CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-29

Local polynomial Whittle estimation of perturbed fractional processes

Per Frederiksen, Frank S. Nielsen and Morten Ørregaard Nielsen

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Aarhus School of Business
University of Aarhus
Handelshøjskolen
Aarhus Universitet

UNIVERSITY OF COPENHAGEN

Local polynomial Whittle estimation of perturbed fractional processes*

Per Frederiksen
Equity Trading & Derivatives
Nordea Markets

Frank S. Nielsen†
University of Aarhus and CREATES

Morten Ørregaard Nielsen
Cornell University and CREATES

June 9, 2008

Abstract

We propose a semiparametric local polynomial Whittle with noise (LPWN) estimator of the memory parameter in long memory time series perturbed by a noise term which may be serially correlated. The estimator approximates the spectrum of the perturbation as well as that of the short-memory component of the signal by two separate polynomials. Including these polynomials we obtain a reduction in the order of magnitude of the bias, but also inflate the asymptotic variance of the long memory estimate by a multiplicative constant. We show that the estimator is consistent for $d \in (0, 1)$, asymptotically normal for $d \in (0, 3/4)$, and if the spectral density is infinitely smooth near frequency zero, the rate of convergence can become arbitrarily close to the parametric rate, $\sqrt{n}$. A Monte Carlo study reveals that the LPWN estimator performs well in the presence of a serially correlated perturbation term. Furthermore, an empirical investigation of the 30 DJIA stocks shows that this estimator indicates stronger persistence in volatility than the standard local Whittle estimator.

JEL Classifications: C22.

Keywords: Bias reduction, local Whittle, long memory, perturbed fractional process, semiparametric estimation, stochastic volatility.

*We are grateful to Torben G. Andersen, Niels Haldrup, Esben Høg, Asger Lunde, and the participants at SETA 2008 for valuable suggestions and comments. This work was partly done while P. Frederiksen was visiting Northwestern University, F. S. Nielsen was visiting Cornell University, and M. Ø. Nielsen was visiting Queen's University and the University of Aarhus; their hospitality is gratefully acknowledged. We are grateful for financial support from the Danish Social Sciences Research Council (grant no. FSE 275-05-0220) and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation).

†Please address correspondence to: Frank S. Nielsen, School of Economics and Management, University of Aarhus, Universitetsparken building 1322, 8000 Aarhus, Denmark; phone: +45 8942 5419; e-mail: fnielsen@creates.au.dk

1 Introduction

We are interested in estimation of the memory parameter in a so-called perturbed fractional process,

$$z_t = y_t + w_t, \tag{1}$$

i.e. a signal-plus-noise model where the signal process y_t is a long memory process with memory parameter d which is perturbed by the additive noise term w_t. These processes have found extensive use in modeling the long memory characteristics of many observed time series. In particular, they are a version of the random walk plus noise or local level unobserved components model, e.g. Harvey (1989), except the signal is a long memory process rather than a random walk.

Another motivation for the perturbed fractional process is the version of the long memory stochastic volatility (LMSV) model for financial returns proposed by Bollerslev & Jubinski (1999),

$$r_t = \kappa\sqrt{e^{y_t+x_t}}u_t, \tag{2}$$

where r_t denotes the return, y_t is the (long memory component of) log-volatility of the returns, x_t is a short-memory process, and y_t, x_t, and u_t are independent to satisfy the requirement that $E(r_t) = 0$. This generalizes the usual LMSV model introduced by Breidt, Crato & de Lima (1998) and Harvey (1998),

$$r_t = \kappa\sqrt{e^{y_t}}u_t, \tag{3}$$

by arguing that allowing for different short-lived news impacts, while imposing a common long memory component, may provide a better characterization of the joint volume-volatility relationship in the context of the Mixture of Distributions Hypothesis, which asserts that stock returns and trading volumes are jointly dependent on the same underlying latent information arrival process. The formulation in (2) allows the volatility to be affected by both long and short-lived news impacts, which is also consistent with the findings of Liesenfeld (2001). It therefore seems natural that an estimator of the memory in $\log r_t^2$ should be able to incorporate both (2) and (3).

The LMSV models (2) and (3) imply that a logarithmic transformation of the squared returns series $\log r_t^2$ becomes a long memory signal-plus-noise process (1) where the signal y_t corresponds to (the long memory component of) the log-volatility of the original returns series and w_t is an additive noise term. In the context of the LMSV model (3), w_t is usually assumed to be *i.i.d.*, but to allow for short-memory persistence in w_t as implied by (2) we will not make that restriction here. In general, when w_t is not assumed to be *i.i.d.*, z_t is referred to as a perturbed fractional process.[1] For reviews of fractionally integrated processes and some applications, see Baillie (1996), Henry & Zaffaroni (2003), or Robinson (2003). In particular, long memory in volatility has received

[1] In the following we use the terms "long memory process" and "fractionally integrated process" or just "fractional process" synonymously, although strictly speaking a fractional process is just a particular form of a long memory process.

considerable interest recently.[2]

If we assume that the log-volatility process $\{y_t\}$ and the noise process $\{w_t\}$ are independent, the spectral density of z_t can be written as

$$f_z(\lambda) = \lambda^{-2d}\phi_y(\lambda) + \phi_w(\lambda) = \lambda^{-2d}G\left(\frac{\phi_y(\lambda)}{\phi_y(0)} + \lambda^{2d}\frac{\phi_w(\lambda)}{\phi_y(0)}\right), \tag{4}$$

where $f_y(\lambda) = \lambda^{-2d}\phi_y(\lambda)$ is the spectrum of the signal y_t, $\phi_w(\lambda)$ is the spectrum of the noise term w_t, and d is the degree of long memory in y_t (or equivalently in z_t).

The assumption of independence between the processes $\{y_t\}$ and $\{w_t\}$ rules out the so-called leverage effect. This assumption is common in the so-called random walk plus noise unobserved components models, and has also been imposed by Breidt et al. (1998), Deo & Hurvich (2001), and Arteche (2004), among others, in the LMSV model. To accommodate the leverage effect, we could allow contemporaneous correlation, while the return process remains a martingale difference sequence by replacing y_t with y_{t-1} in (2). An additional assumption of distributional symmetry around $(0,0)$ would imply that the spectral density decomposition in (4) holds, see Hurvich, Moulines & Soulier (2005). Alternatively, the model could be modified along the lines of model (P2) of Hurvich et al. (2005).

In semiparametric spectral estimation of long memory models, the spectrum (4) is typically approximated using the periodogram of the data near the zero frequency, i.e. for frequencies up to $\lambda_m = 2\pi m/n$ only, where n is the sample size and m is a user-chosen bandwidth number, see sections 2 and 3 below, which tends to infinity slower than n such that $\lambda_m \to 0$. Although the popular log-periodogram regression (LPR) estimator of Geweke & Porter-Hudak (1983) and Robinson (1995*b*) and the local Whittle (LW) estimator of Künsch (1987) and Robinson (1995*a*) both preserve consistency and asymptotic normality when applied to perturbed fractional processes, as shown recently by Deo & Hurvich (2001) and Arteche (2004), these estimators can be severely biased since they do not take the perturbation into account. Indeed, for non-perturbed processes (where $\phi_w(\lambda) = 0$) the bias of the standard semiparametric frequency domain estimators is of order $O(\lambda_m^2)$, whereas the leading bias term when $\phi_w(\lambda) \neq 0$ is of order $O(\lambda_m^{2d})$. As shown in Deo & Hurvich (2001) and Arteche (2004), this bias is typically negative and can be very large (note that $d < 1$). Therefore, estimating long memory in perturbed time series can be a challenging task, and calls for an estimator which explicitly accounts for the perturbation.

Sun & Phillips (2003), Hurvich & Ray (2003), and Hurvich et al. (2005) have proposed such estimators with $\phi_y(\lambda)$ and $\phi_w(\lambda)$ approximated by constants as $\lambda \to 0$, see section 2 below. On the other hand, we propose an estimator where we allow both the spectrum of the perturbation and the spectrum of the short-memory component of the signal, i.e. $\phi_w(\lambda)$ and $\phi_y(\lambda)$, to be approximated by polynomials $h_w(\boldsymbol{\theta}_w, \lambda)$ and $h_y(\boldsymbol{\theta}_y, \lambda)$ of (finite and even) orders $2R_w$ and $2R_y$ near the zero

[2] See, e.g., Ding, Granger & Engle (1993), Baillie, Bollerslev & Mikkelsen (1996), Comte & Renault (1998), Ray & Tsay (2000), Andersen, Bollerslev, Diebold & Ebens (2001), Andersen, Bollerslev, Diebold & Labys (2001, 2003), Wright (2002), Hurvich & Ray (2003), and Arteche (2004) among others.

frequency, instead of constants, thereby obtaining a bias reduction depending on the smoothness of $\phi_w(\lambda)$ and $\phi_y(\lambda)$ near the origin. The approach taken here in modeling the short-run dynamics by a polynomial was introduced by Andrews & Sun (2004) for non-perturbed processes, but is novel in the context of perturbed fractional processes. To maintain generality, $\phi_w(\lambda)$ and $\phi_y(\lambda)$ are only characterized by regularity conditions near frequency zero instead of imposing specific functional forms.

The LMSV model (3) often assumes that the noise term is *i.i.d.* in which case $\phi_w(\lambda) = \sigma_w^2/(2\pi)$ is a constant. This case is of independent interest and is considered in simulations and in an empirical study in Frederiksen & Nielsen (2007). In that paper $\phi_y(\lambda)$ is approximated by a polynomial and $\phi_w(\lambda)$ by a constant as $\lambda \to 0$ thus focusing on exactly the LMSV model (3). However, the theory for their estimator is developed in the present paper.

Thus, to allow serial dependence in the noise as in (2) above we include both polynomials, $h_y(\boldsymbol{\theta}_y, \lambda)$ and $h_w(\boldsymbol{\theta}_w, \lambda)$. Furthermore, empirical studies have typically found that the noise term has much higher (long-run) variance than the short-memory component of the signal. Indeed, Breidt et al. (1998) and Hurvich & Ray (2003) find that the noise term may be as much as 10 or 20 times as variable as the short-memory component of the signal. Thus, careful modeling of the noise term is important and this consideration has lead us to approximate the spectrum of the noise term by a polynomial instead of a constant as $\lambda \to 0$.

Our results show that introducing $h_y(\boldsymbol{\theta}_y, \lambda)$ and $h_w(\boldsymbol{\theta}_w, \lambda)$ inflates the asymptotic variance of the long memory estimator, $\hat{d}$, by a multiplicative constant which depends on the true long memory parameter, d. However, the inflation decreases when d increases, and we obtain a reduction in the order of magnitude of the bias if $\phi(\lambda)$ is sufficiently smooth near frequency zero. We show that the estimator is consistent for $d \in (0, 1)$, asymptotically normal for $d \in (0, 3/4)$, and if $\phi(\lambda)$ is infinitely smooth near frequency zero, the rate of convergence can become arbitrary close to the parametric rate, $n^{1/2}$. This constitutes a rate of convergence improvement relative to Sun & Phillips (2003), Hurvich & Ray (2003), and Hurvich et al. (2005) who are only able to obtain a semiparametric rate of convergence $m^{1/2}$, which is much slower than the parametric rate due to the minimal requirement that $m/n \to 0$.

We present the results of a Monte Carlo study which shows the usefulness of the proposed LPWN estimator. Compared to standard estimators, such as Hurvich & Ray's (2003) local Whittle with noise (LWN) estimator, the LPWN estimator is able to achieve considerable bias reductions in practice, especially in cases with short-run dynamics in both the signal and noise components. We also include an empirical application to the 30 DJIA stocks where the LPWN estimator indicates stronger persistence in volatility than the standard estimators, and for most of the stocks produce estimates of d in the nonstationary region.

The remainder of the paper is organized as follows. In the next section we discuss semiparametric spectral estimation of long memory for perturbed processes and formally define the proposed local Whittle estimator. In section 3 we establish consistency and asymptotic normality of the estimator.

Section 4 investigates the finite sample performance in simulations, and section 5 presents an empirical study of daily log-squared returns series of the 30 DJIA stocks. Section 6 concludes. The proofs of our theorems are gathered in the appendix.

2 Local Whittle estimation of perturbed fractional processes

Semiparametric frequency-domain estimators are essentially based on the local approximation

$$f_z(\lambda) \sim G\lambda^{-2d} \text{ as } \lambda \to 0, \tag{5}$$

where G is a constant and the symbol "$\sim$" means that the ratio of the left and right hand sides tends to one in the limit. Thus, the estimators enjoy robustness to short-run dynamics, since they use only information from periodogram ordinates in the vicinity of the origin.

The local Whittle (LW) estimation method by Künsch (1987) and Robinson (1995*a*) has become popular because of its likelihood interpretation, nice asymptotic properties, and mild assumptions. It is defined as the minimizer of the (negative) local Whittle likelihood function

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

where $m = m(n)$ is a bandwidth number which tends to infinity as $n \to \infty$ but at a slower rate than n, $\lambda_j = 2\pi j/n$ are the Fourier frequencies, and $I_z(\lambda) = (2\pi n)^{-1}|\sum_{t=1}^{n} z_t e^{it\lambda}|^2$ is the periodogram of z_t. Note that the estimator is invariant to a non-zero mean since $j = 0$ is left out of the minimization. Concentrating (6) with respect to G, the estimator of d is

$$\hat{d}_{LW} = \arg\min_{d}\left[\log\hat{G}(d) - 2d\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\right], \quad \hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_z(\lambda_j).$$

It was shown by Robinson (1995*a*) that

$$\sqrt{m}(\hat{d}_{LW} - d) \xrightarrow{d} N(0, 1/4), \tag{7}$$

and later by Velasco (1999) that the range of consistency is $d \in (-1/2, 1]$ and the range of asymptotic normality is $d \in (-1/2, 3/4)$.

To reduce the asymptotic bias of the standard LW estimator, Andrews & Sun (2004) have suggested to replace the constant, $\log G$, in (6) by the polynomial $\theta_0 - \sum_{r=1}^{R}\theta_r\lambda_j^{2r}$. That is, by modeling the logarithm of the spectral density of the short-run component by a polynomial instead of a constant in the vicinity of the origin. This leads to the following (negative) likelihood function,

$$Q(G,d,\boldsymbol{\theta}) = \frac{1}{m}\sum_{j=1}^{m}\left[\log\left(\lambda_j^{-2d}G\exp\left(-\sum_{r=1}^{R}\theta_r\lambda_j^{2r}\right)\right) + \frac{I_z(\lambda_j)}{\lambda_j^{-2d}G\exp\left(-\sum_{r=1}^{R}\theta_r\lambda_j^{2r}\right)}\right],$$

such that

$$\begin{aligned}
(\hat{d}_{LPW}, \hat{\boldsymbol{\theta}}) &= \underset{d\in(-1/2,1/2),\boldsymbol{\theta}\in\Theta}{\arg\min} \left[\log \hat{G}(d,\boldsymbol{\theta}) - 2d\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j - \frac{1}{m}\sum_{j=1}^{m}\sum_{r=1}^{R}\theta_r\lambda_j^{2r} \right], \\
\hat{G}(d,\boldsymbol{\theta}) &= \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}\exp\left(\sum_{r=1}^{R}\theta_r\lambda_j^{2r}\right) I_z(\lambda_j),
\end{aligned}$$

where Θ is a compact and convex set in $\mathbb{R}^R$. As shown by Andrews & Sun (2004), this method does, however, increase the asymptotic variance of $\hat{d}$ in (7) by a multiplicative constant.

For non-perturbed fractional processes, the asymptotic bias of $\hat{d}_{LW}$ and $\hat{d}_{LPW}$ is of order $O(\lambda_m^2)$ and $O(\lambda_m^{\min\{s,2+2R\}})$, respectively, where s is a measure of the smoothness of the spectral density near frequency zero, see below. However, for perturbed fractional processes the bias is of order $O(\lambda_m^{2d})$ and, as shown by e.g. Hurvich & Ray (2003) and Arteche (2004), this bias is typically negative and can be very severe.

For perturbed fractional processes we have the spectral representation (4) rather than (5). There are two main consequences: first, the extra additive term in (4) needs to be taken into account to avoid serious asymptotic bias as mentioned above, and second the rate of convergence of the estimators is reduced if the extra term is not modeled. The latter follows because the choice of bandwidth parameter is severely constrained for perturbed fractional processes when the perturbation term in (4) is not modeled. Thus, for non-perturbed processes the bandwidth requirement is typically $m = o(n^{4/5})$, whereas for perturbed processes it is $m = o(n^{2d/(1+2d)})$ (apart from logarithmic terms). Since $d \leq 1$ and the estimator is $\sqrt{m}-$consistent this is a serious constraint.

