

Busch, Marie; Sibbertsen, Philipp

Article

An overview of modified semiparametric memory estimation methods

Econometrics

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Busch, Marie; Sibbertsen, Philipp (2018) : An overview of modified semiparametric memory estimation methods, *Econometrics*, ISSN 2225-1146, MDPI, Basel, Vol. 6, Iss. 1, pp. 1-21,
<https://doi.org/10.3390/econometrics6010013>

This Version is available at:

<https://hdl.handle.net/10419/195450>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Article

An Overview of Modified Semiparametric Memory Estimation Methods

Marie Busch and Philipp Sibbertsen *

Institute of Statistics, Faculty of Economics and Management, Leibniz University Hannover, D-30167 Hannover, Germany; busch@statistik.uni-hannover.de

* Correspondence: sibbertsen@statistik.uni-hannover.de; Tel.: +49-511-762-3783; Fax: +49-511-762-3923

Received: 13 December 2017; Accepted: 22 February 2018; Published: 12 March 2018

Abstract: Several modified estimation methods of the memory parameter have been introduced in the past years. They aim to decrease the upward bias of the memory parameter in cases of low frequency contaminations or an additive noise component, especially in situations with a short-memory process being contaminated. In this paper, we provide an overview and compare the performance of nine semiparametric estimation methods. Among them are two standard methods, four modified approaches to account for low frequency contaminations and three procedures developed for perturbed fractional processes. We conduct an extensive Monte Carlo study for a variety of parameter constellations and several DGPs. Furthermore, an empirical application of the log-absolute return series of the S&P 500 shows that the estimation results combined with a long-memory test indicate a spurious long-memory process.

Keywords: spurious long memory; semiparametric estimation; low frequency contamination; perturbation; Monte Carlo simulation

JEL Classification: C13; C14; C22

1. Introduction

Over the last decades, the long memory literature focusses on an effect called spurious long memory. This effect is characterized by a contaminated short-memory process exhibiting the typical properties of a long-memory process, such as a hyperbolically decaying autocorrelation function and an unbounded spectral density at the origin. Several types of contamination exist and we follow the classification of [Hou and Perron \(2014\)](#). We distinguish between short-run and additive noise components which influence the bias and efficiency of the standard estimator in finite samples. Further, low frequency contaminations modeled by time trends and level shifts can cause inconsistent estimators and are a source of the aforementioned spurious long memory (cf. [Diebold and Inoue 2001](#); [Granger and Hyung 2004](#)). [Haldrup and Nielsen \(2007\)](#) present strongly biased standard estimators for contaminated processes in their simulation study and emphasize the need of modified estimators. Besides the theoretical effects of low frequency contaminations, [McCloskey and Perron \(2013\)](#) emphasize the relevance to macroeconomic and financial applications.

Motivated by these findings, several modified estimation methods of the memory parameter have been developed. In this survey paper, we concentrate on semiparametric estimation methods and do not consider the literature of fully parametric estimation techniques, such as [McCloskey \(2013\)](#) or [McCloskey and Hill \(2017\)](#). Further, we omit semiparametric Fourier- and wavelet-based estimation approaches and refer to a survey paper by [Faÿ et al. \(2009\)](#).

Generally, all modified estimation methods aim to decrease the upward bias of the memory parameter in cases of low frequency contaminations or an additive noise component, especially in situations with a short-memory process being contaminated. Several types of modifications have been

proposed in the literature and we compare the effects of a trimmed periodogram and an adjusted spectral density. The order of the periodogram differs from a typical long-memory process to a contaminated process. Therefore, by adding or replacing components of the original spectral density, a so-called pseudo-spectral density function is modeled (e.g., [Hou and Perron 2014](#)). Alternatively, trimming the lower frequencies of a periodogram removes the high peak of the periodogram that is caused by the contamination. A common drawback of the modified estimators is the bias-variance trade-off which means that a bias reduction is combined with an increased variance.

To our knowledge, this is the first survey paper on semiparametric estimation methods of the memory parameter considering several types of contaminated processes. Usually, authors of modified estimation methods compare the results to the original non-modified estimation method. However, a comparison between different modified estimation methods would be of great interest. Further, the simulation studies of the modified estimation procedures are often limited to a few data generating processes (DGPs) and to a small set of parameter constellations, except for [Hou and Perron \(2014\)](#). We contribute to this literature by providing an extensive Monte Carlo study for a variety of parameter constellations and several DGPs. Additionally, we review the performance of the estimation methods for perturbed fractional processes since the general idea is closely related to an additive noise component as it has been considered by [Hou and Perron \(2014\)](#).

The idea of the paper is to give an overview of different estimation methods of pure and contaminated short and long-memory processes and to compare them under several scenarios. We aim to provide some user-guidelines in the end or a recommendation on which method should be preferred or which one is performing most stable across the wide setup in the simulation study. In the following, we briefly introduce the estimation methods. Two widespread semiparametric estimation methods are considered for pure long-memory processes, the log-periodogram estimator (also known as GPH estimator) of [Geweke and Porter-Hudak \(1983\)](#) and the local Whittle estimation approach of [Kuensch \(1987\)](#). In this survey paper, we compare the modified estimators of [Smith \(2005\)](#), [Iacone \(2010\)](#), [McCloskey and Perron \(2013\)](#) and [Hou and Perron \(2014\)](#). [Smith \(2005\)](#) develops a biased corrected version of the log-periodogram estimator by adding an additional regressor to the ordinary least squares regression equation. The trimming technique of [Iacone \(2010\)](#) is based on the local Whittle estimator, whereas [McCloskey and Perron \(2013\)](#) use the GPH estimator. As stated by [Yamamoto and Perron \(2013\)](#), the high peaks in the periodogram at low frequencies induced by the level shifts or trends are robustly removed by the trimming procedure of [McCloskey and Perron \(2013\)](#) and [Iacone \(2010\)](#). Modifying the estimator of [Kuensch \(1987\)](#) by adding a term to the spectral density that considers the contaminated process results in the estimator of [Hou and Perron \(2014\)](#).

In addition to the modified estimation methods of the memory parameter, we consider three estimation methods for perturbed fractional processes. In contrast to a pure fractional process, the short-run dynamics and long-run behavior of a perturbed fractional process are no longer driven by the same innovation. As defined in [Frederiksen et al. \(2012\)](#), a perturbed fractional process can be expressed by a signal-plus-noise model with the signal process being a long-memory process, which is perturbed by an additive noise term. We focus on local Whittle based estimators for perturbed fractional processes and refer to [Sun and Phillips \(2003\)](#) or [Arteche \(2006\)](#) for contributions based on the log-periodogram estimator.

Extending this survey paper by the estimators of [Andrews and Sun \(2004\)](#), [Hurvich et al. \(2005\)](#) and [Frederiksen et al. \(2012\)](#) we are considering an additive noise component as well as short-run effects as another form of contamination. All three estimation methods modify the form of the spectral density. [Andrews and Sun \(2004\)](#) replace a constant term in the spectral density by a polynomial structure to model the short-run dynamics. [Hurvich et al. \(2005\)](#) extend the estimator of [Hurvich and Ray \(2003\)](#) that is adding an additional term for capturing the effect of the noise term to the low frequency behavior of the spectral density. The estimator of [Frederiksen et al. \(2012\)](#) models the short-run dynamics as well as the influence of an additive noise component by a polynomial and allows for serial dependence in the noise term, in contrast to [Frederiksen and Nielsen \(2008\)](#).

According to our Monte Carlo results, we conclude to use either the estimator of [Iacone \(2010\)](#) or [Hou and Perron \(2014\)](#) in cases of low frequency contaminations. Although the estimator of [McCloskey and Perron \(2013\)](#) yields comparable estimation results, it suffers from a higher variance. The performance of the estimator of [Smith \(2005\)](#) in cases of high short-run dynamics or perturbation is comparable to the standard perturbation methods, although the estimator has originally been developed for low frequency contamination. However, in the latter situation, the estimator is no alternative to its competitors. Furthermore, the estimator of [Hurvich et al. \(2005\)](#) and [Frederiksen et al. \(2012\)](#) yield almost identical findings, which are in most situations better than the results of [Andrews and Sun \(2004\)](#). Based on the Monte Carlo results, we recommend applying the low frequency estimators when it is unclear whether a low frequency contamination or a perturbed fractional process is present.

Three out of the seven modified semiparametric estimators that we consider have been applied to empirical data in the original paper, such as the Dow Jones Industrial Average or the S&P 500 series. [Smith \(2005\)](#) and [Frederiksen et al. \(2012\)](#) find upward biased estimator suggesting a long-memory process, whereas [McCloskey and Perron \(2013\)](#) find evidence for a spurious long-memory process. In an empirical application, we apply the methods robust to perturbation as well as to low frequency contaminations to daily log-absolute returns of the S&P 500 which gives us additional insights which of these effects if not both are of relevance in our data set. We conclude that the analyzed time series follows rather a spurious long-memory process than a perturbed fractional process. This finding is further supported by a test of [Qu \(2011\)](#) that rejects the null hypothesis of a true long-memory process at the 1% significance level.

The structure of the paper is as follows. In the next section, we introduce a variety of modified semiparametric estimation methods of the memory parameter. In Section 3, we analyze the small sample performance of the different estimation approaches in a Monte Carlo study. Section 4 presents the empirical application and Section 5 concludes. Supplementary Materials with detailed Monte Carlo results are provided on the authors webpages.

2. Semiparametric Estimation Methods

In this section, we introduce several concepts of standard and modified estimation methods. We start with general long-memory definitions and briefly review the two standard estimation methods. An overview of modified local Whittle and log-periodogram estimation methods follows which have been designed to capture low frequency contaminations and to reduce the bias of the standard methods. Afterwards, we consider approaches which focus on perturbed fractional processes and therefore account for short-run and additive noise effects.

