained with the help of simulations is due to the fact that reliable inference on the memory parameter d is not possible under general conditions. Some asymptotic results can be obtained under very restrictive conditions though. Unfortunately, convergence is typically very slow (recall the discussion in Section 2.2). Indeed, Pötscher (2002) showed that many common estimation problems in statistics and econometrics, which include the estimation of d , are ill-posed in the sense that the minimax risk is bounded from below by a positive constant independent of n and does, therefore, not converge to zero as $n \rightarrow \infty$. In particular, he found that for any estimator \hat{d}_n for d based on a sample of size n from a Gaussian process with spectral density f :

$$\sup_{f \in \mathcal{F}} E|\hat{d}_n - d|^r \geq \frac{1}{2^r} > 0, \quad (44)$$

where $1 \leq r < \infty$ and \mathcal{F} is the set of all ARFIMA spectral densities ($p \geq 0$, $q \geq 0$), ARFI spectral densities ($p \geq 0$, $q = 0$), or FIMA spectral densities ($p = 0$, $q \geq 0$). Furthermore, he showed that for every $f_0 \in \mathcal{F}$, (44) holds also “locally,” when the supremum is taken over an arbitrarily small L_1 -neighborhood of f_0 . Finally, he established that confidence intervals for d coincide with the entire parameter space for d with high probability and are therefore uninformative. Nevertheless, it may be possible to formally derive the statistical properties of our new estimator for a rather narrow class of processes such as low order ARFI processes. However, this is left for future research. The current paper provides just a proof of concept.

In our empirical investigation of high-frequency data of an index ETF, we have applied the competing estimators to 1-min log returns and absolute 1-min log returns separately for each day. The results are quite stable over time and across estimation methods. The few deviations are due to conventional smoothing methods and can easily be explained by the size of their bias as shown

in Table 2. We may, therefore, safely conclude that significant long-range dependencies are present only in the intraday volatility but not in the intraday returns. These findings are genuine and not just due to daily or weekly periodic patterns because similar results are obtained when daily and weekly differences are investigated instead of the original intraday returns.

Author Contributions: Both authors contributed equally to the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors thank the academic editor and three anonymous reviewers for helpful comments and suggestions. Open Access Funding by the University of Vienna.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Adenstedt, Rolf K. 1974. On Large-Sample Estimation for the Mean of Stationary Random Sequence. *The Annals of Statistics* 2: 1095–107. [\[CrossRef\]](#)
- Andrews, Donald W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59: 817–58. [\[CrossRef\]](#)
- Auer, Benjamin R. 2016a. On the performance of simple trading rules derived from the fractal dynamics of gold and silver price fluctuations. *Finance Research Letters* 16: 255–67. [\[CrossRef\]](#)
- Auer, Benjamin R. 2016b. On time-varying predictability of emerging stock market returns. *Emerging Markets Review* 27: 1–13. [\[CrossRef\]](#)
- Barkoulas, John T., and Christopher F. Baum. 1996. Long-term dependence in stock returns. *Economic Letters* 53: 253–59. [\[CrossRef\]](#)
- Barkoulas, John T., Christopher F. Baum, and Nickolas Travlos. 2000. Long memory in the Greek stock market. *Applied Financial Economics* 10: 177–84. [\[CrossRef\]](#)
- Batten, Jonathan A., and Peter G. Szilagyi. 2007. Covered interest parity arbitrage and long-term dependence between the US dollar and the Yen. *Physica A* 376: 409–21. [\[CrossRef\]](#)
- Batten, Jonathan A., Craig A. Ellis, and Thomas A. Fethertson. 2008. Sample period selection and long-term dependence: New evidence from the Dow Jones Index. *Chaos, Solitons Fractals* 36: 1126–40. [\[CrossRef\]](#)
- Batten, Jonathan, Cetin Ciner, Brian M. Lucey, and Peter G. Szilagyi. 2013. The structure of gold and silver spread returns. *Quantitative Finance* 13: 561–70. [\[CrossRef\]](#)
- Cajueiro, Daniel O., and Benjamin M. Tabak. 2004. The Hurst exponent over time: Testing the assertion that emerging markets are becoming more efficient. *Physica A* 336: 521–37. [\[CrossRef\]](#)
- Carbone, Anna, Giuliano Castelli, and H. Eugene Stanley. 2004. Time-dependent Hurst exponent in financial time series. *Physica A* 344: 267–71. [\[CrossRef\]](#)
- Chen, Gemai, Bovas Abraham, and Shelton Peiris. 1994. Lag window estimation of the degree of differencing in fractionally integrated time series models. *Journal of Time Series Analysis* 15: 473–87. [\[CrossRef\]](#)
- Cheung, Yin-Wong, and Kon S. Lai. 1995. A search for long memory in international stock market returns. *Journal of International Money and Finance* 14: 597–615. [\[CrossRef\]](#)
- Crato, Nuno. 1994. Some international evidence regarding the stochastic behaviour of stock returns. *Applied Financial Economics* 4: 33–39. [\[CrossRef\]](#)
- Crato, Nuno, and Pedro J. F. de Lima. 1994. Long-range dependence in the conditional variance of stock returns. *Economics Letters* 45: 281–85. [\[CrossRef\]](#)
- Davis, Robert, and David Harte. 1987. Tests of the Hurst effect. *Biometrika* 74: 95–101. [\[CrossRef\]](#)
- Geweke, John, and Susan Porter-Hudak. 1983. The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4: 221–38. [\[CrossRef\]](#)
- Granger, Clive W. J., and Roselyne Joyeux. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1: 15–29. [\[CrossRef\]](#)
- Grau-Carles, Pilar. 2000. Empirical evidence of long-range correlations in stock returns. *Physica A* 287: 396–404. [\[CrossRef\]](#)
- Greene, Myron T., and Bruce D. Fielitz. 1977. Long term dependence in common stock returns. *Journal of Financial Economics* 4: 339–49. [\[CrossRef\]](#)