To allow for (moderate) nonstationarity in volatility we generalize (1) as

$$z_t = \begin{cases} y_t + w_t & \text{if } d \in (0, 1/2), \\ \sum_{s=1}^{t} x_s + w_t & \text{if } d \in [1/2, 1), \end{cases} \tag{8}$$

where, if $d \in [1/2, 1)$, x_t has spectrum of the form $f_x(\lambda) = \lambda^{-2d_x}\phi_x(\lambda)$ with memory parameter $d_x = d - 1$. Defining $y_t = \sum_{s=1}^{t} x_s$ if $d \in [1/2, 1)$, this approach allows $z_t = y_t + w_t$ to possibly be nonstationary with memory parameter $d \in (0, 1)$. Velasco (1999), Hurvich & Ray (2003), and Hurvich et al. (2005) also assume this type of process. Since $\{\sum_{s=1}^{t} x_s\}$ is nonstationary[3] z_t does not have a spectral density if $d \in [1/2, 1)$ but it has a pseudo spectral density, see e.g. Hurvich & Ray (1995) and Velasco (1999). Thus, we may define

$$\begin{aligned}
f_z(\lambda) &= \begin{cases} f_y(\lambda) + f_w(\lambda) & \text{if } d \in (0, 1/2), \\ \left|1 - e^{i\lambda}\right|^{-2} f_x(\lambda) + f_w(\lambda) & \text{if } d \in [1/2, 1), \end{cases} \\
&= \lambda^{-2d} G\left(\frac{\phi_y(\lambda)}{\phi_y(0)} + \lambda^{2d}\frac{\phi_w(\lambda)}{\phi_y(0)}\right),
\end{aligned} \tag{9}$$

[3] In the nonstationary case, $\{\sum_{s=1}^{t} x_s\}$ is a type I fractional process in the terminology of Marinucci & Robinson (1999).

where we maintain the assumption of independence between $\{y_t\}$ and $\{w_t\}$.

Taking (9) into account we propose to approximate (4) locally near the zero frequency by[4]

$$g(\lambda) = \lambda^{-2d} G\left(1 + h_y(\boldsymbol{\theta}_y, \lambda) + \lambda^{2d} h_w(\boldsymbol{\theta}_w, \lambda)\right), \tag{10}$$

where $h_y(\boldsymbol{\theta}_y, \lambda) = \sum_{r=1}^{R_y} \theta_{y,r} \lambda^{2r}$, $h_w(\boldsymbol{\theta}_w, \lambda) = \sum_{r=0}^{R_w} \theta_{w,r} \lambda^{2r}$. If $R_y = 0$ we set $h_y(\boldsymbol{\theta}_y, \lambda) = 0$. Defining also the polynomial $h(d, \boldsymbol{\theta}, \lambda) = h_y(\boldsymbol{\theta}_y, \lambda) + \lambda^{2d} h_w(\boldsymbol{\theta}_w, \lambda)$ with $\boldsymbol{\theta} = (\boldsymbol{\theta}_y', \boldsymbol{\theta}_w')'$ this yields the (concentrated) likelihood

$$Q(d, \boldsymbol{\theta}) = \log \hat{G}(d, \boldsymbol{\theta}) + \frac{1}{m} \sum_{j=1}^{m} \log\left(\lambda_j^{-2d}\left(1 + h(d, \boldsymbol{\theta}, \lambda_j)\right)\right), \tag{11}$$

$$\hat{G}(d, \boldsymbol{\theta}) = \frac{1}{m} \sum_{j=1}^{m} \frac{\lambda_j^{2d} I_z(\lambda_j)}{1 + h(d, \boldsymbol{\theta}, \lambda_j)}. \tag{12}$$

Thus, we propose to minimize (11) over the admissible set $D \times \Theta$,

$$(\hat{d}, \hat{\boldsymbol{\theta}}) = \underset{(d,\theta) \in D \times \Theta}{\arg\min}\, Q(d, \boldsymbol{\theta}),$$

where Θ is a compact and convex set in $\mathbb{R}^{R+1}$, $R = R_y + R_w$, and $D = [d_1, d_2]$ with $0 < d_1 < d_2 < 1$. We call this estimator the local polynomial Whittle with noise (LPWN) estimator.

Note that $h(\boldsymbol{\theta}, \lambda) = 0$ is the standard local Whittle specification in (6), which does not explicitly account for the perturbation. For $R_y = R_w = 0$ we get $h(\boldsymbol{\theta}, \lambda) = \theta$, where $\phi_y(\lambda)$ and $\phi_w(\lambda)$ in (4) are both modeled locally by constants. This is the local Whittle with noise (LWN) estimator of Hurvich & Ray (2003) and Hurvich et al. (2005) (parameterization (P1)). Thus, our model parameterization includes the standard LW estimator and the LWN estimator as special cases. Furthermore, the model with $R_w = 0$, where the noise is modeled by a constant near the zero frequency, is analyzed empirically and in simulations by Frederiksen & Nielsen (2007), using the asymptotic theory provided in this paper.

3 Asymptotic properties

In this section we first introduce the assumptions needed to establish consistency and asymptotic normality of the proposed estimator for the perturbed fractional process, and consequently we present the main results in two theorems. In the following, true values of the parameters are denoted by subscript zero and $\lfloor x \rfloor$ denotes the integer part of a real number x. We also define a function $\phi(\lambda)$ to be smooth of order s at $\lambda = 0$ if, in a neighborhood of $\lambda = 0$, $\phi(\lambda)$ is $\lfloor s \rfloor$ times continuously differentiable with $\lfloor s \rfloor$ −derivative, $\phi^{(\lfloor s \rfloor)}$, satisfying $|\phi^{(\lfloor s \rfloor)}(\lambda) - \phi^{(\lfloor s \rfloor)}(0)| \le C |\lambda|^{s - \lfloor s \rfloor}$ for some constant $C < \infty$. To simplify the presentation, we list only one set of assumptions even though these could be relaxed somewhat for the consistency proof, see e.g. Hurvich et al. (2005).

[4]Note that $\phi_y(\lambda)$ and $\phi_w(\lambda)$ are symmetric around $\lambda = 0$ and are therefore approximated by even polynomials.

A1 The noise process $\{w_t\}$ is independent of the signal process $\{y_t\}$.

A2 The spectral density of z_t is $f_z(\lambda) = \lambda^{-2d_0} G_0 \frac{\phi_y(\lambda)}{\phi_y(0)} + \phi_w(\lambda)$, where $\phi_y(\lambda)$ and $\phi_w(\lambda)$ are real, even, positive, continuous functions on $[-\pi, \pi)$ and $d_0 \in D = [d_1, d_2]$ with $0 < d_1 < d_2 < 1$.

A3 The functions $\phi_y(\lambda)$ and $\phi_w(\lambda)$ are smooth of orders s_y and s_w at $\lambda = 0$, where $s_y > 2R_y$, $s_w > 2R_w$, and $s_y, s_w \geq 1$.

Assumption A1 is the independence assumption used above to write the spectral density of z_t as the sum of the (pseudo) spectral densities of y_t and w_t. Assumption A3 is a smoothness condition on the functions $\phi_y(\lambda)$ and $\phi_w(\lambda)$ similar to that applied by Andrews & Sun (2004). Note that Assumption A3 holds for all $s_y < \infty$ when, e.g., y_t is a finite order ARFIMA process, and for all $s_w < \infty$ when, e.g., w_t is a finite order ARMA process. Under Assumption A3 we establish the following Taylor series expansions of $\phi_y(\lambda)$ and $\phi_w(\lambda)$ around $\lambda = 0$ (recall that odd order derivatives of even functions are zero at frequency zero),

$$\frac{\phi_y(\lambda)}{\phi_y(0)} = 1 + \sum_{r=1}^{\lfloor s_y/2 \rfloor} \theta_{y,r} \lambda^{2r} + O(\lambda^{s_y}) = 1 + h_y(\boldsymbol{\theta}_y, \lambda) + O(\lambda^{\min\{s_y, 2+2R_y\}}) \text{ as } \lambda \to 0,$$

and

$$\frac{\phi_w(\lambda)}{\phi_y(0)} = \frac{\phi_w(0)}{\phi_y(0)} + \sum_{r=1}^{\lfloor s_w/2 \rfloor} \theta_{w,r} \lambda^{2r} + O(\lambda^{s_w}) = h_w(\boldsymbol{\theta}_w, \lambda) + O(\lambda^{\min\{s_w, 2+2R_w\}}) \text{ as } \lambda \to 0,$$

where $\theta_{y,r} = \frac{1}{(2r)!\phi_y(0)} \frac{\partial^{2r}}{\partial \lambda^{2r}} \phi_y(\lambda)\big|_{\lambda=0}$ and $\theta_{w,r+1} = \frac{1}{(2r)!\phi_y(0)} \frac{\partial^{2r}}{\partial \lambda^{2r}} \phi_w(\lambda)\big|_{\lambda=0}$. Hence, the approximation (10) to (9) is

$$\begin{aligned}
\log(f_z(\lambda)/g(\lambda)) &= \log\left(\frac{\phi_y(\lambda)}{\phi_y(0)} + \lambda^{2d} \frac{\phi_w(\lambda)}{\phi_y(0)}\right) - \log(1 + h(d, \boldsymbol{\theta}, \lambda)) \\
&= \log\left(1 + \frac{O(\lambda^{\min\{s_y, 2+2R_y\}}) + \lambda^{2d} O(\lambda^{\min\{s_w, 2+2R_w\}})}{1 + h(d, \boldsymbol{\theta}, \lambda)}\right) \text{ as } \lambda \to 0, \\
\frac{f_z(\lambda)}{g(\lambda)} &= 1 + O(\lambda^{\min\{s_y, 2+2R_y\}}) + \lambda^{2d} O(\lambda^{\min\{s_w, 2+2R_w\}}) \text{ as } \lambda \to 0, \qquad (13)
\end{aligned}$$

and the true values of G and $\boldsymbol{\theta}$ are $G_0 = \phi_y(0)$ and $\boldsymbol{\theta}_0 = (\theta_{0,1}, ..., \theta_{0,R+1})'$, where

$$\begin{aligned}
\theta_{0,r} &= \frac{1}{(2r)!\phi_y(0)} \frac{\partial^{2r}}{\partial \lambda^{2r}} \phi_y(\lambda)\big|_{\lambda=0}, r = 1, \ldots, R_y, \\
\theta_{0,R_y+r+1} &= \frac{1}{(2r)!\phi_y(0)} \frac{\partial^{2r}}{\partial \lambda^{2r}} \phi_w(\lambda)\big|_{\lambda=0}, r = 0, \ldots, R_w.
\end{aligned}$$

A4 (a) The signal y_t has zero mean and admits an infinite order moving average representation $y_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$ (stationary case) or $\Delta y_t = x_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$ (nonstationary case), where $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ and ε_t satisfies, for all t, $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = 1$, $E(\varepsilon_t^3 | \mathcal{F}_{t-1}) = \mu_3 < \infty$, and $E(\varepsilon_t^4 | \mathcal{F}_{t-1}) = \mu_4 < \infty$ almost surely, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\varepsilon_s, s < t\}$.

(b) There exists a random variable ε with $E(\varepsilon^2) < \infty$ such that for all $\tau > 0$ and some $K > 0$, $P(|\varepsilon_t| > \tau) < KP(|\varepsilon| > \tau)$.

(c) In a neighborhood of the origin, $\frac{\partial}{\partial \lambda}\alpha(\lambda) = O(|\alpha(\lambda)|/\lambda)$ as $\lambda \to 0$, where $\alpha(\lambda) = \sum_{k=0}^{\infty} \alpha_k e^{ik\lambda}$.

A5 (a) The noise w_t has zero mean and admits an infinite order moving average representation $w_t = \sum_{j=0}^{\infty} \beta_j \eta_{t-j}$, where $\sum_{j=0}^{\infty} \beta_j^2 < \infty$ and η_t satisfies, for all t, $E(\eta_t | \mathcal{F}_{t-1}) = 0$, $E\left(\eta_t^2 \middle| \mathcal{F}_{t-1}\right) = 1$, $E\left(\eta_t^3 \middle| \mathcal{F}_{t-1}\right) = \mu_3 < \infty$, and $E\left(\eta_t^4 \middle| \mathcal{F}_{t-1}\right) = \mu_4 < \infty$ almost surely, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\eta_s, s < t\}$.

(b) There exists a random variable ε with $E(\varepsilon^2) < \infty$ such that for all $\tau > 0$ and some $K > 0$, $P(|\eta_t| > \tau) < KP(|\varepsilon| > \tau)$.

(c) In a neighborhood of the origin, $\frac{\partial}{\partial \lambda}\beta(\lambda) = O(|\beta(\lambda)|/\lambda)$ as $\lambda \to 0$, where $\beta(\lambda) = \sum_{k=0}^{\infty} \beta_k e^{ik\lambda}$.

Since our estimator is a function of the periodogram at nonzero frequencies only, we assume without loss of generality[5] that the signal process y_t has zero mean. Importantly, Assumptions A4 and A5 allow for non-Gaussian processes. Note that Assumptions A1-A4 plus the assumption that w_t is white noise with finite fourth moment imply the assumptions needed on y_t and w_t to prove consistency and asymptotic normality (if, in addition, $d_2 < 3/4$) of the LWN estimator of Hurvich & Ray (2003). It follows from Theorems 1 and 2 below that their results for the LWN estimator are also valid for our more general assumptions on w_t in Assumption A5.

A6 Θ is a compact and convex subset of $\mathbb{R}^{R+1}$ and $\boldsymbol{\theta}_0$ lies in the interior of Θ.

We are now ready to prove consistency of our estimator. As mentioned above, some of our assumptions could be relaxed somewhat to prove this theorem, but we have preferred to list only one set of assumptions which will be used also for the proof of asymptotic normality below. The proofs of both theorems are given in the appendix.

Theorem 1 *If Assumptions A1-A6 hold and the bandwidth $m = m(n)$ is such that*

$$\frac{1}{m} + \frac{m}{n} \to 0, \tag{14}$$

then $\hat{d} - d_0 = o_P((\log n)^{-5})$.

Note that the theorem proves consistency only for the estimator of the memory parameter (at logarithmic rate). There is no proof of consistency for the estimators of the polynomial parameters in $\boldsymbol{\theta}$. The strategy of proof in Hurvich et al. (2005) would require next a separate proof of consistency of the polynomial parameters, however, we follow instead the method of proof in Andrews & Sun (2004) which does not require an intermediate result on the consistency of $\hat{\boldsymbol{\theta}}$. Thus, we give next the joint asymptotic normality of $\hat{d}$ and $\hat{\boldsymbol{\theta}}$.

[5] In the nonstationary case the zero mean assumption implies that z_t is free of linear trends which does entail a loss of generality in that case.

Theorem 2 *Let Assumptions A1-A6 hold with d_0 in the interior of $D = [d_1, d_2]$, $0 < d_1 < d_2 < 3/4$, and suppose the bandwidth $m = m(n)$ is such that*

$$\frac{m^{1+4R_y}}{n^{4R_y}} + \frac{m^{1+4(d_0+R_w)}}{n^{4(d_0+R_w)}} \to \infty \text{ and } \frac{m^{2\varphi_y+1}}{n^{2\varphi_y}} + \frac{m^{2\varphi_w+4d_0+1}}{n^{2\varphi_w+4d_0}} \to 0, \tag{15}$$

where $\varphi_a = \min\{s_a, 2+2R_a\}, a = y, w$. Then $\hat{d}$ and $\hat{\boldsymbol{\theta}}$ are both consistent and

$$\mathbf{B}_n \begin{pmatrix} \hat{d} - d_0 \\ \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Omega}_{R_y,R_w}^{-1}), \quad \boldsymbol{\Omega}_{R_y,R_w} = \begin{pmatrix} 4 & \boldsymbol{\mu}'_{R_y} & \boldsymbol{\nu}'_{R_w} \\ \boldsymbol{\mu}_{R_y} & \boldsymbol{\Gamma}_{R_y} & \boldsymbol{\psi}'_{R_y,R_w} \\ \boldsymbol{\nu}_{R_w} & \boldsymbol{\psi}_{R_w,R_y} & \boldsymbol{\Psi}_{R_w} \end{pmatrix},$$

where $\mathbf{B}_n = \mathbf{B}_n(d_0)$ is the $(R+2) \times (R+2)$ deterministic diagonal matrix with diagonal elements

$$(\mathbf{B}_n)_{11} = \sqrt{m}, \; (\mathbf{B}_n)_{k+1,k+1} = \sqrt{m}\lambda_m^{2k} \text{ for } k = 1, \ldots, R_y,$$
$$\text{and } (\mathbf{B}_n)_{k+R_y+2,k+R_y+2} = \sqrt{m}\lambda_m^{2d_0+2k} \text{ for } k = 0, \ldots, R_w,$$

$\boldsymbol{\mu}_{R_y}$ and $\boldsymbol{\nu}_{R_w} = \boldsymbol{\nu}_{R_w}(d_0)$ are the vectors

$$(\boldsymbol{\mu}_{R_y})_k = \frac{-4k}{(1+2k)^2} \text{ for } k = 1, \ldots, R_y \text{ and } (\boldsymbol{\nu}_{R_w})_{k+1} = \frac{-4(d_0+k)}{(1+2d_0+2k)^2} \text{ for } k = 0, \ldots, R_w,$$

$\boldsymbol{\Gamma}_{R_y}$ and $\boldsymbol{\Psi}_{R_w} = \boldsymbol{\Psi}_{R_w}(d_0)$ are the $R_y \times R_y$ and $(R_w+1) \times (R_w+1)$ matrices

$$\begin{aligned} (\boldsymbol{\Gamma}_{R_y})_{ik} &= \frac{4ik}{(1+2i+2k)(1+2i)(1+2k)} \text{ for } i, k = 1, \ldots, R_y, \\ (\boldsymbol{\Psi}_{R_w})_{i+1,k+1} &= \frac{4(d_0+i)(d_0+k)}{(1+2i+2k+4d_0)(1+2i+2d_0)(1+2k+2d_0)} \text{ for } i, k = 0, \ldots, R_w, \end{aligned}$$

and $\boldsymbol{\psi}_{R_w,R_y} = \boldsymbol{\psi}_{R_w,R_y}(d_0)$ is the $(R_w+1) \times R_y$ matrix

$$(\boldsymbol{\psi}_{R_w,R_y})_{i+1,k} = \frac{4k(d_0+i)}{(1+2d_0+2k+2i)(1+2d_0+2i)(1+2k)} \text{ for } i = 0, \ldots R_w, k = 1, \ldots, R_y.$$

If $R_y = R_w = 0$ define $\boldsymbol{\Omega}_{0,0} = \begin{pmatrix} 4 & \nu'_0 \\ \nu_0 & \Psi_0 \end{pmatrix}$.