2.1. The Model and Standard Estimation Methods

A pure long-memory process is typically modeled as an *ARFIMA* (p, d, q) process which has been developed by [Granger and Joyeux \(1980\)](#) and [Hosking \(1981\)](#). It is defined as

$$\Phi(L)(1-L)^d y_t = \Theta(L)\epsilon_t,$$

where $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$, $\Phi(L)$ and $\Theta(L)$ are the autoregressive and moving average polynomials, respectively, with all roots lying outside the unit circle. The order of fractional integration is given by d , L denotes the backshift operator and $(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)L^j}{\Gamma(-d)\Gamma(j+1)}$ with the Gamma function $\Gamma(z)$. The interval $0 < d < 1/2$ implies long memory and short memory exists for $d = 0$ with an exponentially decaying autocorrelation function and a bounded spectral density function for zero frequencies. The process is stationary and invertible for the interval $-1/2 < d < 1/2$ as well as mean reverting for $d < 1$. Typically, a long-memory process is defined by not summable autocovariances, $\sum_{k=-\infty}^{\infty} |\gamma(k)| = \infty$. Further, the spectral density function of a stationary *ARFIMA*

process is given by $f(\lambda) = G\lambda^{-2d}$ for frequency λ close to the origin and a slowly varying function G (cf. Doukhan et al. 2002).

In the following, we introduce the data generating process with a constant term c and y_t being a short or long-memory process,

$$x_t = c + y_t + Au_t + Bw_t. \quad (1)$$

The low frequency contamination and additive noise component are captured by u_t and w_t , where A is a dummy variable which is 1 if we include low frequency contaminations in the DGP and zero otherwise. The dummy variable B is its respective equivalent for a possible perturbation process. We consider several types of low frequency contamination, among them are a sinus time trend and randomly (non-)stationary as well as deterministic level shifts. The perturbation process is captured by different forms of ARMA processes with varying coefficients and variances of the error term. The explicit structure of u_t and w_t will be discussed in Section 3 in more detail. The two standard estimation methods are designed for a process of the form $x_t = y_t$. The four estimators especially designed for low frequency components consider the data generating process $x_t = y_t + u_t$, except for Smith (2005) who excludes the short-memory process in y_t . Additionally, Hou and Perron (2014) extend the process to $x_t = y_t + u_t + w_t$ with w_t being a zero mean white noise process with variance σ_w^2 . The three estimators developed for perturbed fractional processes are exhibiting the following DGP structure $x_t = y_t + w_t$ with w_t being an ARMA process. Finally, a joint definition of the periodogram for the individual processes in Equation (1) with $I_v(\lambda) = (2\pi T)^{-1} |\sum_{t=1}^T v_t \exp(-it\lambda)|^2$ with $v = x, y, u$ and T being the sample size is provided by Hou and Perron (2014).

Before we compare the modified estimation methods, we briefly summarize the properties of two most commonly applied semiparametric estimation procedures of Geweke and Porter-Hudak (1983) and Kuensch (1987). They serve as benchmark approaches, since the modified versions should generally produce a smaller bias in the presence of contaminations. Further, the efficiency loss of the modified versions in situations without contaminations can be illustrated by comparing their performance to the standard methods. Geweke and Porter-Hudak (1983) propose the log-periodogram estimator as one of the earliest estimator for a semiparametric model. The GPH estimator is based on a regression model which is obtained by replacing the spectral density of a process by the logarithm of its periodogram,

$$\log I_y(\lambda_j) = c + dX_j + e_j, \quad \lambda_j = 2\pi j/T, \quad j = 1, \dots, m, \quad (2)$$

with a constant term c and an error term e_j . The explanatory variable is given by $X_j = -\log(2 - 2\cos(\lambda_j))$ with the bandwidth parameter m and the Fourier frequency λ_j . Minimizing the sum of squared residuals of the above regression model with respect to the slope coefficient yield

$$\hat{d}_{GPH} = -\frac{1}{2} \frac{\sum_{j=1}^m (Y_j - \bar{Y}) \log I_y(\lambda_j)}{\sum_{j=1}^m (Y_j - \bar{Y})^2},$$

with $Y_j = \log |1 - \exp(-i\lambda_j)| = \log |2 \sin(\lambda_j/2)|$ and $\bar{Y} = 1/m \sum_{j=1}^m Y_j$. Robinson (1995) shows that the asymptotic distribution equals $\sqrt{m}(\hat{d}_{GPH} - d) \xrightarrow{d} \mathcal{N}(0, \frac{\pi^2}{24})$ for $d \in (-1/2, 1/2)$. For the following asymptotic properties of the different estimation methods we assume the stationary interval of the memory parameter, if not stated differently. The local Whittle method of Kuensch (1987), also known as the Gaussian semiparametric estimation method, is often applied in the literature. The estimator is based on the local Whittle likelihood function

$$Q(G, d) = \frac{1}{m} \sum_{j=1}^m \left\{ \log(G\lambda_j^{-2d}) + \frac{I_y(\lambda_j)}{G\lambda_j^{-2d}} \right\},$$

with $j = 1, \dots, \lfloor T/2 \rfloor$. Minimizing $Q(G, d)$ with respect to G leads to the profiled likelihood function

$$R(d) = \log \left(\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d} I_y(\lambda_j) \right) - \frac{2d}{m} \sum_{j=1}^m \log \lambda_j. \tag{3}$$

The local Whittle estimator is given by $\hat{d}_{LW} = \arg \min_d R(d)$ and converges to a normal distribution according to $\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} \mathcal{N} \left(0, \frac{1}{4} \right)$.

2.2. Modified Log-Periodogram Estimator of Smith (2005)

The form of the spectral density is influenced by low frequency contaminations, however, standard estimation methods are not considering this change and biased estimators emerge. Smith (2005) introduces a modified log-periodogram estimator. For rare level shifts, the part of the adjusted spectral density that is mainly responsible for the high bias of the standard GPH method equals $-\log(p^2 + \lambda^2)$. Since the level shift probability p is unknown an estimate $p_T = km/T$ is introduced with $k > 0$. The estimation approach considers the new form of the spectral density with p_T by including the above component of the adjusted spectral density as a regressor in Equation (2). The regression model of the log-periodogram estimator is extended by the additional regressor to

$$\log I_x(\lambda_j) = c + dX_j + \tau Z_j + e_j,$$

with X_j as defined above and $Z_j = -\log \left((km/T)^2 + \lambda_j^2 \right)$. The modified log-periodogram estimator is given by

$$\hat{d}_{SMI} = (\tilde{X}'M\tilde{X})^{-1} \tilde{X}'M \log I_x(\lambda_j),$$

with $\tilde{X} = X - \bar{X}$, the annihilator matrix $M = I - P = I - \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'$ with P denoting the projection matrix and $\tilde{Z} = Z - \bar{Z}$. Simulation studies of Smith (2005) recommends to set $k = 3$. The limiting distribution is similar to the log-periodogram distribution with $\sqrt{m}(\hat{d}_{SMI} - d) \xrightarrow{d} \mathcal{N} \left(0, \frac{\pi^2}{24z_k} \right)$. However, the bias reduction implicates a variance increase which is captured by the scale factor z_k , with $z_3 = 0.31$ for $k=3$.

2.3. Modified Local Whittle Estimator of Iacone (2010)

In contrast to modifying the form of the spectral density, Iacone (2010) proposes a trimmed version of the Gaussian semiparametric estimator of Kuensch (1987). The idea of the procedure is to balance the frequencies of the periodogram. On the one hand, enough frequencies should be considered so that the stationary ARFIMA process dominates in this region and on the other hand lower frequencies for which the contamination component is most influential should be excluded. Therefore, a trimming parameter l is introduced to the local Whittle likelihood function and the loss function in Equation (3) can be rewritten as,

$$R(d) = \log \left\{ \frac{1}{m-l+1} \sum_{j=l}^m \lambda_j^{2d} I_x(\lambda_j) \right\} - \frac{2d}{m-l+1} \sum_{j=l}^m \log \lambda_j.$$

The modified local Whittle estimator is given by $\hat{d}_{IAC} = \arg \min_d R(d)$ and by setting $l = 1$ the local Whittle estimator is obtained. In general, the trimming and bandwidth parameter need to be set to $l = \max\{1, \lfloor K_l T^\alpha \rfloor\}$ and $m = \max\{l + 1, \lfloor K_m T^\beta \rfloor\}$ with $0 \leq \alpha < \beta < 1$ for some $K_l, K_m \in (0, \infty)$ and $1 \leq l < m \leq \lfloor T/2 \rfloor$. Under certain assumptions of the trimming and bandwidth parameter, the estimator of [Iacone \(2010\)](#) follows the limiting distribution of the local Whittle estimator without an inflated variance caused by the trimming technique, $\sqrt{m}(\hat{d}_{IAC} - d) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{4}\right)$.

2.4. Modified Log-Periodogram Estimator of [McCloskey and Perron \(2013\)](#)

In addition to [Smith \(2005\)](#), who introduces a biased correction form of the GPH estimator, [McCloskey and Perron \(2013\)](#) provide a trimmed version of the popular log-periodogram estimator of [Geweke and Porter-Hudak \(1983\)](#). One major difference between these two estimation methods is the restriction on the process. The modified GPH estimator of [Smith \(2005\)](#) is designed for a short-memory process, whereas [McCloskey and Perron \(2013\)](#) extend their estimator to short or long-memory processes. The standard log-periodogram regression from Equation (2) is adjusted by the trimming parameter l and given by

$$\log I_x(\lambda_j) = c + dX_j + e_j,$$

with $j = l, \dots, m$. Applying the ordinary least squares method, the modified GPH estimator can be expressed as

$$\hat{d}_{MCP} = -\frac{1}{2} \frac{\sum_{j=l}^m (Y_j - \bar{Y}) \log I_x(\lambda_j)}{\sum_{j=l}^m (Y_j - \bar{Y})^2}, \quad (4)$$

with $Y_j = \log|1 - \exp(-i\lambda_j)|$ and $\bar{Y} = (m - l + 1)^{-1} \sum_{j=l}^m Y_j$. In addition to the general requirements on the trimming and bandwidth parameter mentioned above for the previous estimator, [McCloskey and Perron \(2013\)](#) constrained the trimming parameter to a strictly positive α value. Further, an additional condition on the two user-chosen parameters is expressed by $(m \log m)/T + (l \log^2 m)/m + (T^{1-2d} \log^4 m)/(l^{2-2d}) \rightarrow 0$ as $T \rightarrow \infty$. To satisfy this condition α is restricted to certain values depending on the memory parameter. For the interval of particular interest, $0 \leq d < 1/2$, the exponent of the trimming parameter needs to be set to $\alpha = 1/2 + \varepsilon$ with $\varepsilon > 0$. The optimal choice of the upper and lower bound of the sum in (4) is crucial for this estimation method.