- Hassler, Uwe. 1993. Regression of spectral estimators with fractionally integrated time series. *Journal of Time Series Analysis* 14: 369–80. [\[CrossRef\]](#)
- Hauser, Michael A., and Erhard Reschenhofer. 1995. Estimation of the fractionally differencing parameter with the R/S method. *Computational Statistics & Data Analysis* 20: 569–79.
- Henry, Ólan T. 2002. Long memory in stock returns: Some international evidence. *Applied Financial Economics* 12: 725–29. [\[CrossRef\]](#)
- Hosking, Jonathan R. M. 1981. Fractional differencing. *Biometrika* 68: 165–76. [\[CrossRef\]](#)
- Hunt, R. L., M. Shelton Peiris, and N. C. Weber. 2003. The bias of lag window estimators of the fractional difference parameter. *Journal of Applied Mathematics and Computing* 12: 67–79. [\[CrossRef\]](#)
- Hurst, Harold E. 1951. Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* 116: 770–99.
- Hurvich, Clifford M., and Kaizo I. Beltrao. 1993. Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *Journal of Time Series Analysis* 14: 455–72. [\[CrossRef\]](#)
- Hurvich, Clifford M., Rohit Deo, and Julia Brodsky. 1998. The mean square error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis* 19: 19–46. [\[CrossRef\]](#)
- Künsch, Hans-Rudolf. 1986. Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability* 23: 1025–30. [\[CrossRef\]](#)
- Lo, Andrew. 1991. Long-term memory in stock market prices. *Econometrica* 59: 1279–313. [\[CrossRef\]](#)
- Lobato, Ignacio N., and N. E. Savin. 1998. Real and spurious long-memory properties of stock-market data. *Journal of Business & Economic Statistics* 16: 261–68.
- Mandelbrot, Benoît. 1971. When can price be arbitrated efficiently? A limit to the validity of the random walk and martingale models. *The Review of Economics and Statistics* 53: 225–36. [\[CrossRef\]](#)
- Mandelbrot, Benoît. 1972. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Economic and Social Measurement* 1: 259–90.
- Mandelbrot, Benoît. 1975. Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31: 271–85. [\[CrossRef\]](#)
- Mandelbrot, Benoît, and James Wallis. 1969. Computer experiments with fractional Gaussian noises. Parts 1, 2, 3. *Water Resources Research* 4: 909–18. [\[CrossRef\]](#)
- Mangat, Manveer K., and Erhard Reschenhofer. 2019. Testing for long-range dependence in financial time series. *Central European Journal of Economic Modelling and Econometrics* 11: 93–106.
- Newey, Whitney K., and Kenneth D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–8. [\[CrossRef\]](#)
- Peiris, M. Shelton, and J. R. Court. 1993. A note on the estimation of degree of differencing in long memory time series analysis. *Probability and Mathematical Statistics* 14: 223–29.
- Pötscher, Benedikt M. 2002. Lower risk bounds and properties of confidence sets for ill-posed estimation problems with applications to spectral density and persistence estimation, unit roots, and estimation of long memory parameters. *Econometrica* 70: 1035–65. [\[CrossRef\]](#)
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reisen, Valderio A. 1994. Estimation of the fractional difference parameter in the ARIMA(p,d,q) model using the smoothed periodogram. *Journal of Time Series Analysis* 15: 335–50. [\[CrossRef\]](#)
- Reisen, Valderio A., Bovas Abraham, and Silvia Lopes. 2001. Estimation of parameters in ARFIMA processes: A simulation study. *Communications in Statistics: Simulation and Computation* 30: 787–803. [\[CrossRef\]](#)
- Reschenhofer, Erhard, and Manveer K. Mangat. 2020. Detecting long-range dependence with truncated ratios of periodogram ordinates. *Communications in Statistics—Theory and Methods*. [\[CrossRef\]](#)
- Reschenhofer, Erhard, Manveer K. Mangat, and Thomas Stark. 2020. Improved estimation of the memory parameter. *Theoretical Economics Letters* 10: 47–68. [\[CrossRef\]](#)

- Robinson, Peter M. 1995. Log-periodogram regression of time series with long range dependence. *Annals of Statistics* 23: 1048–72. [[CrossRef](#)]
- Souza, Sergio, Benjamin M. Tabak, and Daniel O. Cajueiro. 2008. Long-range dependence in exchange rates: The case of the European monetary system. *International Journal of Theoretical and Applied Finance* 11: 199–223. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).