First of all, we note that by setting $R_y = R_w = 0$ we obtain as a special case the results for the LWN estimator of Hurvich & Ray (2003). Secondly, the leading $(R_y+1) \times (R_y+1)$ submatrix of $\boldsymbol{\Omega}_{R_y,R_w}$ is the same as that obtained by Andrews & Sun (2004). Third, we note that the asymptotic variance of $\sqrt{m}(\hat{d} - d_0)$ is free of the polynomial parameters $\boldsymbol{\theta}_0$, but it depends on d_0. Moreover, the use of the polynomials $h_y(\boldsymbol{\theta}_y, \lambda)$ and $h_w(\boldsymbol{\theta}_w, \lambda)$ increases the asymptotic variance of $\hat{d}$ by a multiplicative constant compared to LWN estimator of Hurvich & Ray (2003) (easily seen by use of the formula for the inverse of a partitioned matrix). Andrews & Sun (2004) obtain a similar result for their local polynomial Whittle (LPW) estimator in a non-volatility model.

The first condition in (15) guarantees that all the elements of the scaling matrix $\mathbf{B}_n$ diverge as $n \to \infty$, which is a minimal condition for consistency. The second condition restricts the expansion rate of the bandwidth to control bias and ensure that the estimator uses only relevant information from periodogram ordinates sufficiently near the zero frequency. Alternatively, we can view the bandwidth conditions in (15) separately for the signal process and the noise process. In this way we would write the conditions as

$$\frac{m^{1+4R_y}}{n^{4R_y}} \to \infty, \frac{m^{2\varphi_y+1}}{n^{2\varphi_y}} \to 0 \text{ and } \frac{m^{1+4(d_0+R_w)}}{n^{4(d_0+R_w)}} \to \infty, \frac{m^{2\varphi_w+4d_0+1}}{n^{2\varphi_w+4d_0}} \to 0.$$

It is now easy to see that the bandwidth conditions for both the signal process and the noise process are always compatible because $s_y > 2R_y$ and $s_w > 2R_w$, respectively, by Assumption A3.

Note that the second condition in (15) implies that if $\phi_y(\lambda)$ and $\phi_w(\lambda)$ are infinitely smooth near frequency zero then any (R_y, R_w) can be chosen and the estimator is $n^{1/2-\delta}$ consistent for all $\delta > 0$. Hence, in that case, the rate of convergence is arbitrarily close to the parametric rate. Thus, the condition (15) allows the bandwidth m to be much larger than for the LWN estimator and the standard LW estimator, which require that (assuming $s_y \geq 2, s_w \geq 2$) $m^5 n^{-4} \to 0$ and $m^{4d_0+1} n^{-4d_0} \to 0$, respectively, see Hurvich & Ray (2003) and Arteche (2004). Therefore, Theorem 2 provides an improvement in the rate of convergence relative to existing estimators of the memory parameter for perturbed fractional processes. This comes at the cost of an increase in the asymptotic variance by a multiplicative constant, but this is clearly more than off-set by the faster rate of convergence, at least asymptotically. For example, in the empirically relevant case of $d_0 = 0.4$, which is a typical value of d_0 for financial volatility series, the LW estimator is at most $n^{0.31}$-consistent and the LWN estimator is at most $n^{0.4}$-consistent, whereas our estimator can be arbitrarily close to $n^{0.5}$-consistent if the spectral density is sufficiently smooth near the zero frequency.

Finally, as in Andrews & Sun (2004) we could calculate the asymptotic bias which is of order $O((m/n)^{\varphi_y} + (m/n)^{2d_0+\varphi_w})$, where $\varphi_a = \min\{s_a, 2 + 2R_a\}, a = y, w$, see the proof of Lemma 3(e) in the appendix. This is in contrast to the orders $O((m/n)^2)$ and $O((m/n)^{2d_0})$ for the LWN and LW estimators in Hurvich & Ray (2003) and Arteche (2004). Thus, as in Andrews & Sun (2004) for the pure long memory case, the order of magnitude of the asymptotic bias is smaller when modeling the (smooth) spectral density of the short-memory component locally by a polynomial instead of a constant.

4 Finite sample comparison

In this section we present simulation results to examine the finite sample bias and root mean squared error (RMSE) performance of our LPWN estimator. In particular, we want to examine the accuracy with realistic sample sizes and short-run contamination in both signal and noise.

Our LPWN estimator is implemented with (R_y, R_w) equal to $(1, 0)$, $(0, 1)$, and $(1, 1)$, denoted LPWN(Ry, R_w), and is compared with the LW, LPW, and LWN estimators. From Hurvich & Ray

(2003) we know that the LWN estimator is superior to the LW estimator in terms of bias and RMSE in the context of the standard LMSV model. Furthermore, Hurvich et al. (2005) show that the polynomial log-periodogram regression estimator of Andrews & Guggenberger (2003) suffers from severe bias in the case of perturbed fractional processes and the LPW estimator is expected to perform similarly. Therefore, to conserve space we only compare the LWN and LPWN estimators in the Monte Carlo setup. The results for the LW and LPW estimators are avaible from the authors upon request.

4.1 Monte Carlo setup

We simulate model (2), i.e.

$$z_t = y_t + x_t + w_t, \tag{16}$$

where $\{y_t\}$ is the signal process and $\{x_t + w_t\}$ is the perturbation process. We model $\{x_t\}$ as an ARMA process and $\{w_t\}$ as

$$w_t = \log u_t^2, \quad u_t \sim NID(0,1). \tag{17}$$

Note that the variance of w_t is $\sigma_w^2 = \pi^2/2$ regardless of the variance of u_t. The signal process $\{y_t\}$ and the ARMA part $\{x_t\}$ of the perturbation process follow different DGPs. For brevity, we consider five different DGPs for the signal and ARMA perturbation processes. The general setup for $\{y_t\}$ and $\{x_t\}$ is

$$(1-\alpha_y L)(1-L)^d y_t = \left(1+\beta_y L\right)\eta_t, \quad \eta_t \sim NID(0,\sigma_\eta^2), \tag{18}$$

$$(1-\alpha_x L)\, x_t = (1+\beta_x L)\,\varepsilon_t, \quad \varepsilon_t \sim NID(0,1), \tag{19}$$

with parameter configurations

$$\begin{aligned}
\text{Model I} &: \alpha_y = \beta_y = \alpha_x = \beta_x = 0,\\
\text{Model II} &: \alpha_y = \beta_y = \beta_x = 0, \alpha_x \in \{-0.8, 0.5\},\\
\text{Model III} &: \alpha_y = \beta_y = \alpha_x = 0, \beta_x \in \{-0.8, 0.8\},\\
\text{Model IV} &: \beta_y = \beta_x = 0, (\alpha_y, \alpha_x) \in \{(-0.8, 0.5), (-0.8, 0.8)\},\\
\text{Model V} &: \alpha_y = \alpha_x = 0, (\beta_y, \beta_x) \in \{(-0.8, -0.8), (-0.8, 0.8)\}.
\end{aligned}$$

We remark that in all the models the noise-to-signal ratio is given as

$$nsr = \frac{f_x(0) + f_w(0)}{f_{(1-L)^d y_t}(0)} = \frac{\frac{(1+\beta_x)^2}{(1-\alpha_x)^2} + \frac{\pi^2}{2}}{\sigma_\eta^2 \frac{(1+\beta_y)^2}{(1-\alpha_y)^2}}. \tag{20}$$

For each Monte Carlo DGP we generated 1000 artificial time series with a sample size of 1024, 2048, 4096, and 8192.[6] For all estimators we set the bandwidth as $m = \lfloor n^a \rfloor$, where

[6] The number of observations is chosen as a power of two in order to use the fast Fourier transform in calculating the periodogram. This speeds up the estimations considerably compared to using the discrete Fourier transform.

$a \in \{0.6, 0.7, 0.8\}$. The parameter of interest, d, is set equal to either 0.4 or 0.6. For the noise-to-signal ratio, we choose $nsr \in \{5, 10, 20\}$, and the variance σ_η^2 is set as a function of $\alpha_y, \alpha_x, \beta_y$, and β_x such that the nsr has the desired value. The values of d, nsr, (α_y, β_y), (α_x, β_x), and the sample sizes are chosen to reflect empirical findings on long memory in volatility (see the references in the introduction for some examples). The chosen parameter values for the short-run contamination in the signal and the noise are also inspired by the results from the empirical (parametric) analysis of the DJIA stocks in section 5 below.

The signal $\{y_t\}$ is generated by the circulant embedding method as described in Davies & Harte (1987), i.e. the stationary type I fractionally integrated process in the terminology of Marinucci & Robinson (1999), see also Beran (1994, pp. 215-217). To generate nonstationary series with $d \geq 1/2$, we simulate the ARFIMA process with integration order $d - 1$ and cumulate the resulting series. Numerical optimization was carried out in Matlab v7.2 using the BFGS and DFP optimization routines and selecting the one with the best log-likelihood value. The initial values were set as follows. For the LWN estimator we used the LW estimate, $\hat{d}_{LW}$, if it was in the interior of the admissible space of d, i.e. $[0.01, 0.99]$, c.f. Assumption A2. Otherwise, d was set equal to 0.1. As starting value for the LPWN estimators we used the LWN estimate if it was in the admissible interval, otherwise d was set equal to 0.1.[7] As initial values for the polynomial parameters we used 1 for all estimators.

To conserve space we present only a subset of the results. The left-out results ($d = 0.6$, $n = 1024$, and $m = \lfloor n^{0.6} \rfloor$) are qualitatively very similar to the ones presented, and are available upon request.

4.2 Monte Carlo results

Tables 1-9 display the results of the simulation study and show how the two different sources of bias, i.e. the additive noise term and the contamination from the short-memory dynamics in both the signal and the noise, affect the estimators.

[Table 1 about here]

In the case where there is no contamination by short-run dynamics in the signal or noise, i.e. Model I with results displayed in Table 1, the bias is small for all estimators. The theoretical inflation of the variances from $h(\lambda, d, \boldsymbol{\theta})$ is also noticeable in the RMSEs. Additionally, the RMSE decreases as either the sample size or bandwidth increase. The only case with any noticeable bias is for the LPWN(1,1) estimator with $nsr = 20$, smallest sample size, and highest bandwidth.

[Tables 2 and 3 about here]

In Tables 2 and 3 we consider model II, i.e the signal is an ARFIMA$(0, d, 0)$ process and the noise is an ARMA process with coefficients $(\alpha_x, \beta_x) = (0.5, 0)$ and $(\alpha_x, \beta_x) = (-0.8, 0)$, respectively.

[7] We tried different starting values for d in these cases and the results were indistinguishable.

Here we would presume that the LPWN(0,1) estimator is the better choice. We clearly see that we are able to obtain considerable reduction in bias relative to the LWN estimator, especially for the positive AR root case in Table 2. In that case we find that all three LPWN estimators outperform the LWN estimator in terms of bias, and for the highest bandwidth choice, the LPWN(0,1) estimator is often also superior in terms of RMSE. In the model with a negative AR root in Table 3 the results are very similar to those in Table 1.

[Tables 4 and 5 about here]

We consider next Model III, i.e. where there is MA contamination in the noise, with results presented in Tables 4 and 5. The results for this model are similar to those in Tables 1 and 3. That is, for this model there is only little bias in the LWN estimator and no bias in the LPWN estimators. For the highest bandwidth choice, LWN and LPWN have similar RMSE.

[Tables 6 and 7 about here]

Tables 6 and 7 contain results for Model IV, where $(\alpha_y, \beta_y) = (-0.8, 0), (\alpha_x, \beta_x) = (0.5, 0)$ and $(\alpha_y, \beta_y) = (-0.8, 0), (\alpha_x, \beta_x) = (-0.8, 0)$, respectively. In the case of Table 6 the LWN estimator suffers from very high bias and the LPWN estimators are able to reduce this bias considerably. In particular, the LPWN(1,1) estimator is nearly unbiased in most cases. For the high bandwidth the RMSEs are similar for all estimators. In Table 7, where the contamination is by a negative root, the performance of the LWN estimator is similar to that of the LPWN estimators.

[Tables 8 and 9 about here]

Results for Model V where $(\alpha_y, \beta_y) = (0, -0.8), (\alpha_x, \beta_x) = (0, 0.8)$ and $(\alpha_y, \beta_y) = (0, -0.8), (\alpha_x, \beta_x) = (0, -0.8)$ are shown in Tables 8 and 9, respectively.[8] The LWN estimator suffers from very severe bias in Model V, and consequently its RMSE is also higher than for the previous models. On the other hand, the LPWN estimators have relatively low biases, and in particular the LPWN(1,1) estimator appears essentially unbiased. When compared in terms of RMSE the LPWN estimators are superior in both tables as well. Thus, we have a considerable reduction in bias for all LPWN estimators compared to the LWN estimator, and we also have quite a remarkable reduction in RMSE.

To sum up, the Monte Carlo study shows the usefulness of estimators that explicitly take the short-run dynamics in the perturbation into account, i.e. the LPWN estimators where $(R_y, R_w) = (0, 1)$ and $(R_y, R_w) = (1, 1)$, although the LPWN estimator with $(R_y, R_w) = (1, 0)$ also performs well. All three estimators generally have much smaller biases than the LWN estimator and are fairly insensitive to the persistence in the perturbation and to the contamination from short-memory dynamics in the signal.

[8] In a few cases for the LWN estimator (marked with asterisks) we had convergence problems due to boundary issues resulting in a markedly bimodal finite-sample distribution. In these cases we set the initial value for the polynomial parameter to 10, which resolved the issue.

5 Long memory in DJIA stock volatility

This section analyzes the long memory in daily log-squared returns series of the 30 DJIA stocks corrected for the effects of stock splits and dividends from January 1 1990 to March 31 2008, for a sample of $n = 4753$. To avoid the problem of taking logarithm of zero we based the analysis on adjusted log-squared returns using the method of Fuller (1996, pp. 495-496), i.e. we analyze

$$\log \tilde{r}_t^2 = \log \left(r_t^2 + \alpha \right) - \frac{\alpha}{r_t^2 + \alpha},$$

where $\alpha = \frac{0.02}{n} \sum_{t=1}^{n} r_t^2$. We estimate the long memory in $\log \tilde{r}_t^2$ using the proposed LPWN estimator. We implement the estimator with (R_y, R_w) equal to $(1, 0)$, $(0, 1)$, and $(1, 1)$, and with starting values etc. as in the Monte Carlo study above. For comparison we also report the standard LW, LPW, and LWN estimates. For all estimators we set the bandwidth as $m = \lfloor n^a \rfloor$, where $a \in \{0.6, 0.7, 0.8\}$.

[**Table 10 about here**]

Table 10 presents the results for the LW, LPW, and LWN estimators. As expected from theory, the LW and LPW estimators appear downward biased and are decreasing in the bandwidth. For the LWN estimator the memory estimates of some of the stocks are in the stationary region, but for the most part they are in the nonstationary region.

[**Table 11 about here**]

In Table 11 we present the results for the three variants of the LPWN estimator, i.e. for (R_y, R_w) equal to $(1, 0)$, $(0, 1)$, and $(1, 1)$. First of all, as expected from theory and the simulations above, it is clear that this estimator does not suffer from the downward bias present in the LW and LPW estimators. Second, we note that the three different implementations of the estimator agree with each other for most of the stocks and bandwidth choices. Thirdly, the LPWN estimates are of the same order of magnitude as the LWN estimates, although a little higher on average.

To emphasize the importance of the polynomial approximation of the signal process $\{y_t\}$ and the pertubation process $\{x_t + w_t\}$, we also fitted an extended parametric LMSV-ARFIMA$(1, d, 1)$ model, where the extension is that the noise is modeled by an ARMA process. That is, we model the periodogram of $\log \tilde{r}_t^2$ using the Whittle likelihood framework of Fox & Taqqu (1986) and Breidt et al. (1998), where the fitted model has spectral density

$$f_z(\lambda) = \frac{\sigma_\eta^2}{2\pi} \left(2 \sin \frac{\lambda}{2} \right)^{-2d} \frac{\left(1 + 2\beta_y \cos \lambda + \beta_y^2\right)}{\left(1 - 2\alpha_y \cos \lambda + \alpha_y^2\right)} + \frac{\sigma_\varepsilon^2}{2\pi} \frac{\left(1 + 2\beta_x \cos \lambda + \beta_x^2\right)}{\left(1 - 2\alpha_x \cos \lambda + \alpha_x^2\right)}. \tag{21}$$

In Table 12 the resulting estimates are reported, where we have removed insignificant ARMA terms from both the signal and the noise.