Under restrictive assumptions concerning the trimming and bandwidth parameter, the limiting distribution of the trimmed GPH estimator shows no efficiency loss in cases of low frequency contaminations so that $\sqrt{m}(\hat{d}_{MCP} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right)$ is still valid. However, in the absence of a contamination component, the variance increases and, compared to the trimmed estimator of [Iacone \(2010\)](#), the variance increase is sometimes extremely high.

2.5. Modified Local Whittle Estimator of [Hou and Perron \(2014\)](#)

This approach considers the new form of the approximated spectral density when low frequency contaminations are present. In contrast to [Iacone \(2010\)](#) and [McCloskey and Perron \(2013\)](#), this method considers all data and no trimming technique is required. Adding the new term $G_u \lambda^{-2}/T$ to the standard form of the spectral density, a so-called pseudo spectral density can be written as

$$f_x(\lambda) = G\lambda^{-2d} + G_u \lambda^{-2}/T = G(\lambda^{-2d} + (G_u/G)\lambda^{-2}/T) = G(\lambda^{-2d} + \theta_u \lambda^{-2}/T),$$

where θ_u is the signal-to-noise ratio. The new parameter θ_u controls the influence of contaminations at low frequencies. The pseudo spectral density aims to provide a good approximation of the true

spectral density in the existence of low frequency contaminations. The pseudo likelihood function is given by

$$Q(G, d, \theta_u) = \frac{1}{m} \sum_{j=1}^m \left\{ \log f_x(\lambda_j) + \frac{I_x(\lambda_j)}{f_x(\lambda_j)} \right\} \tag{5}$$

and the modified local Whittle estimator is defined as $\hat{d}_{HP} = \arg \min_{d, \theta_u} R(d, \theta_u)$ for $d \in (-1/2, 1)$ and $\theta_u \in (0, \infty)$ with

$$R(d, \theta_u) = \log \left(\frac{1}{m} \sum_{j=1}^m \frac{I_x(\lambda_j)}{\lambda_j^{-2d} + \theta_u \lambda_j^{-2} / T} \right) + \frac{1}{m} \sum_{j=1}^m \log \left(\lambda_j^{-2d} + \theta_u \lambda_j^{-2} / T \right).$$

In the absence of low frequency contamination, the form of the pseudo likelihood function in Equation (5) reduces to the standard local Whittle likelihood function. Hence, no asymptotic efficiency loss exists with $\sqrt{m}(\hat{d}_{HP} - d) \xrightarrow{d} \mathcal{N} \left(0, \frac{1}{4} \right)$ for $d \in [0, 1/2)$. In contrast to the aforementioned authors, [Hou and Perron \(2014\)](#) are the first who provide an extension to account for additive noise and short-memory effects. The modifications are based on the concepts of [Andrews and Sun \(2004\)](#), [Hurvich et al. \(2005\)](#) and [Frederiksen et al. \(2012\)](#) which follow in the next subsections.

2.6. Local Polynomial Whittle Estimator of [Andrews and Sun \(2004\)](#)

In addition to the modified estimation methods of the memory parameter, we consider three estimation methods for perturbed fractional processes. [Arteche \(2004\)](#) shows that when the local Whittle and log-periodogram estimator are applied to perturbed fractional processes they are biased but remain consistent and asymptotical normal. Therefore, several estimation methods have been developed to account for the perturbation. The following three estimation methods modify the local Whittle estimator of [Kuensch \(1987\)](#).

In contrast to a pure fractional process, the short-run dynamics and long-run behavior of a perturbed fractional process are no longer driven by the same innovation. As defined in [Frederiksen et al. \(2012\)](#), a perturbed fractional process can be expressed by a signal-plus-noise model with the signal process being a long-memory process, which is perturbed by an additive noise term. Following the authors, we use an ARMA process for the perturbation process.

The estimation method of [Andrews and Sun \(2004\)](#) adjusts the form of the spectral density by providing a polynomial structure. For frequencies close to the origin the term $\log G$ can be approximated by $\log G - \gamma(\theta_s, \lambda)$ with the polynomial $\gamma(\theta_s, \lambda) = \sum_{k=1}^r \theta_k \lambda^{2k}$. This leads to the following form of the spectral density $f(\lambda) = G \lambda^{-2d} = G \lambda^{-2d} \exp(-\gamma(\theta_s, \lambda))$ and the resulting estimator is known as local polynomial Whittle (LPW). The log-likelihood function of the estimator is given by

$$Q(G, d, \theta_s) = \frac{1}{m} \sum_{j=1}^m \left\{ \log \left[G \lambda_j^{-2d} \exp(-\gamma(\theta_s, \lambda_j)) \right] + \frac{I_x(\lambda_j)}{G \lambda_j^{-2d} \exp(-\gamma(\theta_s, \lambda_j))} \right\}, \tag{6}$$

where $\gamma(\theta_s, \lambda_j) = \sum_{k=1}^r \theta_k \lambda_j^{2k}$ and $\theta_s = (\theta_1, \dots, \theta_r)'$. The local polynomial Whittle estimator is defined as $\hat{d}_{LPW} = \arg \min_{d, \theta_s} R(d, \theta_s)$ and θ_s will be in a set of \mathbb{R} with

$$R(d, \theta_s) = \log \left(\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d} I_x(\lambda_j) \exp(\gamma(\theta_s, \lambda_j)) \right) - \frac{2d}{m} \sum_{j=1}^m \log \lambda_j - \frac{1}{m} \sum_{j=1}^m \gamma(\theta_s, \lambda_j) + 1,$$

and, by setting $\theta_s = 0$ in Equation (6), the standard local Whittle estimator is obtained. The polynomial degree r influences on the one hand the variance and on the other hand the convergence rate of the estimator. We are more interested in a reduced variance than in a convergence rate close to a parametric one. Therefore, we choose a small polynomial degree and set $r = 1$.

The variance of the LPW estimator increases by a constant c_r that depends on the polynomial degree r since a polynomial structure of the spectral density is considered. The asymptotic properties can be summarized by $\sqrt{m}(\hat{d}_{LPW} - d) \xrightarrow{d} \mathcal{N}(0, \frac{c_r}{4})$ where the constant term in the variance is increasing with r and equals $c_1 = 9/4$ for $r = 1$.

2.7. Local Whittle with Noise Estimator of [Hurvich et al. \(2005\)](#)

[Hurvich et al. \(2005\)](#) suggest to replace the constant component G in the spectral density by a more complex function $G(1 + h(d, \theta_\rho, \lambda))$. We focus on the following parameterizations of $h(d, \theta_\rho, \lambda) = \theta_\rho \lambda^{2d}$. The local Whittle with noise (LWN) estimator of [Hurvich et al. \(2005\)](#) is nested in the estimator of [Frederiksen et al. \(2012\)](#) and the additional parameter θ_ρ is labeled according to them. The above parameterization of $h(d, \theta_\rho, \lambda)$ excludes serial correlation between the noise term and the long-memory process, unlike the method of [Frederiksen et al. \(2012\)](#). With the improved form of the spectral density and the additional parameter θ_ρ , the log-likelihood function equals

$$Q(G, d, \theta_\rho) = \sum_{j=1}^m \left\{ \log \left[G \lambda_j^{-2d} (1 + h(d, \theta_\rho, \lambda_j)) \right] + \frac{I_x(\lambda_j)}{G \lambda_j^{-2d} (1 + h(d, \theta_\rho, \lambda_j))} \right\}.$$

Minimizing $Q(G, d, \theta_\rho)$ with respect to G leads to the profiled likelihood function

$$R(d, \theta_\rho) = \log \left(\frac{1}{m} \sum_{j=1}^m \frac{I_x(\lambda_j) \lambda_j^{2d}}{1 + h(d, \theta_\rho, \lambda_j)} \right) - \frac{2d}{m} \sum_{j=1}^m \log (\lambda_j (1 + h(d, \theta_\rho, \lambda_j))).$$

The LWN estimator is defined as $\hat{d}_{LWN} = \arg \min_{d, \theta_\rho} R(d, \theta_\rho)$ for $d \in (0, 1)$ and $\theta_\rho \in (0, \infty)$. The standard local Whittle estimator is obtained in the case of $h(d, \theta_\rho, \lambda) = 0$. An important difference to the previously introduced estimation methods is the domain of the memory parameter since the short-memory case is explicitly not allowed, instead d is supposed to be greater than zero. Allowing for no correlation between the signal and noise process, the variance of the LWN estimator is depending on d with $d \in (0, 3/4)$, so that the limiting distribution is expressed by $\sqrt{m}(\hat{d}_{LWN} - d) \xrightarrow{d} \mathcal{N}\left(0, \frac{(1+d)^2}{16d^2}\right)$.

2.8. Local Polynomial Whittle with Noise Estimator of [Frederiksen et al. \(2012\)](#)

[Frederiksen et al. \(2012\)](#) combine the approaches of [Andrews and Sun \(2004\)](#) and [Hurvich et al. \(2005\)](#) to create a new estimation method, the so-called local polynomial Whittle with noise (LPWN) estimator. In contrast to [Hurvich et al. \(2005\)](#), the spectral density of the LPWN estimator is locally approximated by polynomials instead of a constant near the zero frequency. Furthermore, this procedure differs from the aforementioned methods, since it considers the correlation between the signal and the noise process. Therefore, two polynomials are introduced that capture the noise, r_w , and short-memory effects r_y . [Frederiksen et al. \(2012\)](#) propose the following approximation of the spectral density $f_x(\lambda_j) \sim G \lambda^{-2d} h(d, \theta, \lambda)$. The LPWN estimator assumes an additional functional form of the spectral density that is expressed by

$$h(d, \theta, \lambda) = \exp(h_y(\theta_y, \lambda)) + \theta_\rho \lambda^{2d} \exp(h_w(\theta_w, \lambda)) \tag{7}$$

with $\theta = (\theta'_y, \theta_\rho, \theta'_w)'$ and by setting $h(d, \theta, \lambda) = 1$ the spectral density of the local Whittle estimator is obtained. The exponents in Equation (7) exhibit the polynomial structure of the earlier introduced

estimator of Andrews and Sun (2004) with $h_a(\theta_a, \lambda) = \sum_{k=1}^{r_a} \theta_{a,k} 2^k$ where $a = y, w$ and $h_a(\theta_a, \lambda) = 0$ for $r_a = 0$. Additionally, by setting both polynomials r_y and r_w to zero the LWN estimator of Hurvich et al. (2005) is obtained with $h(d, \theta, \lambda) = 1 + \theta_\rho \lambda^{2d}$. Further, the authors distinguish between three different types of LPWN(r_y, r_w) estimators, the LPWN(0,1), LPWN(1,0) and LPWN(1,1). The LPWN estimator is obtained by $\hat{d}_{LPWN} = \arg \min_{d, \theta} R(d, \theta)$ for $d \in (0, 1)$ and $\theta \in (\mathbb{R}^y \times (0, \infty) \times \mathbb{R}^w)$ with the objective function

$$R(d, \theta) = \log \left(\frac{1}{m} \sum_{j=1}^m \frac{\lambda_j^{2d} I_x(\lambda_j)}{h(d, \theta, \lambda_j)} \right) - \frac{2d}{m} \sum_{j=1}^m \log \lambda_j + \frac{1}{m} \sum_{j=1}^m \log h(d, \theta, \lambda_j).$$

The polynomials increase the asymptotic variance of $\sqrt{m}(\hat{d}_{LPWN} - d)$ by a multiplicative constant depending only on d and not on the parameters θ , with $d \in (0, 3/4)$. The LPWN estimator is normally distributed with a zero mean and exhibits a more complex form of the variance compared to the aforementioned methods, which can be found in Frederiksen et al. (2012).

3. Monte Carlo Analysis

This section analyzes the performance of the previously presented estimators under different scenarios. First, we introduce the Monte Carlo setup with different kind of DGPs and parameter settings for the estimation methods. Afterwards, we present the simulation results by focusing on the bias and root mean squared error (RMSE) of the estimators.

3.1. Monte Carlo Setup

We compare the performance of the estimation methods under nine different DGPs, which are given in Table 1. The DGP types are obtained by adapting parts of the Monte Carlo structure provided by Qu (2011), Frederiksen et al. (2012) and Hou and Perron (2014). The process y_t is modeled by an ARFIMA (1,d,0) with $(1 - \alpha_y L)(1 - L)^d y_t = \varepsilon_t$, where ε_t is a zero mean white noise process if not otherwise defined. We standardize the ARFIMA process by its own standard deviation and consider the following values of the autoregressive and persistence coefficient with $d = 0.0, 0.2, 0.4, 0.6$ and $\alpha_y = 0.0, 0.3, 0.6$. By choosing $d = 0.6$, we violate against the stationarity assumption for a variety of estimators to check the robustness in the case of non-stationarity. The sample size is set to $T = 256, 512, 1024, 4096$ and the results are based on 1000 Monte Carlo replications.

Table 1. Overview of data generating processes

no contamination	$x_t = y_t$	
sinus trend	$x_t = y_t + u_t$	$u_t = 3 \sin(4\pi t/T)$
stationary RLS	$x_t = y_t + u_t$	$u_t = (1 - \pi\phi)u_{t-1} + \delta_t$ $\delta_t = \pi_t \eta_t \quad \eta_t \stackrel{iid}{\sim} N(0, 1) \quad \pi_t \stackrel{iid}{\sim} \text{Bernoulli}(10/T, 1) \quad \phi = 1$
non-stationary RLS	$x_t = y_t + u_t$	$u_t = (1 - \pi\phi)u_{t-1} + \delta_t$ $\delta_t = \pi_t \eta_t \quad \eta_t \stackrel{iid}{\sim} N(0, 1) \quad \pi_t \stackrel{iid}{\sim} \text{Bernoulli}(10/T, 1) \quad \phi = 0$
deterministic level shift	$x_t = y_t + u_t$	$u_t = \sum_{i=1}^B c_i \mathbb{1}(T_{i-1} < t \leq T_i)$ $0 < c_i < \infty \quad 0 = T_0 < T_1 < \dots < T_{B-1} < T_B = T$
GARCH (1,1)	$x_t = y_t + u_t$	$u_t = \sigma_t v_t \quad v_t \stackrel{iid}{\sim} N(0, 1) \quad \sigma_t^2 = 0.1u_{t-1}^2 + 0.8\sigma_{t-1}^2$
ARMA(0,0)	$x_t = y_t + w_t$	$w_t = \zeta_t \quad \zeta_t \stackrel{iid}{\sim} N(0, 4)$
ARMA(0,1)	$x_t = y_t + w_t$	$w_t = (1 + \beta_w L)\zeta_t \quad \beta_w = 0.6 \quad \zeta_t \stackrel{iid}{\sim} N(0, 1) \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma_{\varepsilon_t}^2)$
stationary RLS plus noise	$x_t = y_t + u_t + w_t$	$u_t = (1 - \pi\phi)u_{t-1} + \delta_t$ $\delta_t = \pi_t \eta_t \quad \eta_t \stackrel{iid}{\sim} N(0, 1) \quad \pi_t \stackrel{iid}{\sim} \text{Bernoulli}(10/T, 1) \quad \phi = 1$ $w_t = \zeta_t \quad \zeta_t \stackrel{iid}{\sim} N(0, 4)$

In the following, we provide further information on the different types of DGPs given in Table 1. Unreported simulation results of a linear trend as contamination show almost no difference to the performance of a sinus trend and are therefore omitted. The components π_t and η_t of the random level shift (RLS) processes are mutually independent and the shift probability equals $p = 10/T$. For the deterministic level shift process, we consider five shifts at fixed user-chosen dates with $c_i = -2, 2, -4, 2, 1, -2$ for $i = 1, \dots, 6$ and $T_i/T \rightarrow \tau_i$ with $0 < \tau_1 < \dots < \tau_5 < 1$. DGPs 7 and 8 are especially designed for the perturbed fractional estimation methods. The ARMA(0,0) process is basically an additive white noise process with a large variance. Unreported simulation results of an ARMA(1,0) process with $\alpha_w = 0.6$ are showing no clear pattern and recommendations are hard to formulate. Therefore, we omit the results and instead consider the ARMA(0,1) process. The last DGP combines the two groups of contaminations with a low frequency component, in the form of a stationary RLS, and an additive noise term as a perturbation contamination.

Following Frederiksen et al. (2012), we set the noise-to-signal ratio equal to five and by rearranging the noise-to-signal ratio the variance of the ARFIMA (1,d,0) process is obtained by

$$\sigma_\varepsilon^2 = \frac{(1 - \alpha_y)^2 \cdot \frac{(1 + \beta_w)^2}{(1 - \alpha_w)^2}}{5}.$$

We conduct the simulation study for all three forms of the LPWN estimator of Frederiksen et al. (2012). The bias and RMSE results of the three different types are very similar, especially the LPWN(0,1) and LPWN(1,0) show no significant difference. For the Monte Carlo study and the empirical example, we present the results of the LPWN(1,1) since we are varying the autoregressive coefficient for every DGP and also analyze several noise structures as perturbation. In general, we follow the commonly used parameter settings in the literature to reflect the empirical findings in the financial volatility. For all estimation methods, we choose the following values of the bandwidth $m = \lfloor T^\delta \rfloor$ with $\delta = 0.60, 0.70, 0.79$. The additional parameter of Smith (2005) is set to $k = 3$. To satisfy the trimming and bandwidth condition of Iacone (2010), we choose $m = \lfloor 0.9T^\delta \rfloor$ and $l = \lfloor 0.2T^{\delta-0.05} \rfloor$. We consider the adaptive procedure of McCloskey and Perron (2013), since this extension exhibits a smaller variance. The first step of the adaptive method, $i = 1$, applies the setting of the trimmed GPH estimator with $l = \lfloor K_l T^{1/2+\varepsilon} \rfloor$. For the next iterations, $i > 1$, the trimming changes to $l = \lfloor K_l T^{(1-2\hat{d}_{i-1})/(2-2\hat{d}_{i-1})+\varepsilon} \rfloor$ with \hat{d} as a consistent estimator of d . Two different convergence criteria are used, either the procedure stops as $|\hat{d}_i - \hat{d}_{i-1}| < 0.01$ or when $i > 9$ with the final value of \hat{d} . We set $K_l = 1$ and $\varepsilon = 0.03$ for the estimator of McCloskey and Perron (2013) and the GPH estimator is used as starting value. The numerical optimization of the Monte Carlo study is based on R using the L-BFGS-B algorithm developed by Byrd et al. (1995).

3.2. Monte Carlo Results

In the following, the DGPs are numbered consecutively according to the order in Table 1. The bias and root mean squared error results of all estimation methods and for the corresponding DGPs are given in the Supplementary Materials. Before presenting some detailed results, we are summarizing the major findings of the simulation study.

The two standard estimation methods are very robust compared to the modified methods in two cases. Considering the first DGP without contamination and the ARFIMA (1,d,0) plus GARCH (1,1), the log-periodogram and Gaussian estimator are both characterized by a smaller bias and RMSE especially for $\phi \leq 0.3$. In the case of a larger bandwidth and $\phi = 0.6$, the modified estimators of Andrews and Sun (2004) and Hou and Perron (2014) perform slightly better or equivalent in terms of the RMSE than the local Whittle estimator. The same results hold for the modified estimator of Smith (2005) compared to the GPH estimator.