[**Insert Table 12 about here**]

The estimated values of d from the parametric results are in line with those from the LWN and LPWN estimators in Tables 10 and 11. Furthermore, there is significant (at 10% level) short-run dynamics in the signal (19 out of 30 cases), in the noise (16 out of 30 cases), and in both the signal and noise (13 out of 30 cases). The estimated (long-run) nsr's can be calculated from the parameter estimates as in (20), and are for most of the stocks in the vicinity of $10 - 30$, although there are cases where the nsr is very high because σ_η^2 is very small and insignificant. Taking the high nsr's and significant short-run dynamics in both the signal and the noise into consideration stresses the importance of the LPWN estimators.

6 Concluding remarks

In this paper we have proposed a semiparametric local polynomial Whittle with noise estimator of the degree of long memory, d, in financial volatility time series perturbed by dynamic short-run noise. The estimator allows the spectrum of the perturbation and that of the short-memory component of the signal to be modeled as finite even polynomials, instead of constants near the zero frequency. This is shown to yield a bias reduction depending on the smoothness of the spectra. However, including the polynomials inflates the asymptotic variance of $\hat{d}$ by a multiplicative constant which depends on the true long memory parameter, d.

We have shown that the estimator is consistent for $d \in (0,1)$, asymptotically normal for $d \in (0, 3/4)$, and if the spectral density is sufficiently smooth near frequency zero the rate of convergence becomes arbitrary close to the parametric rate, $\sqrt{n}$.

A Monte Carlo study revealed that the proposed local polynomial Whittle with noise estimator is able to achieve considerable bias reductions in practice compared to standard (e.g., local Whittle with noise) estimators, especially in cases with short-run dynamics in both the signal and noise components. In an empirical investigation of the 30 DJIA stocks the local polynomial Whittle with noise estimator indicated stronger persistence in volatility than standard estimators, and for most of the stocks produced estimates of d in the nonstationary region.

Appendix A: Proof of Theorem 1

This proof follows the proofs of Theorem 3.1 and Lemma C.2 of Hurvich et al. (2005). As in the proofs of Theorem 1 of Robinson (1995a) and Theorem 3.1 of Hurvich et al. (2005), to show consistency of $\hat{d}$ we need to separately prove that $\lim_{n\to\infty} P(\hat{d} \in D_1) = 0$ and that $(\hat{d} - d_0)\mathbf{1}(\hat{d} \in D_2) \xrightarrow{P} 0$, where $\mathbf{1}(A)$ is the indicator function of the set A, $D_1 = (-\infty, d_0 - 1/2 + \epsilon) \cap D$, $D_2 = [d_0 - 1/2 + \epsilon, +\infty) \cap D$, and $\epsilon < 1/4$ is a positive real number to be set later.

Let $\alpha_k(d, \boldsymbol{\theta}) = \frac{1+h_k(d_0,\boldsymbol{\theta}_0)}{1+h_k(d,\boldsymbol{\theta})}$. Then the proof that $(\hat{d} - d_0)\mathbf{1}(\hat{d} \in D_2) \xrightarrow{P} 0$ follows as in Hurvich et al. (2005, pp. 1303-1305) by showing that

$$Z_m = \sum_{k=1}^{m} \frac{k^{2(d-d_0)}\alpha_k(d,\boldsymbol{\theta})}{\sum_{j=1}^{m} j^{2(d-d_0)}\alpha_j(d,\boldsymbol{\theta})} \left(\frac{I_z(\lambda_k)}{f_z(\lambda_k)} - 1 \right) = o_P(1) \tag{22}$$

uniformly on $(d,\boldsymbol{\theta}) \in D_2 \times \Theta$ and that

$$R_m(d,\boldsymbol{\theta}) = \log \left(1 + \frac{\sum_{k=1}^{m} k^{2(d-d_0)} \left(\alpha_k(d,\boldsymbol{\theta}) - 1\right)}{\sum_{j=1}^{m} j^{2(d-d_0)}} \right) - \frac{1}{m}\sum_{k=1}^{m} \log\left(1 + \left(\alpha_k(d,\boldsymbol{\theta}) - 1\right)\right) = o(1) \tag{23}$$

uniformly on $(d,\boldsymbol{\theta}) \in D \times \Theta$.

Note that there exists a constant $C > 0$ such that

$$\sup_{(d,\boldsymbol{\theta})\in D\times\Theta} \sup_{k=1,\ldots,m} \left|\alpha_k(d,\boldsymbol{\theta}) - 1\right| = \sup_{(d,\boldsymbol{\theta})\in D\times\Theta} \sup_{k=1,\ldots,m} \left| \frac{h_k(d_0,\boldsymbol{\theta}_0) - h_k(d,\boldsymbol{\theta})}{1 + h_k(d,\boldsymbol{\theta})} \right| \le C\,(m/n)^{2d_1},$$

since Θ is compact and $d \ge d_1 > 0$, see Lemma 4. Now we use that $\log(1+x) = x + O(x^2)$ as $x \to 0$ to obtain

$$\sup_{(d,\boldsymbol{\theta})\in D\times\Theta} |R_m(d,\boldsymbol{\theta})| \le C \sup_{(d,\boldsymbol{\theta})\in D\times\Theta} \sup_{k=1,\ldots,m} \left|\alpha_k(d,\boldsymbol{\theta}) - 1\right| \le C\,(m/n)^{2d_1} = o(1)\,.$$

To show (22) we apply Proposition A.1 of Hurvich et al. (2005), which holds here since our Assumptions A1-A6 imply their Assumptions (H1)-(H3) with the exception that we allow serially correlated peturbation terms. It is, however, easily shown that replacing their Assumption (H2) with our Assumption A5, their Proposition A.1 still holds. The only other change is that the term $(k/n)^{\min(\beta,d_0)}$ in their eq. (F.15) should be replaced by $(k/n)^{\varphi_y} + (k/n)^{\varphi_w}$ due to the more accurate approximation of $f_z(\lambda)$ offered by our function $g(\lambda)$ in (10) due to the included polynomials, see also Lemma 5 below. Thus, according to their Proposition A.1, letting

$$c_k = \frac{k^{2(d-d_0)}\alpha_k(d,\boldsymbol{\theta})}{\sum_{j=1}^{m} j^{2(d-d_0)}\alpha_j(d,\boldsymbol{\theta})},$$

then for $\epsilon \in (0,1)$, $K \in (0,\infty)$, and all $k \in \{1,\ldots,m-1\}$, we need to show that

$$|c_k - c_{k+1}| \le K m^{-\epsilon} k^{\epsilon-2}, |c_m| \le K m^{-1}$$

uniformly on $(d,\boldsymbol{\theta}) \in D_2 \times \Theta$, which implies (22).

Note that, uniformly on $(d,\boldsymbol{\theta}) \in D_2 \times \Theta$, we have that $\sum_{j=1}^{m} j^{2(d-d_0)}\alpha_j(d,\boldsymbol{\theta}) \ge C m^{2(d-d_0)+1}$ and

$$\begin{aligned}
& |k^{2(d-d_0)}\alpha_k(d,\boldsymbol{\theta}) - (k+1)^{2(d-d_0)}\alpha_{k+1}(d,\boldsymbol{\theta})| \\
\le\; & |k^{2(d-d_0)} - (k+1)^{2(d-d_0)}|\alpha_k(d,\boldsymbol{\theta}) + (k+1)^{2(d-d_0)}|\alpha_k(d,\boldsymbol{\theta}) - \alpha_{k+1}(d,\boldsymbol{\theta})| \\
\le\; & (k+a)^{2(d-d_0)-1}C + (k+1)^{2(d-d_0)}C(\lambda_{k+1} - \lambda_k)\lambda_{k+a}^{2d-1}, a \in [0,1] \\
\le\; & Ck^{2(d-d_0)-1},
\end{aligned}$$

where the first inequality is the triangle inequality and the second follows from the mean value theorem and Lemma 4. It follows that

$$\begin{aligned}
\sup_{(d,\boldsymbol{\theta})\in D_2\times\Theta}\left|\frac{k^{2(d-d_0)}\alpha_k(d,\boldsymbol{\theta})-(k+1)^{2(d-d_0)}\alpha_{k+1}(d,\boldsymbol{\theta})}{\sum_{j=1}^m j^{2(d-d_0)}\alpha_j(d,\boldsymbol{\theta})}\right| &\leq \sup_{(d,\boldsymbol{\theta})\in D_2\times\Theta} C\left|\frac{k^{2(d-d_0)-1}}{m^{2(d-d_0)+1}}\right| \leq Ck^{2\epsilon-2}m^{-2\epsilon}, \\
\sup_{(d,\boldsymbol{\theta})\in D_2\times\Theta}\left|\frac{m^{2(d-d_0)}\alpha_m(d,\boldsymbol{\theta})}{\sum_{j=1}^m j^{2(d-d_0)}\alpha_j(d,\boldsymbol{\theta})}\right| &\leq Cm^{-1},
\end{aligned}$$

which proves (22).

The proof that $\lim_{n\to\infty} P(\hat{d}\in D_1)=0$ follows exactly as in Hurvich et al. (2005, pp. 1305-1306) since their Proposition A.1 holds in our case as well. Thus we have shown that $\hat{d}\overset{P}{\to} d_0$. To strengthen this result to $\hat{d}-d_0=o_P((\log n)^{-5})$ we use the proof of Lemma C.2 of Hurvich et al. (2005) without change.

Appendix B: Proof of Theorem 2

For the proof of Theorem 2 we need the score and Hessian (both multiplied by m) of (11):

$$\begin{aligned}
\mathbf{S}_n(d,\boldsymbol{\theta}) &= \hat{G}(d,\boldsymbol{\theta})^{-1}\sum_{j=1}^m\left(\frac{GI_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}-\frac{1}{m}\sum_{k=1}^m\frac{GI_z(\lambda_k)}{g_k(d,\boldsymbol{\theta})}\right)\mathbf{X}_j, \\
\mathbf{H}_n(d,\boldsymbol{\theta}) &= \mathbf{H}_{1n}(d,\boldsymbol{\theta})+\mathbf{H}_{2n}(d,\boldsymbol{\theta}), \\
\mathbf{H}_{1n}(d,\boldsymbol{\theta}) &= \hat{G}(d,\boldsymbol{\theta})^{-2}\left(\hat{G}(d,\boldsymbol{\theta})\sum_{j=1}^m\frac{GI_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\mathbf{X}_j\mathbf{X}_j'-m\left(\frac{1}{m}\sum_{j=1}^m\frac{GI_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\mathbf{X}_j\right)\left(\frac{1}{m}\sum_{j=1}^m\frac{GI_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\mathbf{X}_j\right)'\right), \\
\mathbf{H}_{2n}(d,\boldsymbol{\theta}) &= \hat{G}(d,\boldsymbol{\theta})^{-1}\sum_{j=1}^m\left(\frac{GI_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}-\frac{1}{m}\sum_{k=1}^m\frac{GI_z(\lambda_k)}{g_k(d,\boldsymbol{\theta})}\right)\frac{\partial\mathbf{X}_j}{\partial(d,\boldsymbol{\theta}')},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{X}_j &= (X_{1j},\mathbf{X}_{2j}',\mathbf{X}_{3j}')', \\
X_{1j} &= 2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}, \\
\mathbf{X}_{2j} &= \left(\frac{-\lambda_j^2}{(1+h_j(d,\boldsymbol{\theta}))},\ldots,\frac{-\lambda_j^{2R_y}}{(1+h_j(d,\boldsymbol{\theta}))}\right)', \\
\mathbf{X}_{3j} &= \left(\frac{-\lambda_j^{2d}}{(1+h_j(d,\boldsymbol{\theta}))},\ldots,\frac{-\lambda_j^{2d+2R_w}}{(1+h_j(d,\boldsymbol{\theta}))}\right)',
\end{aligned}$$

$h_j(d,\boldsymbol{\theta})=h(d,\boldsymbol{\theta},\lambda_j)$, $g_j(d,\boldsymbol{\theta})=\lambda_j^{-2d}G(1+h_j(d,\boldsymbol{\theta}))$, and $D_m(\eta)=\{d\in D:(\log m)^5|d-d_0|<\eta\}$ for $\eta>0$. Note that $\mathbf{X}_j$ is the vector of partial derivatives of $-\log g_j(d,\boldsymbol{\theta})$. The matrix $\mathbf{H}_{2n}(d,\boldsymbol{\theta})$

is symmetric and has (i,l)'th and (l,i)'th elements

$$\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_{z}\left(\lambda_{j}\right)}{g_{j}\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_{z}\left(\lambda_{k}\right)}{g_{k}\left(d,\boldsymbol{\theta}\right)}\right)\left(\mathbf{X}_{j}\right)_{i}\left(\mathbf{X}_{j}\right)_{l},\ i,l=2,\ldots,R+2,$$

$$-\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_{z}\left(\lambda_{j}\right)}{g_{j}\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_{z}\left(\lambda_{k}\right)}{g_{k}\left(d,\boldsymbol{\theta}\right)}\right)\left(\mathbf{X}_{j}\right)_{i}\frac{2h_{w}(\boldsymbol{\theta}_{w},\lambda_{j})\lambda_{j}^{2d}\log\lambda_{j}}{(1+h_{j}(d,\boldsymbol{\theta}))},\ i=2,\ldots,R_{y}+1,l=1,$$

$$\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_{z}\left(\lambda_{j}\right)}{g_{j}\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_{z}\left(\lambda_{k}\right)}{g_{k}\left(d,\boldsymbol{\theta}\right)}\right)\left(\mathbf{X}_{j}\right)_{i}2\log\lambda_{j}\left(1-\frac{h_{w}(\boldsymbol{\theta}_{w},\lambda_{j})\lambda_{j}^{2d}}{(1+h_{j}(d,\boldsymbol{\theta}))}\right),\ i=R_{y}+2,\ldots,R+2,l=1,$$

$$\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_{z}\left(\lambda_{j}\right)}{g_{j}\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_{z}\left(\lambda_{k}\right)}{g_{k}\left(d,\boldsymbol{\theta}\right)}\right)\left(\mathbf{X}_{j}\right)_{R_{y}+2}4h_{w}(\boldsymbol{\theta}_{w},\lambda_{j})\left(\log\lambda_{j}\right)^{2}\left(1-\frac{h_{w}(\boldsymbol{\theta}_{w},\lambda_{j})\lambda_{j}^{2d}}{(1+h_{j}(d,\boldsymbol{\theta}))}\right),\ i=l=1.$$

We also define the matrix

$$\mathbf{J}_{n}=\sum_{j=1}^{m}\left(\mathbf{X}_{j}-\frac{1}{m}\sum_{k=1}^{m}\mathbf{X}_{k}\right)\left(\mathbf{X}_{j}-\frac{1}{m}\sum_{k=1}^{m}\mathbf{X}_{k}\right)'.$$

We next state a lemma adapted from Andrews & Sun (2004), henceforth abbreviated AS. The proof is given in the next section.

Lemma 3 *Under the assumptions of Theorem 2 we have, as $n\to\infty$,*

(a) $\mathbf{B}_{n}^{-1}\mathbf{J}_{n}\mathbf{B}_{n}^{-1}\to\boldsymbol{\Omega}_{R_{y},R_{w}}$,

(b) $\left\Vert\mathbf{B}_{n}^{-1}\left(\mathbf{H}_{1n}\left(d_{0},\boldsymbol{\theta}_{0}\right)-\mathbf{J}_{n}\right)\mathbf{B}_{n}^{-1}\right\Vert=o_{P}\left(1\right)$ *and* $\left\Vert\mathbf{B}_{n}^{-1}\mathbf{H}_{2n}\left(d_{0},\boldsymbol{\theta}_{0}\right)\mathbf{B}_{n}^{-1}\right\Vert=o_{P}\left(1\right),$

(c) $\sup_{\boldsymbol{\theta}\in\Theta}\left\Vert\mathbf{B}_{n}^{-1}\left(\mathbf{H}_{kn}\left(d_{0},\boldsymbol{\theta}\right)-\mathbf{H}_{kn}\left(d_{0},\boldsymbol{\theta}_{0}\right)\right)\mathbf{B}_{n}^{-1}\right\Vert=o_{P}\left(1\right),\ k=1,2,$

(d) $\sup_{d\in D_{m}(\eta_{n}),\boldsymbol{\theta}\in\Theta}\left\Vert\mathbf{B}_{n}^{-1}\left(\mathbf{H}_{kn}\left(d,\boldsymbol{\theta}\right)-\mathbf{H}_{kn}\left(d_{0},\boldsymbol{\theta}\right)\right)\mathbf{B}_{n}^{-1}\right\Vert=o_{P}\left(1\right),\ k=1,2,$ *for all sequences of constants* $\{\eta_{n}\}_{n\geq1}$ *for which* $\eta_{n}=o\left(1\right)$,

(e) $\mathbf{B}_{n}^{-1}\mathbf{S}_{n}\left(d_{0},\boldsymbol{\theta}_{0}\right)\xrightarrow{d}N\left(\mathbf{0},\boldsymbol{\Omega}_{R_{y},R_{w}}\right).$

Since the LPWN likelihood (11) is a continuous function on a compact set the LPWN estimator exists. From Lemma 3 we know by Lemma 1 of AS that there exists a solution to the first order conditions with probability tending to one, and that the solution satisfies the convergence result in Theorem 2, see also Lemmas 1 and 2 of AS. If the (negative) likelihood function is strictly convex and twice differentiable then the solution to the first order conditions is unique and minimizes (11) and hence equals the LPWN estimator.