For a sinus trend or deterministic level shift as contamination type, the modified estimator of Hou and Perron (2014) is outperforming the local Whittle estimator in terms of bias and also for the

majority cases of RMSE. Additionally, for the two variations of the random level shift process, \hat{d}_{HP} performs well, especially for lower values of the memory parameter ($d \leq 0.2$).

The two trimmed versions of the local Whittle and GPH estimator are both exhibiting a very similar pattern over the DGPs. Similar to the estimator of [Hou and Perron \(2014\)](#), the estimators of [Iacone \(2010\)](#) and [McCloskey and Perron \(2013\)](#) also have their strength in the presence of level shifts, irrespective of whether they occur randomly (non-)stationary or deterministic. The good performance in the aforementioned situations is restricted to smaller values of d and to less short-run effects.

While the estimators of [Hou and Perron \(2014\)](#), [Iacone \(2010\)](#) and [McCloskey and Perron \(2013\)](#) often improve the standard estimators in the short-memory case or without short-run dynamics, the methods of [Smith \(2005\)](#), [Andrews and Sun \(2004\)](#), [Hurvich et al. \(2005\)](#) and [Frederiksen et al. \(2012\)](#) outperform their respective standard methods for larger values of the persistence parameter ($d \geq 0.4$) as well as for a higher short-run parameter ($\phi \geq 0.3$). This finding is not very surprising for the three estimation methods designed for perturbation, however the good performance of \hat{d}_{SMI} in these situations is rather unexpected. In particular, the DGPs especially constructed for the perturbation setup are dominated by the perturbation estimators and by [Smith \(2005\)](#). Moreover, unreported simulation studies stress the robustness of the estimation method of [Smith \(2005\)](#) to short-run effects with values larger than $\phi = 0.6$.

After a broad overview of the simulation results, we focus on three different contamination types in more detail: a sinus trend, a stationary random level shift and an additive noise term. Since the parameter intervals of the Monte Carlo study are very comprehensive, we concentrate on a sample size of $T = 1024$ for the following three tables and distinguish between two bandwidths. Further, we omit the persistence value of $d = 0.6$ for the first two detailed DGPs and for the third contamination type we add the memory value $d = 0.6$ but exclude the short-memory case.

The results of a sinus trend as contamination are presented in Table 2. The best performance in terms of lowest bias and RMSE is obtained by the GPH and modified GPH estimator of [McCloskey and Perron \(2013\)](#). Further, the trimmed local Whittle estimator of [Iacone \(2010\)](#) exhibits a comparable small bias and RMSE for $\phi = 0$. In the next step, we compare the modified methods to the respective standard estimators. The modified estimator of [Hou and Perron \(2014\)](#) outperforms the local Whittle estimator for all combinations in terms of bias and RMSE. Generally, the estimator improves the performance with a larger bandwidth, except for situations with a larger autoregressive parameter of the ARFIMA model. The performance of the estimator of [Iacone \(2010\)](#) is very similar with an even smaller bias and RMSE for $\phi = 0.0$ and for $\phi = 0.3$ combined with a smaller bandwidth. Although the trimmed GPH estimator of [McCloskey and Perron \(2013\)](#) performs well in terms of the bias for $\phi \leq 0.3$, the inflated variance increases the RMSE compared to the GPH estimator in most cases. The estimator of [Smith \(2005\)](#) performs better with an increasing autoregressive parameter and for a larger bandwidth. However, the three estimation methods developed for perturbed fractional processes are not an alternative to the standard methods. For the trimmed local Whittle estimator, the reduction of the RMSE is in the range of 60–80% for $\phi = 0.0$ and remains within 30–75% for $\phi = 0.3$. In contrast to the aforementioned estimator, the estimator of [Hou and Perron \(2014\)](#) has a smaller reduction of the RMSE for $\phi = 0.0$ with 20–65% due to the higher bias. For $\phi = 0.3$, the reduction of the RMSE of \hat{d}_{HP} is very similar to the estimator of [Iacone \(2010\)](#) with 35–80% and remains within 20–75% for $\phi = 0.6$, except for $d = 0.4$. The estimator of [Smith \(2005\)](#) provides a smaller reduction of the RMSE compared to the estimator of [Hou and Perron \(2014\)](#) for the highest value of the autoregressive parameter with 5–50%.

Table 2. DGP2:ARFIMA (1,d,0) plus sinus trend for T = 1024.

d	$\phi = 0.0$						$\phi = 0.3$						$\phi = 0.6$					
	0.0		0.2		0.4		0.0		0.2		0.4		0.0		0.2		0.4	
δ	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79
Bias																		
d_{LW}	0.661	0.524	0.457	0.350	0.297	0.215	0.614	0.527	0.432	0.375	0.300	0.268	0.614	0.596	0.470	0.485	0.374	0.420
d_{GPH}	0.119	0.075	0.095	0.063	0.080	0.047	0.153	0.171	0.131	0.158	0.114	0.148	0.267	0.369	0.246	0.355	0.234	0.349
d_{HP}	-0.311	-0.147	-0.288	-0.145	-0.275	-0.160	-0.158	0.059	-0.154	0.048	-0.159	0.037	0.121	0.371	0.119	0.374	0.161	0.413
d_{IAC}	0.002	0.002	-0.007	-0.008	-0.007	-0.017	0.070	0.185	0.057	0.173	0.063	0.167	0.293	0.519	0.288	0.513	0.293	0.503
d_{MCP}	-0.082	-0.015	-0.041	-0.007	0.020	0.006	0.060	0.159	0.045	0.131	0.096	0.148	0.252	0.371	0.250	0.362	0.335	0.503
d_{SMI}	0.287	0.179	0.234	0.146	0.187	0.113	0.257	0.155	0.215	0.125	0.171	0.101	0.227	0.190	0.185	0.163	0.151	0.143
d_{LPW}	0.928	0.719	0.665	0.501	0.448	0.329	0.830	0.650	0.593	0.457	0.413	0.314	0.735	0.625	0.539	0.469	0.401	0.364
d_{LWN}			0.790	0.789	0.582	0.507			0.789	0.674	0.546	0.372			0.673	0.499	0.437	0.423
d_{LPWN11}			0.790	0.789	0.581	0.507			0.789	0.674	0.545	0.372			0.673	0.499	0.436	0.423
RMSE																		
d_{LW}	0.661	0.524	0.457	0.350	0.299	0.216	0.615	0.527	0.432	0.375	0.302	0.270	0.615	0.596	0.470	0.486	0.376	0.421
d_{GPH}	0.134	0.086	0.113	0.076	0.098	0.065	0.163	0.176	0.143	0.164	0.128	0.155	0.273	0.371	0.253	0.357	0.241	0.352
d_{HP}	0.332	0.161	0.311	0.158	0.302	0.173	0.191	0.080	0.185	0.075	0.195	0.072	0.150	0.375	0.151	0.378	0.224	0.415
d_{IAC}	0.104	0.079	0.108	0.076	0.109	0.079	0.128	0.199	0.116	0.190	0.118	0.184	0.312	0.525	0.308	0.518	0.310	0.507
d_{MCP}	2.981	0.105	0.178	0.079	0.127	0.066	0.230	0.171	0.134	0.142	0.148	0.155	0.269	0.373	0.288	0.368	0.397	0.535
d_{SMI}	0.310	0.195	0.257	0.164	0.219	0.139	0.280	0.172	0.240	0.147	0.203	0.130	0.252	0.203	0.215	0.180	0.188	0.164
d_{LPW}	0.929	0.720	0.666	0.501	0.450	0.330	0.831	0.650	0.594	0.458	0.416	0.316	0.736	0.625	0.541	0.470	0.404	0.366
d_{LWN}			0.790	0.789	0.582	0.511			0.789	0.679	0.549	0.376			0.678	0.499	0.441	0.424
d_{LPWN11}			0.790	0.789	0.582	0.510			0.789	0.679	0.548	0.376			0.678	0.499	0.441	0.424

Table 3 presents the results for a stationary random level shift as contamination. In general, the results of the non-stationary RLS in the Supplementary Materials have a similar pattern but with a larger bias due to the non-stationarity and, thus, an increased RMSE. Three out of the four modified estimators developed especially for low frequency contaminations, such as a random level shifts, improve the standard methods, especially for smaller persistence values. The highest reduction in the RMSE for the estimation methods of Hou and Perron (2014), Iacone (2010) and McCloskey and Perron (2013) is obtained for $\phi = 0.3$ with 50–55%, 35–50% and 1–35%, respectively, and with less improvement for an increased short-run effect. The performance of \hat{d}_{SMI} is rather unexpected and shows no improvement against the GPH estimator. The poor results of Smith (2005) are stressed by the fact that the estimator of Andrews and Sun (2004) outperforms \hat{d}_{SMI} , even though the estimator of Smith (2005) has originally been developed for random level shifts, unlike \hat{d}_{LPW} . The three estimation methods not explicitly developed for low frequency contaminations are offering no improvement. Nevertheless, compared to the sinus trend above the bias and RMSE decreased and the findings are getting closer to the other modified estimators, especially for a higher short-run dynamic.

After two typical examples of low frequency contaminations in the form of a time trend and level shift, we consider the effect of an additive noise term with the results given in Table 4. Overall, the estimators of Smith (2005), Hurvich et al. (2005) and Frederiksen et al. (2012) are the three best performing estimation methods in terms of bias and RMSE. In the following, we compare the modified methods to the respective standard estimators. As expected, the three estimation methods developed for perturbed fractional processes are dominating the standard methods in terms of a lower bias and RMSE, although a RMSE reduction is restricted to larger memory values of $d \geq 0.4$. Additionally, the estimator of Smith (2005) performs well compared to the GPH estimator. The reduction of the RMSE for the estimators of Smith (2005) and Andrews and Sun (2004) is decreasing over the ϕ interval and ranging from 20–30% to 15–30% and finally to 10–30%. A similar but slightly larger reduction of the RMSE is obtained for the estimation methods of Hurvich et al. (2005) and Frederiksen et al. (2012) with 20–60%, 15–60% and still 5–60% for the same aforementioned region of ϕ . The largest reduction is gained for $d = 0.6$ and $\delta = 0.79$, irrespective of the previous estimation methods. In the presence of an additive noise term, the two standard methods are more robust than the four modified estimators developed for low frequency contaminations, except for Smith (2005).