Thus, all that remains is to show that the Hessian is positive definite which proves convexity. The positive definiteness of $\mathbf{H}_{1n}$ follows as in eq. (5.1) of AS. Compared to AS we have the additional term $\mathbf{H}_{2n}$. For $\mathbf{H}_{2n}$ we know that $\left\Vert\mathbf{B}_{n}^{-1}\mathbf{H}_{2n}(d,\boldsymbol{\theta})\mathbf{B}_{n}^{-1}\right\Vert=o_{P}\left(1\right)$ uniformly on $(d,\boldsymbol{\theta})\in D_{m}\left(\eta_{n}\right)\times\Theta$ by Lemma 3(b)-(d) and the triangle inequality. Since $\hat{d}\in D_{m}\left(\eta_{n}\right)$ with probability tending to one by Theorem 1, this shows that $\mathbf{H}_{n}$ is positive definite with probability tending to one, which concludes the proof.

Appendix C: Proof of Lemma 3

We now turn to the proof of Lemma 3, which follows the method of proof for Lemma 2 of AS, with modifications to allow $d \geq 1/2$ (following Velasco (1999)) and to accommodate the additive noise term in the spectral density (see Lemma 5), and with an additional proof for each of (b), (c), and (d) of negligibility of the term $\mathbf{H}_{2n}(d, \boldsymbol{\theta})$.

C.1 Proof of (a)

Part (a) of the lemma follows by approximating sums by integrals, see, e.g., Lemma 2 of Andrews & Guggenberger (2003).

C.2 Proof of (b), first statement

The proof roughly follows that of Lemma 2(b) in AS, except now b can be non-integer (equal to d or $2d$) in their eq. (A.6), which we write a little differently as

$$\begin{aligned}
\tilde{G}_{a,b,c}(d,\boldsymbol{\theta}) &= m^{-1}\sum_{j=1}^{m}\frac{\lambda_j^{2d}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))^{c+1}}\left(2\log j - \frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a\left(\frac{j}{m}\right)^{2b}, \\
\hat{G}_{a,b}(d,\boldsymbol{\theta}) &= m^{-1}\sum_{j=1}^{m}\lambda_j^{2d}I_z(\lambda_j)\,(2\log j)^a\left(\frac{j}{m}\right)^{2b}, \\
J_{a,b} &= Gm^{-1}\sum_{j=1}^{m}(2\log j)^a\left(\frac{j}{m}\right)^{2b},
\end{aligned}$$

for $a, c = 0, 1, 2$ and $b = 0, 1, \ldots, 2R_y, d, d+1, \ldots, d+R_w+R_y, 2d, 2d+1, \ldots, 2d+2R_w$. The elements of $\mathbf{B}_n^{-1}\mathbf{H}_{1n}\left(d,\boldsymbol{\theta}\right)\mathbf{B}_n^{-1}$ are (omitting the argument for brevity)

$$\begin{aligned}
(1,1) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{2,0,0}-\tilde{G}_{1,0,0}^2\right), \\
(1,1+k) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{1,k,1}-\tilde{G}_{1,0,0}\tilde{G}_{0,k,1}\right) \text{ for } k=1,\ldots,R_y, \\
(1,2+R_y+k) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{1,k+d,1}-\tilde{G}_{1,0,0}\tilde{G}_{0,k+d,1}\right) \text{ for } k=0,\ldots,R_w, \\
(1+i,1+k) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{0,i+k,2}-\tilde{G}_{0,i,1}\tilde{G}_{0,k,1}\right) \text{ for } i,k=1,\ldots,R_y, \\
(1+i,2+R_y+k) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{0,k+i+d,2}-\tilde{G}_{0,i,1}\tilde{G}_{0,k+d,1}\right) \text{ for } i=1,\ldots,R_y, k=0,\ldots,R_w, \\
(2+R_y+i,2+R_y+k) &: \tilde{G}_{0,0,0}^{-2}\left(\tilde{G}_{0,0,0}\tilde{G}_{0,k+i+2d,2}-\tilde{G}_{0,i+d,1}\tilde{G}_{0,k+d,1}\right) \text{ for } i,k=0,\ldots,R_w,
\end{aligned}$$

and the corresponding elements of $\mathbf{B}_n^{-1}\mathbf{J}_n\left(d,\boldsymbol{\theta}\right)\mathbf{B}_n^{-1}$ are given by the same expressions with $\tilde{G}_{a,b,c}$ replaced by $J_{a,b}$. To prove the first statement of Lemma 3(b) it suffices to show that (since b can take values including d, we distinguish between b and b_0)

$$\Delta_{a,b_0} = \left|\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta}_0) - J_{a,b_0}\right| = o_P((\log m)^{-2}), \tag{24}$$

$$\tilde{\Delta}_{a,b_0,c} = \left|\tilde{G}_{a,b_0,c}(d_0,\boldsymbol{\theta}_0) - \hat{G}_{a,b_0}(d_0,\boldsymbol{\theta}_0)\right| = o_P((\log m)^{-2}). \tag{25}$$

In view of Lemma 5 below, the proof of (A.9) in AS pp. 598-599 works also for our eq. (24) where we find that ($\xi_{k,n}(d)$ is defined in Lemma 5)

$$\Delta_{a,b_0} = O_P\Big((\log m)^a m^{-1}\xi_{m,n}(d_0) + (\log m)^a m^{\varphi_y} n^{-\varphi_y} + (\log m)^a m^{d_0+\varphi_w} n^{-d_0-\varphi_w} \\ + (\log m)^a m^{2d_0} n^{-2d_0} + (\log m)^{a+1} m^{2d_0-1} n^{-d_0} + (\log m)^a m^{-1/2}\Big),$$

which is

$$O_P\Big((\log m)^{a+2/3} m^{-2/3} + (\log m)^a m^{-1/2} n^{-1/4} + (\log m)^a (m/n)^{\min(\varphi_y, d_0+\varphi_w, 2d_0)} \\ +(\log m)^{a+1} m^{2d_0-1} n^{-d_0} + (\log m)^a m^{-1/2}\Big)$$

in the stationary case and

$$O_P\Big((\log m)^{a+2/(5-4d_0)} m^{1/(5-4d_0)-1} + (\log m)^{a+1} m^{2d_0-2} + (\log m)^a m^{(d_0-1)/2} n^{-1/2} (\log n)^{5/4} \\ +(\log m)^{a+1/2} n^{-1/4} m^{d_0-1} + (\log m)^a (m/n)^{\min(\varphi_y, d_0+\varphi_w, 2d_0)} + (\log m)^{a+1} m^{2d_0-1} n^{-d_0} + (\log m)^a m^{-1/2}\Big)$$

in the nonstationary case. Since $d_0 < d_2 < 3/4$ and by (15), clearly $\Delta_{a,b_0} = o_P((\log m)^{-2})$ in both cases.

To prove (25) we write $\tilde{G}_{a,b_0,c}(d_0, \boldsymbol{\theta}_0) - \hat{G}_{a,b_0}(d_0, \boldsymbol{\theta}_0)$ as

$$m^{-1}\sum_{j=1}^{m} \lambda_j^{2d_0} I_z(\lambda_j) \left[\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}_0))^{c+1}} \left(2\log j - \frac{2h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\lambda_j^{2d_0}\log\lambda_j}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)^a - (2\log j)^a\right] \left(\frac{j}{m}\right)^{2b_0}$$
$$= m^{-1}\sum_{j=1}^{m} \lambda_j^{2d_0} I_z(\lambda_j) \left[\frac{1}{1+O((j/n)^{2d_0})} \left(2\log j - \frac{O((j/n)^{2d_0}\log n)}{1+O((j/n)^{2d_0})}\right)^a - (2\log j)^a\right] \left(\frac{j}{m}\right)^{2b_0}$$

by Lemma 4(i). This proves (25) for $a = 0$ since

$$\begin{aligned} \tilde{G}_{0,b_0,c}(d_0,\boldsymbol{\theta}_0) - \hat{G}_{0,b_0}(d_0,\boldsymbol{\theta}_0) &= m^{-1}\sum_{j=1}^{m} \lambda_j^{2d_0} I_z(\lambda_j) \left(\frac{1}{1+O((j/n)^{2d_0})} - 1\right) \left(\frac{j}{m}\right)^{2b_0} \\ &= O_P\left((m/n)^{2d_0} \hat{G}_{0,b_0}(d_0,\boldsymbol{\theta}_0)\right) \\ &= O_P\left((m/n)^{2d_0}\right) \\ &= o_P((\log m)^{-2}) \end{aligned}$$

because d_0 belongs to the interior of the parameter space and is therefore bounded away from zero. When $a \geq 1$ we apply the mean value theorem, i.e. $x^a = y^a + (y-x)a\bar{x}^{a-1}$ for $x \leq \bar{x} \leq y$, such that

$$\left(2\log j - \frac{2h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\lambda_j^{2d_0}\log\lambda_j}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)^a - (2\log j)^a = a\frac{2h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\lambda_j^{2d_0}\log\lambda_j}{(1+h_j(d_0,\boldsymbol{\theta}_0))} O((\log j)^{a-1})$$

uniformly in $j = 1, \dots, m$. This impiles that (25) is

$$
\begin{aligned}
& m^{-1} \sum_{j=1}^{m} \lambda_j^{2d_0} I_z(\lambda_j) \left[aO((j/n)^{2d_0} \log n) O((\log j)^{a-1}) \right] \left(\frac{j}{m} \right)^{2b_0} \\
= \;& O_P((m/n)^{2d_0} (\log n)(\log m)^{a-1} \hat{G}_{0,b_0}(d_0, \boldsymbol{\theta}_0)) \\
= \;& O_P\left((m/n)^{2d_0} (\log n)^a \right) \\
= \;& o_P((\log m)^{-2}).
\end{aligned}
$$

C.3 Proof of (e)

We now prove part (e) since it will be useful in the proof of the remaining statements. By (24) and (25) with $a = b = c = 0$ we get that $\hat{G}\left(d_0, \boldsymbol{\theta}_0\right) = G_0(1 + o_P((\log m)^{-2}))$, so that, apart from smaller order terms,

$$
\begin{aligned}
\mathbf{B}_n^{-1} \mathbf{S}_n\left(d_0, \boldsymbol{\theta}_0\right) &= m^{-1/2} \sum_{j=1}^{m} \left(\frac{I_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} - \frac{1}{m} \sum_{k=1}^{m} \frac{I_z\left(\lambda_k\right)}{g_k\left(d_0, \boldsymbol{\theta}_0\right)} \right) \tilde{\mathbf{X}}_{0,j} \\
&= m^{-1/2} \sum_{j=1}^{m} \left(\frac{I_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} - 1 \right) \left(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right), \qquad (26)
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\mathbf{X}}_j &= (X_{1,j}, \tilde{\mathbf{X}}_{2,j}', \tilde{\mathbf{X}}_{3,j}')', \\
\tilde{\mathbf{X}}_{2,j} &= \left(\frac{-(j/m)^2}{(1 + h_j(d, \boldsymbol{\theta}))}, \dots, \frac{-(j/m)^{2R_y}}{(1 + h_j(d, \boldsymbol{\theta}))} \right)', \\
\tilde{\mathbf{X}}_{3,j} &= \left(\frac{-(j/m)^{2d}}{(1 + h_j(d, \boldsymbol{\theta}))}, \dots, \frac{-(j/m)^{2d + 2R_w}}{(1 + h_j(d, \boldsymbol{\theta}))} \right)',
\end{aligned}
$$

and $\tilde{\mathbf{X}}_{0,j}$ is $\tilde{\mathbf{X}}_j$ evaluated at $(d_0, \boldsymbol{\theta}_0)$.

As in AS p. 601 we write the right-hand side of (26) as $T_{1,n} + T_{2,n} + T_{3,n} + T_{4,n}$, where

$$
\begin{aligned}
T_{1,n} &= m^{-1/2} \sum_{j=1}^{m} \left(\frac{I_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} - 2\pi I_\varepsilon\left(\lambda_j\right) - E\left(\frac{I_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} - 2\pi I_\varepsilon\left(\lambda_j\right) \right) \right) \left(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right), \\
T_{2,n} &= m^{-1/2} \sum_{j=1}^{m} \left(\frac{E I_z\left(\lambda_j\right)}{f_z\left(\lambda_j\right)} - 1 \right) \frac{f_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} \left(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right), \\
T_{3,n} &= m^{-1/2} \sum_{j=1}^{m} \left(2\pi I_\varepsilon\left(\lambda_j\right) - 1 \right) \left(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right), \\
T_{4,n} &= m^{-1/2} \sum_{j=1}^{m} \left(\frac{f_z\left(\lambda_j\right)}{g_j\left(d_0, \boldsymbol{\theta}_0\right)} - 1 \right) \left(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right).
\end{aligned}
$$

Then we show that $T_{3,n} \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{\Omega}_r\right)$ while $T_{i,n} = o_P(1)$ for $i = 1, 2, 4$.

Clearly the proof for $T_{3,n}$ of AS works here as well. We just have to verify that

$$\frac{1}{m}\sum_{j=1}^{m}\zeta_j^2 \to \boldsymbol{\beta}'\boldsymbol{\Omega}_{R_y,R_w}\boldsymbol{\beta},$$

where

$$\zeta_j = \boldsymbol{\beta}'(\tilde{\mathbf{X}}_{0,j} - \frac{1}{m}\sum_{k=1}^{m}\tilde{\mathbf{X}}_{0,k}) \text{ and } \boldsymbol{\Omega}_{R_y,R_w} = \begin{pmatrix} 4 & \boldsymbol{\mu}'_{R_y} & \boldsymbol{\nu}'_{R_w} \\ \boldsymbol{\mu}_{R_y} & \boldsymbol{\Gamma}_{R_y} & \boldsymbol{\psi}'_{R_y,R_w} \\ \boldsymbol{\nu}_{R_w} & \boldsymbol{\psi}_{R_w,R_y} & \boldsymbol{\Psi}_{R_w} \end{pmatrix},$$

which follows from part (a) of the lemma.

To show the result for $T_{1,n}$ we use summation by parts:

$$\begin{aligned} T_{1,n} &= m^{-1/2}\sum_{k=1}^{m-1}\left(\tilde{\mathbf{X}}_{0,k} - \tilde{\mathbf{X}}_{0,k+1}\right)\sum_{j=1}^{k}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j) - E\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j)\right)\right) \\ &\quad + \left(\tilde{\mathbf{X}}_{0,m} - \frac{1}{m}\sum_{k=1}^{m}\tilde{\mathbf{X}}_{0,k}\right) m^{-1/2}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j) - E\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j)\right)\right) \\ &= m^{-1/2}\sum_{k=1}^{m-1} O(k^{-1})O_P(\xi_{k,n}(d_0) + k^{\varphi_y+1/2}n^{-\varphi_y} + k^{1/2+2d_0}n^{-2d_0}) \\ &\quad + O(1)m^{-1/2}O_P(\xi_{m,n}(d_0) + m^{\varphi_y+1/2}n^{-\varphi_y} + m^{1/2+2d_0}n^{-2d_0}) \\ &= O_P(m^{-1/2}(\log m)\xi_{m,n}(d_0) + (m/n)^{\min(\varphi_y,2d_0)}), \end{aligned}$$

where $\xi_{k,n}(d)$ is defined in Lemma 5. The second equality above applies Lemma 5 and that $\tilde{\mathbf{X}}_{0,k} - \tilde{\mathbf{X}}_{0,k+1} = O(k^{-1})$ uniformly in $k = 1, \ldots, m$ and $\tilde{\mathbf{X}}_{0,m} - \frac{1}{m}\sum_{k=1}^{m}\tilde{\mathbf{X}}_{0,k} = O(1)$ (follows from approximating sums by integrals, see also AS p. 602). Thus $T_{1,n} = O_P((\log m)^{5/3}m^{-1/6} + (\log m)n^{-1/4} + (m/n)^{\min(\varphi_y,2d_0)})$ in the stationary case and $T_{1,n} = O_P((\log m)^{1+2/(5-4d_0)}m^{-(3-4d_0)/(10-8d_0)} + (\log m)^2 m^{2d_0-3/2} + (\log m)(\log n)^{5/4}n^{-1/2}m^{d_0/2} + (\log m)^{3/2}n^{-1/4}m^{d_0-1/2} + (m/n)^{\min(\varphi_y,2d_0)})$ in the nonstationary case. Since d_0 belongs to the interior of the parameter space it follows that $T_{1,n} = o_P(1)$.

To prove the result for $T_{2,n}$ we use Robinson's (1995*b*) Theorem 2, i.e., that $EI_y(\lambda_j)/f_y(\lambda_j) = 1 + O(j^{-1}(\log j))$ uniformly in $j = 1, \ldots, m$ in the stationary case, as well as Velasco's (1999) Theorem 1, $EI_y(\lambda_j)/f_y(\lambda_j) = 1 + O(j^{2d_0-2}(\log j))$ uniformly in $j = 1, \ldots, m$ in the nonstationary case. Note that, as in AS, the remainder terms are different from those of Robinson (1995*b*) and Velasco (1999) because of the normalization by $f_y(\lambda_j)$ rather than by $G_0\lambda_j^{-2d_0}$. Thus, as in the proof of Lemma 5 we can write

$$\begin{aligned} \frac{EI_z(\lambda_j)}{f_z(\lambda_j)} - 1 &= \frac{f_y(\lambda_j) - f_z(\lambda_j)}{f_z(\lambda_j)}\left(\frac{EI_y(\lambda_j)}{f_y(\lambda_j)} - 1\right) + \left(\frac{EI_y(\lambda_j)}{f_y(\lambda_j)} - 1\right) \\ &\quad + \frac{2\sqrt{f_y(\lambda_j)}}{f_z(\lambda_j)}\frac{E\,\mathrm{Re}(I_{yw}(\lambda_j))}{\sqrt{f_y(\lambda_j)}} + \frac{EI_w(\lambda_j) + f_y(\lambda_j) - f_z(\lambda_j)}{f_z(\lambda_j)}. \end{aligned}$$

Because $EI_w(\lambda_j) = f_w(\lambda_j) + O(j^{-1}(\log j))$ and $f_z(\lambda_j) - f_y(\lambda_j) = f_w(\lambda_j)$, the last term is $O(j^{-1}(\log j)\lambda_j^{2d_0})$. By the same reasoning and by independence of $\{y_t\}$ and $\{w_t\}$, the second to last term is $O_P(\lambda_j^{d_0} j^{-1}(\log j))$ in the stationary case and $O_P(\lambda_j^{d_0} j^{2d_0-2}(\log j))$ in the nonstationary case (see also the proof of Lemma 5 below and the second to last equation on p. 108 of Velasco (1999)). We thus obtain the bounds $EI_z(\lambda_j)/f_z(\lambda_j) - 1 = O(j^{-1}(\log j))$ for the stationary case and $EI_z(\lambda_j)/f_z(\lambda_j) - 1 = O(j^{2d_0-2}(\log j))$ for the nonstationary case, for all $j = 1, \ldots, m$. We also have that $f_z(\lambda_j)/g_j(d_0, \boldsymbol{\theta}_0) - 1 = O((j/n)^{\varphi_y} + (j/n)^{2d_0+\varphi_w})$ for all $j = 1, \ldots, m$ by (13). Therefore, in the stationary case, $T_{2,n}$ can be bounded similarly to (A.24) of AS,

$$\begin{aligned} T_{2,n} &= m^{-1/2} \sum_{j=1}^{m} O(j^{-1}(\log j)) O(1) O(\log m) \\ &= O\left(m^{-1/2}(\log m) \sum_{j=1}^{m} j^{-1}(\log j) \right) \\ &= O((\log m)^3 m^{-1/2}), \end{aligned}$$

using also that $|\tilde{\mathbf{X}}_{0,j} - \frac{1}{m}\sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k}| = O(\log m)$ uniformly in $j = 1, \ldots, m$. In the nonstationary case we find in the same way that

$$\begin{aligned} T_{2,n} &= m^{-1/2} \sum_{j=1}^{m} O(j^{2d_0-2}(\log j)) O(1) O(\log m) \\ &= O\left(m^{-1/2}(\log m) \sum_{j=1}^{m} j^{2d_0-2}(\log j) \right) \\ &= O((\log m)^3 m^{2d_0-3/2}). \end{aligned}$$

In both the stationary and nonstationary cases, $T_{2,n}$ is $o(1)$ since $d_0 < d_2 < 3/4$.

The proof for $T_{4,n}$ follows from summation by parts and the approximation $f_z(\lambda_j)/g_j(d_0, \boldsymbol{\theta}_0) - 1 = O((j/n)^{\varphi_y} + (j/n)^{2d_0+\varphi_w})$ for all $j = 1, \ldots, m$, which implies that

$$\begin{aligned} T_{4,n} &= m^{-1/2} \sum_{k=1}^{m-1} \left(\tilde{\mathbf{X}}_{0,k} - \tilde{\mathbf{X}}_{0,k+1} \right) \sum_{j=1}^{k} \left(\frac{f_z(\lambda_j)}{g_j(d_0, \boldsymbol{\theta}_0)} - 1 \right) \\ &\quad + \left(\tilde{\mathbf{X}}_{0,m} - \frac{1}{m} \sum_{k=1}^{m} \tilde{\mathbf{X}}_{0,k} \right) m^{-1/2} \sum_{j=1}^{m} \left(\frac{f_z(\lambda_j)}{g_j(d_0, \boldsymbol{\theta}_0)} - 1 \right) \\ &= m^{-1/2} \sum_{k=1}^{m-1} O(k^{-1}) \sum_{j=1}^{k} O((j/n)^{\varphi_y} + (j/n)^{2d_0+\varphi_w}) \\ &\quad + O(1) m^{-1/2} \sum_{j=1}^{m} O((j/n)^{\varphi_y} + (j/n)^{2d_0+\varphi_w}) \\ &= O(m^{1/2+\varphi_y} n^{-\varphi_y} + m^{1/2+2d_0+\varphi_w} n^{-2d_0-\varphi_w}). \end{aligned}$$

Condition (15) shows that this is $o_P(1)$.

C.4 Proof of (b), second statement

To prove the second statement of Lemma 3(b) we have to show that

$$\frac{1}{m}\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d,\boldsymbol{\theta}\right)}\right)(\tilde{\mathbf{X}}_j)_i(\tilde{\mathbf{X}}_j)_l,\ i,l=2,\ldots,R+2,$$

$$-\frac{1}{m}\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d,\boldsymbol{\theta}\right)}\right)(\tilde{\mathbf{X}}_j)_i\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}(\log\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))},\ i=2,\ldots,R_y+1,$$

$$\frac{1}{m}\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d,\boldsymbol{\theta}\right)}\right)(\tilde{\mathbf{X}}_j)_i2(\log\lambda_j)\left(1-\frac{h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}}{(1+h_j(d,\boldsymbol{\theta}))}\right),\ i=R_y+2,\ldots,R+2,$$

$$\frac{1}{m}\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d,\boldsymbol{\theta}\right)}\right)(\tilde{\mathbf{X}}_j)_{R_y+2}4h_w(\boldsymbol{\theta}_w,\lambda_j)\left(\log\lambda_j\right)^2\left(1-\frac{h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}}{(1+h_j(d,\boldsymbol{\theta}))}\right),$$

are all negligible when evaluated at $(d_0,\boldsymbol{\theta}_0)$. Note that it suffices to prove the result for the generic term

$$V_n\left(d,\boldsymbol{\theta}\right)=\frac{1}{m}\hat{G}\left(d,\boldsymbol{\theta}\right)^{-1}\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d,\boldsymbol{\theta}\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d,\boldsymbol{\theta}\right)}\right)(\tilde{\mathbf{X}}_j)_{R_y+2}q_j(d,\boldsymbol{\theta}),\tag{27}$$

where $q_j(d_0,\boldsymbol{\theta}_0)$ depends on j but is at most of order $O\left(\log n\right)$ and satisfies $q_{j+1}(d_0,\boldsymbol{\theta}_0)-q_j(d_0,\boldsymbol{\theta}_0)=O(j^{-1})$ uniformly in $j=1,\ldots,m$. Summation by parts on $V_n(d_0,\boldsymbol{\theta}_0)$ yields

$$\begin{aligned}V_n\left(d_0,\boldsymbol{\theta}_0\right) &= \frac{1}{m}\hat{G}\left(d_0,\boldsymbol{\theta}_0\right)^{-1}q_m(d_0,\boldsymbol{\theta}_0)\sum_{j=1}^{m}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d_0,\boldsymbol{\theta}_0\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d_0,\boldsymbol{\theta}_0\right)}\right)(\tilde{\mathbf{X}}_{0,j})_{R_y+2}\\ &\quad+\frac{1}{m}\hat{G}\left(d_0,\boldsymbol{\theta}_0\right)^{-1}\sum_{k=1}^{m-1}(q_k(d_0,\boldsymbol{\theta}_0)-q_{k+1}(d_0,\boldsymbol{\theta}_0))\sum_{j=1}^{k}\left(\frac{GI_z\left(\lambda_j\right)}{g_j\left(d_0,\boldsymbol{\theta}_0\right)}-\frac{1}{m}\sum_{k=1}^{m}\frac{GI_z\left(\lambda_k\right)}{g_k\left(d_0,\boldsymbol{\theta}_0\right)}\right)(\tilde{\mathbf{X}}_{0,j})_{R_y+2}\\ &= m^{-1}q_m(d_0,\boldsymbol{\theta}_0)O_P(m^{1/2})+m^{-1}\sum_{k=1}^{m-1}(q_k(d_0,\boldsymbol{\theta}_0)-q_{k+1}(d_0,\boldsymbol{\theta}_0))O_P(k^{1/2})\\ &= O_P\left(m^{-1/2}(\log n)+m^{-1/2}\right),\end{aligned}$$

where the second equality follows from part (e) of the lemma.

C.5 Proof of (c)

First we prove the result for $\mathbf{H}_{1n}$, where we need to show that

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\tilde{G}_{a,b_0,c}(d_0,\boldsymbol{\theta})-\tilde{G}_{a,b_0,c}(d_0,\boldsymbol{\theta}_0)\right|=o_P((\log m)^{-2})$$

for $a,c=0,1,2$ and $b=0,1,\ldots,2R_y,d,d+1,\ldots,d+R_w+R_y,2d,2d+1,\ldots,2d+2R_w$. By the triangle inequality it suffices to show that

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\tilde{G}_{a,b_0,c}(d_0,\boldsymbol{\theta})-\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta})\right|+\sup_{\boldsymbol{\theta}\in\Theta}\left|\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta})-\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta}_0)\right|+\tilde{\Delta}_{a,b_0,c}=o_P((\log m)^{-2}).\tag{28}$$

We showed in (25) that $\tilde{\Delta}_{a,b_0,c} = o_P((\log m)^{-2})$.

Following the proof of (25), the first term on the left-hand side of (28) is

$$\sup_{\boldsymbol{\theta}\in\Theta}\left| m^{-1}\sum_{j=1}^{m}\lambda_j^{2d_0}I_z(\lambda_j)\left[\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}))^{c+1}}\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d_0}\log\lambda_j}{(1+h_j(d_0,\boldsymbol{\theta}))}\right)^a-(2\log j)^a\right]\left(\frac{j}{m}\right)^{2b_0}\right|$$
$$= m^{-1}\sum_{j=1}^{m}\lambda_j^{2d_0}I_z(\lambda_j)\left[\frac{1}{1+O((j/n)^{2d_0})}\left(2\log j-\frac{O((j/n)^{2d_0}\log n)}{1+O((j/n)^{2d_0})}\right)^a-(2\log j)^a\right]\left(\frac{j}{m}\right)^{2b_0}$$

by Lemma 4(i), which proves the result for the first term of (28) by the same arguments as those applied to (25).

For n sufficiently large, $g_j(d_0,\boldsymbol{\theta}_0) > 0$ for all $j = 1,\ldots,m$, and then the second term on the left-hand side of (28) is

$$\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\Theta}\left| m^{-1}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\left(\frac{g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta})}-1\right)(2\log j)^a\left(\frac{j}{m}\right)^{2b_0}\right| \\
= &\sup_{\boldsymbol{\theta}\in\Theta,j=1,\ldots,m}\left|\frac{g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta})}-1\right| m^{-1}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}(2\log j)^a\left(\frac{j}{m}\right)^{2b_0} \\
= &\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta}_0)\sup_{\boldsymbol{\theta}\in\Theta,j=1,\ldots,m}\left|\frac{h_j(d_0,\boldsymbol{\theta}_0)-h_j(d_0,\boldsymbol{\theta})}{1+h_j(d_0,\boldsymbol{\theta})}\right|,
\end{aligned}$$

noting that all the terms inside the summation on the right-hand side of the second equality are positive. From Lemma 4(ii) and the fact that $\hat{G}_{a,b_0}(d_0,\boldsymbol{\theta}_0) = O_P((\log m)^a)$ by (24), it thus follows that the second term on the left-hand side of (28) is $O_P((\log m)^a(1+o(1))^{-1}\lambda_m^{2d_0})$, which proves (28).

Next we prove the result for $\mathbf{H}_{2n}$. Again, it suffices to show the result for the generic term $V_n(d,\boldsymbol{\theta})$ defined in (27), i.e. we must show that $\sup_{\boldsymbol{\theta}\in\Theta}|V_n(d_0,\boldsymbol{\theta}) - V_n(d_0,\boldsymbol{\theta}_0)| = o_P(1)$. By (24) and (28) we have that

$$\sup_{\boldsymbol{\theta}\in\Theta}\hat{G}(d_0,\boldsymbol{\theta}) = G(1+o_P((\log m)^{-2})), \tag{29}$$

and $\sup_{\boldsymbol{\theta}\in\Theta}|V_n(d_0,\boldsymbol{\theta}) - V_n(d_0,\boldsymbol{\theta}_0)|$ is, apart from a term that is negligible uniformly in $\boldsymbol{\theta}$,

$$\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}-\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\right)\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}q_j(d_0,\boldsymbol{\theta})\right. \\
&\left.-\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)}\right)\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}_0))}q_j(d_0,\boldsymbol{\theta}_0)\right| \\
\leq\ &\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}q_j(d_0,\boldsymbol{\theta})-\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}_0))}q_j(d_0,\boldsymbol{\theta}_0)\right)\right| \quad (30) \\
&+\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{1}{m}\sum_{k=1}^{m}\left(\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)}\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\left(\frac{j}{m}\right)^{2d_0}\right| \quad (31)
\end{aligned}$$

By the triangle inequality, (30) is bounded by

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\left(\frac{j}{m}\right)^{2d_0}\left(\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\right| \tag{32}$$

$$+\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d_0}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}-\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\right)\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\right|. \tag{33}$$

Note that, by inspection of the definition of $q_j(d,\boldsymbol{\theta})$ in (27) we have two prototypical expressions for the difference appearing in (32),

$$\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}=\begin{cases} O\left(\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right), \\ O\left(4\left(h_w(\boldsymbol{\theta}_w,\lambda_j)-h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\right)(m/n)^{2d_0}\left(\log\lambda_j\right)^2\right). \end{cases} \tag{34}$$

Inserting the first term of (34) into (32) we obtain, since for n sufficiently large $g_j(d_0,\boldsymbol{\theta}_0)>0$ for all $j=1,\ldots,m$,

$$\begin{aligned}(32) &= O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\left(\frac{j}{m}\right)^{4d_0}\left(\frac{(1+h_j(d_0,\boldsymbol{\theta}_0))-(1+h_j(d_0,\boldsymbol{\theta}))}{(1+h_j(d_0,\boldsymbol{\theta}))(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\right|\right)\\ &= O_P\left(\hat{G}_{0,2d_0}(d_0,\boldsymbol{\theta}_0)\sup_{\boldsymbol{\theta}\in\Theta,j=1,\ldots,m}\left|\frac{h_j(d_0,\boldsymbol{\theta}_0)-h_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))(1+h_j(d_0,\boldsymbol{\theta}_0))}\right|\right),\end{aligned}$$

and by Lemma 4(ii) it follows that (32) is $O_P(\lambda_m^{2d_0})$ in this case. Inserting the second term of (34) into (32) we obtain

$$\begin{aligned}(32) &= O_P\left(\hat{G}_{0,d_0}(d_0,\boldsymbol{\theta}_0)\sup_{\boldsymbol{\theta}\in\Theta,j=1,\ldots,m}\left|\left(\left(h_w(\boldsymbol{\theta}_w,\lambda_j)-h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\right)(m/n)^{2d_0}\left(\log\lambda_j\right)^2\right)\right|\right)\\ &= O_P\left((m/n)^{2d_0}(\log n)^2\right)\end{aligned}$$

by compactness of Θ. It follows that (32) is $o_P(1)$. Applying summation by parts to (33) we get the bound