The last detailed DGP in Table 5 contains two groups of contamination: a stationary random level shift and an additive noise term. The overall outcome combines the individual results of each contamination type from Table 3 and Table 4. Excluding the short-memory process, the estimators of Smith (2005) and Andrews and Sun (2004) perform best in terms of bias and RMSE. For $d \geq 0.4$ the reduction of the RMSE to the respective standard method of Smith (2005) and Andrews and Sun (2004) equals on average 30%, irrespectively of the short-run influence. On the contrary, the estimators of Iacone (2010) and Hou and Perron (2014) provide the best results for a short-memory process being contaminated. With an increasing influence of the short-run dynamic, the reduction of the RMSE for $d = 0$ decreases on average from 55% to 50% and then to 30% for the estimators of Iacone (2010) and Hou and Perron (2014). Therefore, no procedure provides stable results over the whole range of the persistence parameter in the case of two contamination types.

Table 3. DGP3: ARFIMA (1,d,0) plus stationary RLS for T = 1024.

d	$\phi = 0.0$						$\phi = 0.3$						$\phi = 0.6$					
	0.0		0.2		0.4		0.0		0.2		0.4		0.0		0.2		0.4	
δ	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79
Bias																		
d_{LW}	0.340	0.257	0.159	0.090	0.019	-0.047	0.319	0.270	0.159	0.121	0.033	-0.012	0.339	0.335	0.207	0.197	0.090	0.044
d_{GPH}	0.322	0.225	0.152	0.077	0.020	-0.049	0.304	0.251	0.157	0.120	0.037	-0.008	0.337	0.336	0.210	0.206	0.096	0.054
d_{HP}	0.039	0.017	-0.049	-0.075	-0.128	-0.185	0.089	0.116	0.015	0.026	-0.083	-0.108	0.235	0.282	0.149	0.161	0.034	-0.028
d_{IAC}	0.100	0.035	-0.013	-0.070	-0.102	-0.185	0.121	0.129	0.033	0.037	-0.059	-0.100	0.252	0.328	0.172	0.191	0.062	-0.018
d_{MCP}	0.058	0.026	-0.037	-0.059	-0.089	-0.138	0.106	0.131	0.026	0.050	-0.038	-0.053	0.261	0.306	0.174	0.190	0.082	0.037
d_{SMI}	0.570	0.429	0.337	0.228	0.160	0.086	0.505	0.373	0.292	0.204	0.142	0.083	0.418	0.347	0.244	0.225	0.136	0.135
d_{LPW}	0.479	0.371	0.266	0.181	0.104	0.038	0.428	0.338	0.234	0.171	0.095	0.049	0.374	0.343	0.217	0.215	0.114	0.108
d_{LWN}			0.491	0.459	0.228	0.204			0.403	0.310	0.190	0.148			0.287	0.256	0.161	0.165
d_{LPWN11}			0.455	0.434	0.198	0.182			0.379	0.300	0.170	0.131			0.281	0.247	0.151	0.141
RMSE																		
d_{LW}	0.350	0.266	0.175	0.108	0.068	0.068	0.330	0.277	0.174	0.133	0.071	0.048	0.347	0.339	0.217	0.202	0.106	0.061
d_{GPH}	0.335	0.236	0.174	0.098	0.079	0.074	0.317	0.259	0.176	0.134	0.080	0.053	0.348	0.342	0.223	0.213	0.117	0.074
d_{HP}	0.147	0.072	0.135	0.103	0.175	0.199	0.152	0.133	0.107	0.065	0.136	0.130	0.254	0.287	0.171	0.169	0.098	0.079
d_{IAC}	0.158	0.084	0.115	0.104	0.150	0.201	0.166	0.148	0.113	0.085	0.124	0.125	0.274	0.336	0.202	0.206	0.121	0.082
d_{MCP}	0.224	0.126	0.221	0.121	0.202	0.170	0.230	0.158	0.173	0.097	0.160	0.096	0.291	0.313	0.212	0.200	0.139	0.077
d_{SMI}	0.589	0.445	0.366	0.251	0.206	0.130	0.526	0.389	0.325	0.229	0.194	0.130	0.444	0.364	0.278	0.245	0.186	0.164
d_{LPW}	0.493	0.382	0.286	0.198	0.141	0.082	0.442	0.348	0.255	0.188	0.134	0.086	0.390	0.352	0.237	0.226	0.144	0.125
d_{LWN}			0.520	0.484	0.269	0.238			0.437	0.339	0.233	0.185			0.310	0.269	0.189	0.189
d_{LPWN11}			0.489	0.463	0.244	0.222			0.415	0.330	0.214	0.171			0.304	0.259	0.176	0.166

Table 4. DGP8: ARFIMA (1,d,0) plus ARMA(0,0) for T = 1024.

d	$\phi = 0.0$						$\phi = 0.3$						$\phi = 0.6$					
	0.2		0.4		0.6		0.2		0.4		0.6		0.2		0.4		0.6	
δ	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79
Bias																		
<i>d_{LW}</i>	-0.134	-0.144	-0.227	-0.258	-0.333	-0.393	-0.105	-0.105	-0.196	-0.228	-0.322	-0.380	-0.034	-0.040	-0.150	-0.191	-0.300	-0.368
<i>d_{GPH}</i>	-0.132	-0.141	-0.225	-0.257	-0.339	-0.406	-0.100	-0.102	-0.191	-0.225	-0.329	-0.394	-0.032	-0.036	-0.145	-0.188	-0.304	-0.381
<i>d_{HP}</i>	-0.151	-0.155	-0.273	-0.298	-0.459	-0.504	-0.123	-0.115	-0.243	-0.267	-0.455	-0.498	-0.046	-0.046	-0.191	-0.232	-0.433	-0.492
<i>d_{IAC}</i>	-0.148	-0.154	-0.284	-0.323	-0.469	-0.529	-0.112	-0.108	-0.255	-0.288	-0.458	-0.519	-0.012	-0.032	-0.193	-0.262	-0.435	-0.521
<i>d_{MCP}</i>	-0.176	-0.179	-0.330	-0.336	-0.506	-0.539	-0.132	-0.117	-0.296	-0.297	-0.508	-0.529	-0.050	-0.046	-0.233	-0.259	-0.467	-0.532
<i>d_{SMI}</i>	-0.112	-0.126	-0.153	-0.190	-0.189	-0.271	-0.091	-0.103	-0.125	-0.157	-0.172	-0.256	-0.055	-0.040	-0.083	-0.110	-0.147	-0.234
<i>d_{LPW}</i>	-0.125	-0.136	-0.184	-0.216	-0.246	-0.314	-0.107	-0.107	-0.157	-0.185	-0.233	-0.297	-0.053	-0.037	-0.116	-0.136	-0.207	-0.279
<i>d_{LWN}</i>	0.013	0.003	-0.011	-0.003	-0.013	0.000	0.029	-0.004	-0.004	-0.014	-0.002	0.008	0.035	0.012	-0.009	0.008	0.000	0.015
<i>d_{LPWN11}</i>	-0.010	-0.024	-0.055	-0.049	-0.061	-0.042	0.007	-0.023	-0.040	-0.052	-0.046	-0.034	0.025	-0.001	-0.032	-0.036	-0.044	-0.031
RMSE																		
<i>d_{LW}</i>	0.142	0.148	0.231	0.260	0.335	0.394	0.115	0.110	0.200	0.230	0.324	0.381	0.056	0.050	0.155	0.193	0.301	0.369
<i>d_{GPH}</i>	0.145	0.148	0.233	0.260	0.343	0.408	0.117	0.110	0.199	0.228	0.333	0.396	0.067	0.056	0.154	0.192	0.309	0.383
<i>d_{HP}</i>	0.161	0.160	0.281	0.302	0.469	0.508	0.134	0.120	0.253	0.271	0.466	0.502	0.070	0.057	0.202	0.237	0.448	0.498
<i>d_{IAC}</i>	0.183	0.172	0.303	0.333	0.481	0.535	0.156	0.130	0.278	0.299	0.471	0.525	0.105	0.084	0.219	0.273	0.450	0.528
<i>d_{MCP}</i>	0.248	0.211	0.375	0.349	0.594	0.550	0.378	0.142	0.359	0.312	0.565	0.540	0.223	0.096	0.282	0.279	0.609	0.545
<i>d_{SMI}</i>	0.162	0.149	0.191	0.204	0.215	0.279	0.150	0.129	0.168	0.176	0.200	0.264	0.135	0.088	0.136	0.134	0.176	0.245
<i>d_{LPW}</i>	0.146	0.146	0.198	0.222	0.253	0.316	0.132	0.118	0.173	0.191	0.240	0.299	0.093	0.062	0.133	0.143	0.214	0.281
<i>d_{LWN}</i>	0.193	0.188	0.186	0.175	0.160	0.143	0.186	0.148	0.172	0.152	0.162	0.141	0.130	0.088	0.140	0.128	0.149	0.138
<i>d_{LPWN11}</i>	0.175	0.173	0.186	0.180	0.179	0.163	0.170	0.137	0.166	0.169	0.174	0.158	0.118	0.072	0.156	0.130	0.159	0.153

Table 5. DGP9: ARFIMA (1,d,0) plus stationary RLS plus additive noise for T = 1024.