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{q_m(d_0,\boldsymbol{\theta})}{(1+h_m(d_0,\boldsymbol{\theta}))}\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d_0}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}-\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\right)\right|$$

$$+\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{k=1}^{m-1}\left(\frac{q_k(d_0,\boldsymbol{\theta})}{(1+h_k(d_0,\boldsymbol{\theta}))}-\frac{q_{k+1}(d_0,\boldsymbol{\theta})}{(1+h_{k+1}(d_0,\boldsymbol{\theta}))}\right)\sum_{j=1}^{k}\left(\frac{j}{m}\right)^{2d_0}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}-\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}\right)\right|,$$

where the first term is

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{q_m(d_0,\boldsymbol{\theta})}{(1+h_m(d_0,\boldsymbol{\theta}))}\frac{1}{G}\left(\hat{G}_{0,d_0}(d_0,\boldsymbol{\theta})-\hat{G}_{0,d_0}(d_0,\boldsymbol{\theta}_0)\right)\right|=o_P\left((\log n)(\log m)^{-2}\right)$$

by (24), (28), and $\sup_{\boldsymbol{\theta}\in\Theta} q_m(d_0,\boldsymbol{\theta}) = O(\log n)$, and the second term is

$$
\begin{aligned}
& O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{k=2}^{m-1}\left(\frac{q_k(d_0,\boldsymbol{\theta})}{(1+h_k(d_0,\boldsymbol{\theta}))}-\frac{q_{k+1}(d_0,\boldsymbol{\theta})}{(1+h_{k+1}(d_0,\boldsymbol{\theta}))}\right)\left(\frac{k}{m}\right)^{2d_0}k(\log k)^{-2}\right|\right) \\
= & O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{k=2}^{m-1}\left(\frac{q_k(d_0,\boldsymbol{\theta})-q_{k+1}(d_0,\boldsymbol{\theta})}{(1+h_k(d_0,\boldsymbol{\theta}))}\right)\left(\frac{k}{m}\right)^{2d_0}k(\log k)^{-2}\right|\right) \\
& +O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{k=2}^{m-1}q_{k+1}(d_0,\boldsymbol{\theta})\left(\frac{1}{(1+h_k(d_0,\boldsymbol{\theta}))}-\frac{1}{(1+h_{k+1}(d_0,\boldsymbol{\theta}))}\right)\left(\frac{k}{m}\right)^{2d_0}k(\log k)^{-2}\right|\right),
\end{aligned}
$$

which, using $q_k(d_0,\boldsymbol{\theta})-q_{k+1}(d_0,\boldsymbol{\theta}) = O(k^{-1})$ and $q_{k+1}(d_0,\boldsymbol{\theta}) = O(\log n)$ for any $\boldsymbol{\theta}$, is

$$
\begin{aligned}
& O_P\left(\frac{1}{m}\sum_{k=2}^{m-1}\left(\frac{k}{m}\right)^{2d_0}(\log k)^{-2}+\frac{1}{m}\sum_{k=1}^{m-1}(\log n)\left(\lambda_{k+1}^{2d_0}-\lambda_k^{2d_0}\right)\left(\frac{k}{m}\right)^{2d_0}k(\log k)^{-2}\right) \\
= & O_P\left(\frac{1}{m}\sum_{k=2}^{m-1}\left(\frac{k}{m}\right)^{2d_0}(\log k)^{-2}+(\log n)\lambda_m^{2d_0}\frac{1}{m}\sum_{k=1}^{m-1}\left(\frac{k}{m}\right)^{2d_0}(\log k)^{-2}\right) \\
= & O_P\left((\log m)^{-2}+(\log n)(\log m)^{-2}(m/n)^{2d_0}\right).
\end{aligned}
$$

Thus both terms of (33) are $o_P(1)$ under (15).

Along the same lines we rewrite (31) as

$$
\begin{aligned}
& \sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\left(\frac{j}{m}\right)^{2d_0}\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)}\right| \\
& +\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d_0}\frac{1}{m}\sum_{k=1}^{m}\left(\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}-\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)}\right)\right|
\end{aligned}
$$

and, using the definition of $\hat{G}_{a,b}(d,\boldsymbol{\theta})$, this is equal to

$$
\begin{aligned}
& \sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{\hat{G}_{0,0}(d_0,\boldsymbol{\theta}_0)}{G}\frac{1}{m}\sum_{j=1}^{m}\left(\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\left(\frac{j}{m}\right)^{2d_0}\right| \\
& +\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{G}\left(\hat{G}_{0,0}(d_0,\boldsymbol{\theta})-\hat{G}_{0,0}(d_0,\boldsymbol{\theta}_0)\right)\frac{1}{m}\sum_{j=1}^{m}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d_0}\right| \\
= & O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta}_0)}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\left(\frac{j}{m}\right)^{2d_0}\right|\right) \\
& +o_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|(\log m)^{-2}\frac{1}{m}\sum_{j=1}^{m}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d_0}\right|\right),
\end{aligned}
$$

where the second term is easily seen to be $o_P((\log m)^{-2}(\log n)) = o_P(1)$. By (34), the first term is

$$
\begin{aligned}
& O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{(j/m)^{4d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{(j/m)^{4d_0}}{(1+h_j(d_0,\boldsymbol{\theta}_0))}\right)\right|\right) \\
& +O_P\left(\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(h_w(\boldsymbol{\theta}_w,\lambda_j)-h_w(\boldsymbol{\theta}_{w,0},\lambda_j)\right)(m/n)^{2d_0}(\log\lambda_j)^2\left(\frac{j}{m}\right)^{2d_0}\right|\right) \\
= \; & O_P\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d_0}\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{h_j(d_0,\boldsymbol{\theta}_0)-h_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))\,(1+h_j(d_0,\boldsymbol{\theta}_0))}\right|\right) \\
& +O_P\left((m/n)^{2d_0}\frac{1}{m}\sum_{j=1}^{m}(\log\lambda_j)^2\left(\frac{j}{m}\right)^{2d_0}\right) \\
= \; & O_P\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d_0}\lambda_j^{2d_0}+(m/n)^{2d_0}(\log n)^2\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d_0}\right) \\
= \; & O_P\left((m/n)^{2d_0}(\log n)^2\right),
\end{aligned}
$$

using compactness of Θ and Lemma 4. This is $o_P(1)$ which proves part (c).

C.6 Proof of (d)

Again, we first prove the result for $\mathbf{H}_{1n}$ which follows if

$$
\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\tilde{G}_{a,b,c}(d,\boldsymbol{\theta})-\tilde{G}_{a,b_0,c}(d_0,\boldsymbol{\theta})\right| = o_P((\log m)^{-2}) \tag{35}
$$

for $a, c = 0, 1, 2$ and $b = 0, 1, \ldots, 2R_y, d, d+1, \ldots, d+R_w+R_y, 2d, 2d+1, \ldots, 2d+2R_w$. Defining

$$
\begin{aligned}
\tilde{E}_{a,b,c}(d,\boldsymbol{\theta}) &= \frac{1}{m}\sum_{j=1}^{m}\frac{j^{2d}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))^{c+1}}\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a\left(\frac{j}{m}\right)^{2b}, \\
\hat{E}_{a,b}(d,\boldsymbol{\theta}) &= \frac{1}{m}\sum_{j=1}^{m}\frac{j^{2d}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))}(2\log j)^a\left(\frac{j}{m}\right)^{2b},
\end{aligned}
$$

we need to show that, for all $a, c = 0, 1, 2$ and $b = 0, 1, \ldots, 2R_y, d, d+1, \ldots, d+R_w+R_y, 2d, 2d+1, \ldots, 2d+2R_w$,

$$
Z_{a,b,c}(\eta_n) := \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\tilde{E}_{a,b,c}(d,\boldsymbol{\theta})-\tilde{E}_{a,b_0,c}(d_0,\boldsymbol{\theta})\right| = o_P(n^{2d_0}(\log m)^{-2}),
$$

see also AS p. 600. Note that since b can take values including d, we distinguish between b and b_0 which are obviously the same in case $b = 0, 1, \ldots, 2R_y$. By the triangle inequality it is sufficient to show that

$$\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\tilde{E}_{a,b,c}(d,\boldsymbol{\theta})-\hat{E}_{a,b}(d,\boldsymbol{\theta})\right|+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\hat{E}_{a,b}(d,\boldsymbol{\theta})-\hat{E}_{a,b_0}(d_0,\boldsymbol{\theta})\right|$$

$$+\sup_{\boldsymbol{\theta}\in\Theta}\left|\hat{E}_{a,b_0}(d_0,\boldsymbol{\theta})-\tilde{E}_{a,b_0,c}(d_0,\boldsymbol{\theta})\right|$$

$$=: Z_{1,a,b,c}(\eta_n)+Z_{2,a,b,c}(\eta_n)+Z_{3,a,b_0,c}(\eta_n)=o_P(n^{2d_0}(\log m)^{-2}).$$

The result for $Z_{3,a,b_0,c}(\eta_n)$ follows from part (c) of the lemma since it does not depend on d.

For $Z_{2,a,b,c}(\eta_n)$ we find that

$$\begin{aligned}
& Z_{2,a,b,c}(\eta_n) \\
= & \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^m\left[\left(\frac{j^{2d}I_z(\lambda_j)}{1+h_j(d,\boldsymbol{\theta})}\right)\left(\frac{j}{m}\right)^{2b}-\left(\frac{j^{2d_0}I_z(\lambda_j)}{1+h_j(d_0,\boldsymbol{\theta})}\right)\left(\frac{j}{m}\right)^{2b_0}\right](2\log j)^a\right| \\
= & \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^m\left(j^{2d}-j^{2d_0}\right)I_z(\lambda_j)\frac{1}{1+h_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2b}(2\log j)^a\right| \\
& +\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^m j^{2d_0}I_z(\lambda_j)\left(\frac{1+h_j(d,\boldsymbol{\theta})}{1+h_j(d_0,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2b_0-2b}-1\right)\frac{1}{1+h_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2b}(2\log j)^a\right|.
\end{aligned}$$

Since Θ is compact and $0<d_1\le d\le d_2<\infty$, for n sufficiently large it holds that

$$\inf_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\ldots,m}}|1+h_j(d,\boldsymbol{\theta})|\ge c>0,\quad \sup_{b\ge 0,j=1,\ldots,m}|j/m|^{2b}=1. \tag{36}$$

Thus, the first term of $Z_{2,a,b,c}(\eta_n)$ is bounded by

$$\sup_{d\in D_m(\eta_n)}\left|c^{-1}\frac{1}{m}\sum_{j=1}^m j^{2d_0}I_z(\lambda_j)(2\log j)^a\left|j^{2d-2d_0}-1\right|\right|,$$

which is $o_P(n^{2d_0}(\log m)^{-2})$ as in (A.18) of AS.

The second term of $Z_{2,a,b,c}(\eta_n)$ is bounded by

$$\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^m j^{2d_0}I_z(\lambda_j)\left(\frac{1+h_j(d,\boldsymbol{\theta})}{1+h_j(d_0,\boldsymbol{\theta})}-1\right)\frac{1}{1+h_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2b}(2\log j)^a\right| \tag{37}$$

$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^m j^{2d_0}I_z(\lambda_j)\frac{1+h_j(d,\boldsymbol{\theta})}{1+h_j(d_0,\boldsymbol{\theta})}\left(\left(\frac{j}{m}\right)^{2b_0}-\left(\frac{j}{m}\right)^{2b}\right)\frac{1}{1+h_j(d,\boldsymbol{\theta})}(2\log j)^a\right| \tag{38}$$

and using (36) and Lemma 4(iii) we find that (37) is

$$o_P\left(\frac{1}{m}\sum_{j=1}^m j^{2d_0}I_z(\lambda_j)\lambda_m^{2d_1}(\log m)^a\right)=o_P\left(\frac{1}{m}\sum_{j=1}^m \lambda_j^{2d_0}I_z(\lambda_j)\left(\frac{n}{2\pi}\right)^{2d_0}\lambda_m^{2d_1}(\log m)^a\right).$$

Noting that $m^{-1}\sum_{j=1}^{m}\lambda_j^{2d_0}I_z(\lambda_j)=\hat{G}_{0,0}(d_0,\mathbf{0})=G(1+o_P((\log m)^{-2}))$, (37) is equal to $o_P(n^{2d_0}\lambda_m^{2d_1}(\log m)^a)=o_P(n^{2d_0}(\log m)^{-2})$. By the mean value theorem, $x^a=x^b+(a-b)x^{a^*}(\log x)$ for $a\le a^*\le b$ which implies that

$$\sup_{d\in D_m(\eta_n)}\left|\left(\frac{j}{m}\right)^{2b_0}-\left(\frac{j}{m}\right)^{2b}\right|=O\left(\sup_{d\in D_m(\eta_n)}(b_0-b)(\log m)\right)=O\left((\log m)^{-5}\eta_n(\log m)\right).$$

Thus, applying also (36) and Lemma 4(iii), (38) is

$$\begin{aligned}O_P\left(\frac{1}{m}\sum_{j=1}^{m}j^{2d_0}I_z(\lambda_j)(\log m)^{a-4}\eta_n\right) &= O_P\left(\eta_n n^{2d_0}(\log m)^{a-4}\hat{G}_{0,0}(d_0,\mathbf{0})\right)\\ &= O_P\left(\eta_n n^{2d_0}(\log m)^{a-4}\right)\\ &= o_P\left(n^{2d_0}(\log m)^{-2}\right)\end{aligned}$$

since $\eta_n=o(1)$ and $a\le 2$.

Next, $Z_{1,a,b,c}(\eta_n)$ is

$$\sup_{\substack{d\in D_m(\eta_n)\\ \boldsymbol{\theta}\in\Theta}}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{j^{2d_0}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))}j^{2d-2d_0}\left[\frac{1}{(1+h_j(d,\boldsymbol{\theta}))^c}\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a-(2\log j)^a\right]\left(\frac{j}{m}\right)^{2b}\right|.$$

If $a=0$ the result follows by (36), Lemma 4(iv), $\sup_{d\in D_m(\eta_n),j=1,\ldots,m}j^{2d-2d_0}=O(1)$, and $\frac{1}{m}\sum_{j=1}^{m}j^{2d_0}I_z(\lambda_j)=\hat{G}_{0,0}(d_0,\mathbf{0})(2\pi/n)^{-2d_0}=O_P(n^{2d_0})$. When $a\ge 1$ we apply the mean value theorem as in the proof of (25) such that

$$\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a-(2\log j)^a=O\left((\log j)^{a-1}\lambda_j^{2d}(\log n)\right)\tag{39}$$

uniformly in $\boldsymbol{\theta}\in\Theta$. We then bound $Z_{1,a,b,c}(\eta_n)$ as

$$\begin{aligned}&\sup_{\substack{d\in D_m(\eta_n)\\ \boldsymbol{\theta}\in\Theta}}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{j^{2d_0}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))}j^{2d-2d_0}\left[\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a-(2\log j)^a\right]\left(\frac{j}{m}\right)^{2b}\right|\\ &+\sup_{\substack{d\in D_m(\eta_n)\\ \boldsymbol{\theta}\in\Theta}}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{j^{2d_0}I_z(\lambda_j)}{(1+h_j(d,\boldsymbol{\theta}))}j^{2d-2d_0}\left(\frac{1}{(1+h_j(d,\boldsymbol{\theta}))^c}-1\right)\left(2\log j-\frac{2h_w(\boldsymbol{\theta}_w,\lambda_j)\lambda_j^{2d}\log\lambda_j}{(1+h_j(d,\boldsymbol{\theta}))}\right)^a\left(\frac{j}{m}\right)^{2b}\right|,\end{aligned}$$

where the first term is $\hat{G}_{0,0}(d_0,\mathbf{0})(2\pi/n)^{-2d_0}O_P((\log m)^{a-1}\lambda_m^{2d_1}(\log n))=o_P(n^{2d_0}(\log m)^{-2})$ by (36) and (39) and the second term is $\hat{G}_{0,0}(d_0,\mathbf{0})(2\pi/n)^{-2d_0}o_P(\lambda_m^{2d_1}(\log m)^a)=o_P(n^{2d_0}(\log m)^{-2})$ by (36), (39), and Lemma 4(iv).