d	$\phi = 0.0$						$\phi = 0.3$						$\phi = 0.6$					
	0.0		0.4		0.6		0.0		0.4		0.6		0.0		0.4		0.6	
δ	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79	0.70	0.79
Bias																		
<i>d_{LW}</i>	0.221	0.157	-0.131	-0.188	-0.283	-0.356	0.212	0.167	-0.120	-0.169	-0.275	-0.347	0.225	0.196	-0.086	-0.150	-0.262	-0.339
<i>d_{GPH}</i>	0.205	0.134	-0.137	-0.199	-0.293	-0.374	0.199	0.149	-0.123	-0.178	-0.281	-0.366	0.220	0.190	-0.087	-0.157	-0.269	-0.358
<i>d_{HP}</i>	0.031	0.016	-0.274	-0.307	-0.476	-0.523	0.049	0.048	-0.248	-0.280	-0.465	-0.515	0.118	0.119	-0.195	-0.257	-0.448	-0.517
<i>d_{IAC}</i>	0.041	0.014	-0.272	-0.320	-0.462	-0.532	0.048	0.048	-0.249	-0.290	-0.447	-0.523	0.125	0.133	-0.193	-0.273	-0.431	-0.521
<i>d_{MCP}</i>	0.005	-0.007	-0.317	-0.333	-0.502	-0.543	-0.045	0.043	-0.290	-0.295	-0.443	-0.532	0.118	0.120	-0.217	-0.275	-0.464	-0.533
<i>d_{SMI}</i>	0.399	0.276	0.022	-0.069	-0.101	-0.205	0.368	0.259	0.024	-0.054	-0.089	-0.198	0.325	0.262	0.037	-0.028	-0.086	-0.180
<i>d_{LPW}</i>	0.329	0.237	-0.040	-0.112	-0.174	-0.257	0.303	0.230	-0.035	-0.093	-0.165	-0.250	0.280	0.240	-0.013	-0.068	-0.155	-0.234
<i>d_{LWN}</i>			0.205	0.201	0.077	0.087			0.179	0.174	0.075	0.085			0.149	0.161	0.063	0.094
<i>d_{LPWN11}</i>			0.154	0.164	0.031	0.046			0.135	0.137	0.030	0.049			0.114	0.119	0.019	0.049
RMSE																		
<i>d_{LW}</i>	0.233	0.168	0.146	0.194	0.332	0.385	0.223	0.175	0.134	0.175	0.280	0.349	0.236	0.203	0.103	0.155	0.267	0.342
<i>d_{GPH}</i>	0.220	0.146	0.154	0.206	0.300	0.377	0.213	0.159	0.140	0.185	0.289	0.370	0.234	0.197	0.110	0.164	0.276	0.361
<i>d_{HP}</i>	0.108	0.058	0.292	0.313	0.490	0.528	0.112	0.073	0.267	0.286	0.480	0.520	0.148	0.129	0.219	0.266	0.465	0.522
<i>d_{IAC}</i>	0.115	0.076	0.293	0.329	0.474	0.537	0.116	0.090	0.270	0.301	0.460	0.528	0.166	0.153	0.221	0.284	0.445	0.526
<i>d_{MCP}</i>	0.465	0.112	0.402	0.351	0.573	0.556	2.195	0.099	0.339	0.312	1.551	0.543	0.336	0.147	0.275	0.295	0.497	0.546
<i>d_{SMI}</i>	0.424	0.295	0.132	0.117	0.157	0.224	0.392	0.277	0.130	0.106	0.146	0.215	0.354	0.281	0.134	0.094	0.150	0.198
<i>d_{LPW}</i>	0.345	0.251	0.103	0.131	0.193	0.265	0.320	0.242	0.098	0.114	0.183	0.257	0.299	0.252	0.087	0.091	0.174	0.242
<i>d_{LWN}</i>			0.271	0.260	0.180	0.168			0.252	0.236	0.173	0.170			0.223	0.218	0.175	0.171
<i>d_{LPWN11}</i>			0.244	0.242	0.178	0.168			0.227	0.220	0.170	0.167			0.200	0.197	0.171	0.168

In the following, we give a broad user-guideline by focusing on the RMSE. Usually, an economist is not aware of specific trends, level shifts or perturbation types when empirical data are analyzed. In the following, we compare the effects of two possible scenarios. The first situation is characterized by a perturbation process and instead of using the more appropriate estimators designed for this specific case one falsely applies estimation methods developed for low frequency contamination. The Monte Carlo results indicate a moderate increase of the RMSE for low frequency contaminated estimators, apart from $d = 0.6$. The next scenario summarizes the reverse case, which means that, for a low frequency contaminated process, the estimators developed to capture perturbation are used. Depending on the specific type of low frequency contamination, the Monte Carlo study shows mixed results. Considering a sinus trend or deterministic level shift the RMSE increases substantially for the perturbation estimators, in some cases three times as much as the estimators developed for low frequency contamination. The increase of the RMSE for a (non-)stationary RLS is lower than for the two aforementioned processes but still more than twice as high as the RMSE of low frequency contaminated estimators. The random level shift process plus an additive noise term shows slightly higher root mean squared errors. Considering the three different types of random level shifts an improvement of the RMSE is obtained for $d = 0.6$ when the perturbation estimators are used instead of the methods designed for low frequency contamination. Therefore, we recommend the estimators developed for low frequency contamination, since the increase in a perturbation setup is not as severe as in the reverse case, especially in the stationary region of the memory parameter.

In summary, our results show a good performance of \hat{d}_{SMI} in situations of a high short-run dynamic and in typical perturbation situations, however, in cases of low frequency contaminations, the estimator shows some weaknesses. The estimators of [Iacone \(2010\)](#) and [Hou and Perron \(2014\)](#) perform very similar in terms of bias and RMSE, whereas \hat{d}_{HP} is slightly better. In many situations, the estimator of [McCloskey and Perron \(2013\)](#) cannot compete with the other modified estimators developed for low frequency contamination in terms of RMSE due to the higher variance. Finally, the estimators of [Hurvich et al. \(2005\)](#) and [Frederiksen et al. \(2012\)](#) yield almost identical findings, which are in most situations better than the results of [Andrews and Sun \(2004\)](#).

4. Application

In this section, we analyze the daily log-absolute return series of the S&P 500 from 1 January 1997 to 31 October 2017 ($T = 5245$) to investigate the robustness of the aforementioned results to the choice of estimator. Furthermore, we apply methods robust to perturbation as well as to low frequency contaminations which gives us additional insights which of these effects if not both are of relevance in our data set. We follow the method of [Xu and Perron \(2014\)](#) and calculate the log-absolute returns as $\log(|r_t| + 0.001)$. The log-returns are obtained by differencing the logarithm of the price index, $r_t = \log(p_t) - \log(p_{t-1})$, and the constant 0.001 is added to avoid the problem of taking the logarithm of zero. Figure 1 displays the time series of the S&P 500 with the hyperbolically decaying autocorrelation function and the periodogram with a pole at the origin, as typical characteristics of a long-memory process.

We estimate the persistence parameter with nine different estimation methods and vary the bandwidth parameter according to our Monte Carlo study with $m = \lfloor 1 + T^\delta \rfloor$ and $\delta = 0.60, 0.70, 0.79$. The parameter settings of the corresponding estimation methods remain unchanged to the simulation study above. Table 6 shows the estimation results of \hat{d} and the standard errors which are given in brackets below. Additionally, we apply the univariate test of [Qu \(2011\)](#) to test the null hypothesis of a true long-memory process against spurious long memory. The test statistics are given in the last column of Table 6. We set the trimming parameter of [Qu \(2011\)](#) to $\varepsilon_{Qu} = 0.02$ and the critical value equals 1.517 at the 1% significance level.

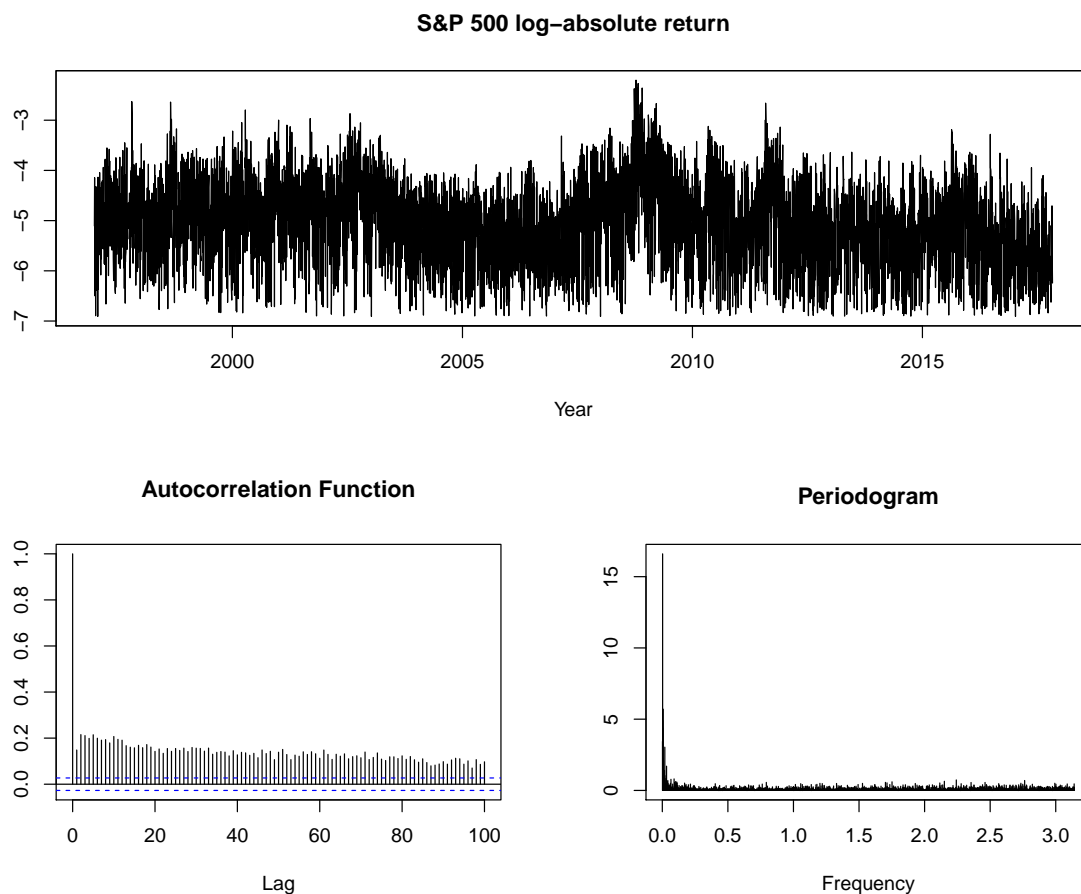


Figure 1. The log-absolute return series of the S&P 500 with the corresponding autocorrelation function and periodogram.

Table 6. \hat{d} estimates and the corresponding standard errors in brackets below for different bandwidths $m = \lfloor 1 + T^\delta \rfloor$ with * denoting significance at 1% level.

δ/\hat{d}	LW	GPH	HP	IAC	MCP	SMI	LPW	LWN	LPWN11	Qu Test
0.60	0.453 (0.038)	0.493 (0.049)	0.353 (0.038)	0.339 (0.086)	0.506 (0.055)	0.570 (0.100)	0.521 (0.057)	0.608 (0.039)	0.603 (0.138)	0.910
0.70	0.383 (0.025)	0.408 (0.032)	0.270 (0.025)	0.313 (0.056)	0.349 (0.041)	0.534 (0.061)	0.489 (0.037)	0.585 (0.026)	0.473 (0.103)	1.836 *
0.79	0.284 (0.017)	0.275 (0.022)	0.133 (0.017)	0.125 (0.038)	0.156 (0.033)	0.483 (0.040)	0.390 (0.025)	0.629 (0.017)	0.589 (0.062)	3.727 *

In general, the memory parameters decrease with an increased bandwidth, with the exception of \hat{d}_{LWN} and \hat{d}_{LPWN11} . This finding gives a first hint that the series might be contaminated since the long-memory component dominates for higher frequencies in a contaminated process (cf. Perron and Qu 2010). The estimators of Iacone (2010), Hou and Perron (2014) and McCloskey and Perron (2013) are downward biased compared to the respective non-modified standard method. Especially for the highest bandwidth ($m = \lfloor 1 + 5245^{0.79} \rfloor = 869$), the reduction of the memory is extremely strong with less than half of the memory value for \hat{d}_{HP} and \hat{d}_{IAC} compared to \hat{d}_{LW} . This downward pattern of the estimators for an increased bandwidth suggests the existence of low frequency contaminations in the log-absolute return series of the S&P 500. This interpretation is in line with the test results of Qu (2011).

The test rejects the null hypothesis for $\delta = 0.70$ and $\delta = 0.79$ at the 1% significance level and therefore suggests that the S&P 500 series follows a spurious long-memory process.

Contrary to this finding, the four other estimation methods present upward biased estimator of \hat{d} in a non-stationary region compared to the local Whittle or GPH estimator. The estimator of Andrews and Sun (2004) increases less than the other estimators and lies only for $\delta = 0.60$ in the non-stationary region. However, the estimators of Smith (2005), Frederiksen et al. (2012) and, especially, Hurvich et al. (2005) show larger memory parameters in the non-stationary region. The estimation results of Hurvich et al. (2005) and Frederiksen et al. (2012) are very similar. As expected from the simulation study, the estimator of Smith (2005) behaves rather unexpected since it is an estimator originally designed to capture low frequency contaminations instead of perturbation.

The estimators developed to capture perturbation are in the non-stationary region and combined with downward biased standard estimators and even more biased modified estimators for low frequency contamination a perturbed fractional process may be possible (cf. Wenger et al. 2017). Our findings are in line with those of McCloskey and Perron (2013) and Frederiksen et al. (2012), whereas the interpretation of the latter authors differs to our interpretation of a spurious long-memory process. We conclude that the analyzed time series follows a spurious long-memory process rather than a perturbed fractional process. In addition to the estimation results, this interpretation is supported by a formal test of Qu (2011) and the current literature of spurious long memory (cf. Lu and Perron 2010 or Qu and Perron 2013). Without the test results of Qu (2011), a perturbed fractional process would be an alternative conclusion.

In summary, the two standard methods indicate a stationary long-memory process, whereas the methods of Iacone (2010), Hou and Perron (2014) and McCloskey and Perron (2013) suggest a weakly stationary long-memory process, especially for the highest bandwidth. The four other modified estimation methods imply a non-stationary long-memory process.

5. Conclusions

In this paper, we analyze the performance of nine semiparametric estimation methods and compare them in a simulation study. The Monte Carlo study suggests to use either the estimator of Iacone (2010) or Hou and Perron (2014) in cases of low frequency contaminations. Although the estimator of McCloskey and Perron (2013) yields comparable estimation results, it suffers from a higher variance. The performance of the estimator of Smith (2005) in cases of high short-run dynamics or perturbation is comparable to the standard perturbation methods, even though the estimator has originally been developed for low frequency contamination. However, in the latter situation, the estimator is no alternative to its competitors. Furthermore, the estimators of Hurvich et al. (2005) and Frederiksen et al. (2012) yield almost identical findings, which are in most situations better than the results of Andrews and Sun (2004). Based on the Monte Carlo results, we recommend applying the low frequency estimators when its unclear whether a low frequency contamination or a perturbed fractional process is present. In our empirical example, we consider the log-absolute return series of the S&P 500. The downward pattern of Iacone (2010), Hou and Perron (2014) and McCloskey and Perron (2013) as the bandwidth increases combined with a rejection of the null hypothesis of a true long-memory process by the test of Qu (2011) suggests a spurious long-memory process. This paper presents the strengths and weaknesses of the different semiparametric estimation methods and provide some user-guidelines.

Supplementary Materials: The following are available online at www.mdpi.com/2225-1146/6/1/13/s1. The tables present the bias and root mean squared error (RMSE) results of nine different estimation methods.

Author Contributions: The authors contributed jointly to the paper.

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

References

- Andrews, Donald W. K., and Yixiao Sun. 2004. Adaptive local polynomial Whittle estimation of long-range dependence. *Econometrica* 72: 569–614.
- Arteche, Josu. 2004. Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics* 119: 131–54.
- Arteche, Josu. 2006. Semiparametric estimation in perturbed long memory series. *Computational Statistics & Data Analysis* 51: 2118–141.
- Byrd, Richard H., Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16: 1190–208.
- Diebold, Francis X., and Atsushi Inoue. 2001. Long memory and regime switching. *Journal of Econometrics* 105: 131–59.
- Doukhan, Paul, George Oppenheim, and Murad Taqqu. 2002. *Theory and Applications of Long-Range Dependence*. Berlin: Springer Science & Business Media.
- Faÿ, Gilles, Eric Moulines, François Roueff, and Murad S. Taqqu. 2009. Estimators of long-memory: Fourier versus wavelets. *Journal of Econometrics* 151: 159–77.
- Frederiksen, Per, Frank S. Nielsen, and Morten Ørregaard Nielsen. 2012. Local polynomial Whittle estimation of perturbed fractional processes. *Journal of Econometrics* 167: 426–47.
- Frederiksen, Per, and Morten Ørregaard Nielsen. 2008. Bias-reduced estimation of long-memory stochastic volatility. *Journal of Financial Econometrics* 6: 496–512.
- 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.
- Granger, Clive W. J., and Namwon Hyung. 2004. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance* 11: 399–421.
- 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.
- Haldrup, Niels, and Morten Ørregaard Nielsen. 2007. Estimation of fractional integration in the presence of data noise. *Computational Statistics & Data Analysis* 51: 3100–14.
- Hosking, Jonathan R. M. 1981. Fractional differencing. *Biometrika* 68: 165–76.
- Hou, Jie, and Pierre Perron. 2014. Modified local Whittle estimator for long memory processes in the presence of low frequency (and other) contaminations. *Journal of Econometrics* 182: 309–28.
- Hurvich, Clifford M., Eric Moulines, and Philippe Soulier. 2005. Estimating long memory in volatility. *Econometrica* 73: 1283–328.
- Hurvich, Clifford M., and Bonnie K. Ray. 2003. The local Whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics* 1: 445–70.
- Iacone, Fabrizio. 2010. Local Whittle estimation of the memory parameter in presence of deterministic components. *Journal of Time Series Analysis* 31: 37–49.
- Kuensch, Hans R. 1987. Statistical aspects of self-similar processes. Paper presented at First World Congress of the Bernoulli Society, Tashkent, Uzbekistan, September 8–14, pp. 67–74.
- Lu, Yang K., and Pierre Perron. 2010. Modeling and forecasting stock return volatility using a random level shift model. *Journal of Empirical Finance* 17: 138–56.
- McCloskey, Adam. 2013. Estimation of the long-memory stochastic volatility model parameters that is robust to level shifts and deterministic trends. *Journal of Time Series Analysis* 34: 285–301.
- McCloskey, Adam, and Jonathan B. Hill. 2017. Parameter estimation robust to low-frequency contamination. *Journal of Business & Economic Statistics* 35: 598–610.
- McCloskey, Adam, and Pierre Perron. 2013. Memory parameter estimation in the presence of level shifts and deterministic trends. *Econometric Theory* 29: 1196–237.
- Perron, Pierre, and Zhongjun Qu. 2010. Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business & Economic Statistics* 28: 275–90.
- Qu, Zhongjun. 2011. A test against spurious long memory. *Journal of Business & Economic Statistics* 29: 423–38.
- Qu, Zhongjun, and Pierre Perron. 2013. A stochastic volatility model with random level shifts and its applications to S&P 500 and NASDAQ return indices. *The Econometrics Journal* 16: 309–39.

- Robinson, Peter M. 1995. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics* 23: 1048–72.
- Smith, Aaron. 2005. Level shifts and the illusion of long memory in economic time series. *Journal of Business & Economic Statistics* 23: 321–35.
- Sun, Yixiao, and Peter C. B. Phillips. 2003. Nonlinear log-periodogram regression for perturbed fractional processes. *Journal of Econometrics* 115: 355–89.
- Wenger, Kai, Christian Leschinski, and Philipp Sibbertsen. 2017. The Memory of Volatility. Unpublished Manuscript, Department of Economics, University of Hannover, Hanover, Germany.
- Xu, Jiawen, and Pierre Perron. 2014. Forecasting return volatility: Level shifts with varying jump probability and mean reversion. *International Journal of Forecasting* 30: 449–63.
- Yamamoto, Yohei, and Pierre Perron. 2013. Estimating and testing multiple structural changes in linear models using band spectral regressions. *The Econometrics Journal* 16: 400–29.



© 2018 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/>).