We proceed to show that $\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\mathbf{B}_n^{-1}\left\|\mathbf{H}_{2n}(d,\boldsymbol{\theta})-\mathbf{H}_{2n}(d_0,\boldsymbol{\theta})\right\|\mathbf{B}_n^{-1}=o_P(1)$ or equivalently that $\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}|V_n(d,\boldsymbol{\theta})-V_n(d_0,\boldsymbol{\theta})|=o_P(1)$. Since we have shown (35) we have that

$\hat{G}(d,\boldsymbol{\theta}) \xrightarrow{P} G$ uniformly in $\boldsymbol{\theta} \in \Theta, d \in D_m(\eta_n)$, so we need to show that the following is $o_P(1)$:

$$\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}-\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d,\boldsymbol{\theta})}\right)\frac{(j/m)^{2d}}{(1+h_j(d,\boldsymbol{\theta}))}q_j(d,\boldsymbol{\theta})\right.$$
$$\left.-\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}-\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\right)\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}q_j(d_0,\boldsymbol{\theta})\right|$$
$$\leq \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\frac{q_j(d,\boldsymbol{\theta})}{(1+h_j(d,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d}-\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d_0}\right)\right| \tag{40}$$
$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{1}{m}\sum_{k=1}^{m}\left(\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d_0}-\frac{I_z(\lambda_k)}{g_k(d,\boldsymbol{\theta})}\frac{q_j(d,\boldsymbol{\theta})}{(1+h_j(d,\boldsymbol{\theta}))}\left(\frac{j}{m}\right)^{2d}\right)\right| \tag{41}$$

By the triangle inequality we get the bounds

$$(40)\leq \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}\left(\left(\frac{j}{m}\right)^{2d}-\left(\frac{j}{m}\right)^{2d_0}\right)\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\right| \tag{42}$$
$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2d}\left(\frac{g_j(d_0,\boldsymbol{\theta})}{g_j(d,\boldsymbol{\theta})}-1\right)\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\right| \tag{43}$$
$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2d}\left(\frac{q_j(d,\boldsymbol{\theta})}{(1+h_j(d,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\right)\right| \tag{44}$$

and

$$(41)\leq \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\left(\frac{j}{m}\right)^{2d_0}-\left(\frac{j}{m}\right)^{2d}\right)\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\right| \tag{45}$$
$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d}\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_k)}{g_k(d_0,\boldsymbol{\theta})}\left(1-\frac{g_k(d_0,\boldsymbol{\theta})}{g_k(d,\boldsymbol{\theta})}\right)\right| \tag{46}$$
$$+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d}\left(\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))}-\frac{q_j(d,\boldsymbol{\theta})}{(1+h_j(d,\boldsymbol{\theta}))}\right)\frac{1}{m}\sum_{k=1}^{m}\frac{I_z(\lambda_j)}{g_k(d,\boldsymbol{\theta})}\right| \tag{47}$$

The required results for (42) and (45) follow using the mean value theorem as in (38), whereas the results for (43) and (46) follow as in (37). For (44) and (47) we note that, by inspection of the definition of $q_j(d,\boldsymbol{\theta})$ in (27), c.f. (34), it is sufficient to show the result for

$$\begin{aligned}\frac{q_j(d,\boldsymbol{\theta})}{(1+h_j(d,\boldsymbol{\theta}))}-\frac{q_j(d_0,\boldsymbol{\theta})}{(1+h_j(d_0,\boldsymbol{\theta}))} &= \frac{(j/m)^{2d}}{(1+h_j(d,\boldsymbol{\theta}))}-\frac{(j/m)^{2d_0}}{(1+h_j(d_0,\boldsymbol{\theta}))}\\ &= \left(\frac{j}{m}\right)^{2d}\left(\frac{1}{(1+h_j(d,\boldsymbol{\theta}))}-\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}))}\right)\\ &\quad+\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\left(\frac{j}{m}\right)^{2d}-\left(\frac{j}{m}\right)^{2d_0}\right).\end{aligned}$$

Inserting this into (44) ((47) follows the same way) we get the bound

$$\begin{aligned}(44) \quad \leq \quad & \sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{4d}\left(\frac{1}{(1+h_j(d,\boldsymbol{\theta}))}-\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}))}\right)\right|\\ &+\sup_{d\in D_m(\eta_n),\boldsymbol{\theta}\in\Theta}\left|\frac{1}{m}\sum_{j=1}^{m}\frac{I_z(\lambda_j)}{g_j(d,\boldsymbol{\theta})}\left(\frac{j}{m}\right)^{2d}\frac{1}{(1+h_j(d_0,\boldsymbol{\theta}))}\left(\left(\frac{j}{m}\right)^{2d}-\left(\frac{j}{m}\right)^{2d_0}\right)\right|,\end{aligned}$$

which we can handle similarly to (37) respectively (38).

Appendix D: Auxiliary lemmas

We now state two useful lemmas, which are used in the proofs of the main theorems. The first is stated without proof and gathers some properties of the function $h_j(d,\boldsymbol{\theta})$, which all follow by compactness of Θ.

Lemma 4 *Let* $h_j(d,\boldsymbol{\theta}) = h(d,\boldsymbol{\theta},\lambda_j) = \sum_{r=1}^{R_y}\theta_{y,r}\lambda^{2r} + \lambda_j^{2d}\sum_{r=0}^{R_w}\theta_{w,r}\lambda^{2r}$, $0<d_1<d_2<1$, *and let* Θ *be compact. Then, as* $n\to\infty$ *and for* $c=0,1,2$,

- *(i)* $\sup_{\boldsymbol{\theta}\in\Theta}|(1+h_j(d_0,\boldsymbol{\theta}))^{c+1}-1| = O(\sup_{\boldsymbol{\theta}\in\Theta}h_j(d_0,\boldsymbol{\theta})) = O((j/n)^{2d_0})$,
- *(ii)* $\inf_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\dots,m}}|1+h_j(d,\boldsymbol{\theta})| = 1+o(1)$ *and* $\sup_{\boldsymbol{\theta}\in\Theta,j=1,\dots,m}|h_j(d_0,\boldsymbol{\theta}_0)-h_j(d_0,\boldsymbol{\theta})| = O(\lambda_m^{2d_0})$,
- *(iii)* $\sup_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\dots,m}}\left|\frac{1+h_j(d,\boldsymbol{\theta})}{1+h_j(d_0,\boldsymbol{\theta})}-1\right| = O\left(\sup_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\dots,m}}\left|\frac{\theta_{r+1}(\lambda_j^{2d}-\lambda_j^{2d_0})}{1+h_j(d_0,\boldsymbol{\theta})}\right|\right) = O(\lambda_m^{2d_1})$,
- *(iv)* $\sup_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\dots,m}}|(1+h_j(d,\boldsymbol{\theta}))^c-1| = O(\sup_{\substack{d\in[d_1,d_2],\boldsymbol{\theta}\in\Theta\\ j=1,\dots,m}}h_j(d,\boldsymbol{\theta})) = O(\lambda_m^{2d_1})$.

The next lemma provides approximations of the periodogram of z_t by that of ε_t, following well known results from, e.g., Robinson (1995a), Velasco (1999), AS, and Hurvich et al. (2005).

Lemma 5 *Let Assumptions A1-A6 hold. Then, as* $n\to\infty$ *and for all* $k=1,\dots,m$,

$$\begin{aligned}&\sum_{j=1}^{k}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)\\ &\qquad= O_P\left(\xi_{k,n}(d_0)+k^{\varphi_y+1}n^{-\varphi_y}+k^{d_0+\varphi_w+1}n^{-d_0-\varphi_w}+k^{1+2d_0}n^{-2d_0}+k^{2d_0}n^{-d_0}(\log k)\right)\end{aligned}$$

and

$$\sum_{j=1}^{k}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)-E\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)\right)$$
$$=O_P\left(\xi_{k,n}(d_0)+k^{\varphi_y+1/2}n^{-\varphi_y}+k^{1/2+2d_0}n^{-2d_0}\right),$$

where

$$\xi_{k,n}(d)=k^{1/3}(\log k)^{2/3}+k^{1/2}n^{-1/4}$$

in the stationary case and

$$\xi_{k,n}(d)=k^{1/(5-4d)}(\log k)^{2/(5-4d)}+k^{2d-1}(\log k)+n^{-1/2}k^{(1+d)/2}(\log n)^{5/4}+n^{-1/4}k^{d}(\log k)^{1/2}$$

in the nonstationary case.

Proof. Note that, in the nonstationary case, Hurvich et al. (2005) examine the difference between the normalized periodograms of z_t and Δy_t (in our notation), whereas we examine the difference between the normalized periodograms of z_t and y_t itself in both the stationary and nonstationary cases.

Define $\tilde{g}_j(d,\boldsymbol{\theta})=\lambda_j^{-2d}G_0\left(1+h_y(\boldsymbol{\theta}_y,\lambda_j)\right)$ and write

$$\sum_{j=1}^{k}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)=\sum_{j=1}^{k}\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}\right) \tag{48}$$
$$+\sum_{j=1}^{k}\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right). \tag{49}$$

In the stationary case (49) is $O_P(k^{1/3}(\log k)^{2/3}+k^{\varphi_y+1}n^{-\varphi_y}+k^{1/2}n^{-1/4})$ by (A.13)(i) of AS, and in the nonstationary case (49) is $O_P(k^{1/(5-4d_0)}(\log k)^{2/(5-4d_0)}+k^{\varphi_y+1}n^{-\varphi_y}+k^{2d_0-1}(\log k)+n^{-1/2}k^{(1+d_0)/2}(\log n)^{5/4}+n^{-1/4}k^{d_0}(\log k)^{1/2})$ by slight modification of Lemma 1 of Velasco (1999) to account for the better approximation of $f_y(\lambda_j)$ by $\tilde{g}_j(d_0,\boldsymbol{\theta}_0)$ due to our polynomial appearing in $\tilde{g}_j(d_0,\boldsymbol{\theta}_0)$ (the required modification is the same as that used by AS to modify (4.8) of Robinson (1995*a*) to obtain their (A.13)(i)). The term (48) is

$$\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)}-\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}=\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)-g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)}\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-1\right) \tag{50}$$
$$+\frac{2\sqrt{h_w(\boldsymbol{\theta}_w,\lambda_j)}\sqrt{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}}{g_j(d_0,\boldsymbol{\theta}_0)}\frac{\operatorname{Re}(I_{yw}(\lambda_j))}{\sqrt{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}\sqrt{h_w(\boldsymbol{\theta}_w,\lambda_j)}} \tag{51}$$
$$+\frac{I_w(\lambda_j)+\tilde{g}_j(d_0,\boldsymbol{\theta}_0)-g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)}, \tag{52}$$

where $I_{ab}(\lambda) = \frac{1}{2\pi n}\sum_{t=1}^{n}\sum_{s=1}^{n} a_t b_s e^{i(s-t)\lambda}$ denotes the cross-periodogram between the two series a_t and b_t. Using summation by parts on (50) we find that

$$\begin{aligned}
&\sum_{j=1}^{k} \frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0) - g_j\,(d_0,\boldsymbol{\theta}_0)}{g_j\,(d_0,\boldsymbol{\theta}_0)} \left(\frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 1\right) \\
= \; & \sum_{j=1}^{k-1} \left(\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0) - g_j\,(d_0,\boldsymbol{\theta}_0)}{g_j\,(d_0,\boldsymbol{\theta}_0)} - \frac{\tilde{g}_{j+1}(d_0,\boldsymbol{\theta}_0) - g_{j+1}\,(d_0,\boldsymbol{\theta}_0)}{g_{j+1}\,(d_0,\boldsymbol{\theta}_0)}\right) \sum_{l=1}^{j} \left(\frac{I_y\,(\lambda_l)}{\tilde{g}_l(d_0,\boldsymbol{\theta}_0)} - 1\right) \\
& + \frac{\tilde{g}_k(d_0,\boldsymbol{\theta}_0) - g_k\,(d_0,\boldsymbol{\theta}_0)}{g_k\,(d_0,\boldsymbol{\theta}_0)} \sum_{j=1}^{k} \left(\frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 1\right),
\end{aligned}$$

which is $O_P((k/n)^{2d_0}(k^{1/3}(\log k)^{2/3} + k^{\varphi_y+1}n^{-\varphi_y} + k^{1/2}n^{-1/4} + k^{1/2}))$ in the stationary case whereas it is $O_P((k/n)^{2d_0}(k^{1/(5-4d_0)}(\log k)^{2/(5-4d_0)} + k^{\varphi_y+1}n^{-\varphi_y} + k^{2d_0-1}(\log k) + n^{-1/2}k^{(1+d_0)/2}(\log n)^{5/4} + n^{-1/4}k^{d_0}(\log k)^{1/2} + k^{1/2}))$ in the nonstationary case, by the same methods as applied previously and using also (4.9) of Robinson (1995*a*) and that $|\tilde{g}_j(d_0,\boldsymbol{\theta}_0)/g_j\,(d_0,\boldsymbol{\theta}_0) - 1| \leq C(j/n)^{2d_0}$. Next, (52) is easily seen to be $O_P((j/n)^{2d_0})$ because $E|I_w(\lambda_j)| = O_P(1)$ uniformly in $j = 1,\ldots,m$. Since $\{y_t\}$ and $\{w_t\}$ are independent (51) is $O_P((j/n)^{d_0}(j^{-1}(\log j) + (j/n)^{\min(\varphi_y,\varphi_w)}))$ in the stationary case by Theorem 2 of Robinson (1995*b*), yielding a contribution to (48) of $O_P((k/n)^{d_0}((\log k) + k^{1+\min(\varphi_y,\varphi_w)}n^{-\min(\varphi_y,\varphi_w)}))$. In the nonstationary case we use Theorem 1 of Velasco (1999) which shows that $\mathrm{Re}(I_{yw}(\lambda_j))|\tilde{g}_j\,(d_0,\boldsymbol{\theta}_0)\,|^{-1/2}|h_w(\boldsymbol{\theta}_w,\lambda_j)|^{-1/2} = O_P((j^{2d_0-2}(\log j) + (j/n)^{\min(\varphi_y,\varphi_w)}))$, yielding a contribution to (48) of $O_P((k/n)^{d_0}(k^{d_0}(\log k) + k^{1+\min(\varphi_y,\varphi_w)}n^{-\min(\varphi_y,\varphi_w)})$ (Velasco's result has to be modified to accommodate multivariate time series, but the modification is simple by comparing e.g. his equation (A.1) with equation (4.3) of Robinson (1995*b*), see also the second to last equation on p. 108 of Velasco (1999)). The difference in the remainder terms relative to Robinson (1995*b*) and Velasco (1999) is due to the different remainder term in the approximation of $f_y(\lambda_j)$ by $\tilde{g}_j\,(d_0,\boldsymbol{\theta}_0)$ due to our polynomial appearing in $\tilde{g}_j\,(d_0,\boldsymbol{\theta}_0)$.

To prove the second result we write

$$\begin{aligned}
&\sum_{j=1}^{k} \left(\frac{I_z\,(\lambda_j)}{g_j\,(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon\,(\lambda_j) - E\left(\frac{I_z\,(\lambda_j)}{g_j\,(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon\,(\lambda_j)\right)\right) \\
= \; & \sum_{j=1}^{k} \left(\frac{I_z\,(\lambda_j)}{g_j\,(d_0,\boldsymbol{\theta}_0)} - \frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - E\left(\frac{I_z\,(\lambda_j)}{g_j\,(d_0,\boldsymbol{\theta}_0)} - \frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}\right)\right) \qquad (53) \\
& + \sum_{j=1}^{k} \left(\frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon\,(\lambda_j) - E\left(\frac{I_y\,(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon\,(\lambda_j)\right)\right). \qquad (54)
\end{aligned}$$

By (A.21) of AS, (54) is $O_P(k^{1/3}(\log k)^{2/3} + k^{\varphi_y+1/2}n^{-\varphi_y} + k^{1/2}n^{-1/4})$ in the stationary case, and by (slight modification of) Lemma 1 of Velasco (1999), (54) is $O_P(k^{1/(5-4d_0)}(\log k)^{2/(5-4d_0)} + k^{\varphi_y+1/2}n^{-\varphi_y} + k^{2d_0-1}(\log k) + n^{-1/2}k^{(1+d_0)/2}(\log n)^{5/4} + n^{-1/4}k^{d_0}(\log k)^{1/2})$ in the nonstationary case. For eq. (53) we write

$$
\begin{aligned}
&\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - \frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - E\left(\frac{I_z(\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - \frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}\right) \\
= &\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0) - g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)} \left[\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j)\right) - E\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)} - 2\pi I_\varepsilon(\lambda_j)\right)\right] \quad (55) \\
&+\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0) - g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)} \left(2\pi I_\varepsilon(\lambda_j) - 1\right) \quad (56) \\
&+\frac{2\sqrt{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}}{g_j(d_0,\boldsymbol{\theta}_0)} \frac{\mathrm{Re}\left(I_{yw}(\lambda_j) - EI_{yw}(\lambda_j)\right)}{\sqrt{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}} \quad (57) \\
&+\frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} \left[\left(\frac{I_w(\lambda_j)}{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)} - 2\pi I_\eta(\lambda_j)\right) - E\left(\frac{I_w(\lambda_j)}{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)} - 2\pi I_\eta(\lambda_j)\right)\right] \quad (58) \\
&+\frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} \left(2\pi I_\eta(\lambda_j) - 1\right), \quad (59)
\end{aligned}
$$

using also that $h_w(\boldsymbol{\theta}_{w,0},\lambda_j) = g_j(d_0,\boldsymbol{\theta}_0) - \tilde{g}_j(d_0,\boldsymbol{\theta}_0)$.

Using summation by parts we find that (59) is

$$
\begin{aligned}
\sum_{j=1}^{k} \frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} \left(2\pi I_\eta(\lambda_j) - 1\right) &= \sum_{j=1}^{k-1} \left(\frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_j)}{g_j(d_0,\boldsymbol{\theta}_0)} - \frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_{j+1})}{g_{j+1}(d_0,\boldsymbol{\theta}_0)}\right) \sum_{l=1}^{j} \left(2\pi I_\eta(\lambda_l) - 1\right) \\
&\quad + \frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)} \sum_{j=1}^{k} \left(2\pi I_\eta(\lambda_j) - 1\right) \\
&= \sum_{j=1}^{k-1} \left| \frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_j) g_{j+1}(d_0,\boldsymbol{\theta}_0) - h_w(\boldsymbol{\theta}_{w,0},\lambda_{j+1}) g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)\, g_{j+1}(d_0,\boldsymbol{\theta}_0)} \right| O_P(j^{1/2}) \\
&\quad + \frac{h_w(\boldsymbol{\theta}_{w,0},\lambda_k)}{g_k(d_0,\boldsymbol{\theta}_0)} O_P(k^{1/2}) \\
&= O_P\left(\sum_{j=1}^{k-1} j^{2d_0 - 1/2} n^{-2d_0}\right) + O_P(k^{1/2+2d_0} n^{-2d_0}) \\
&= O_P(k^{1/2+2d_0} n^{-2d_0}),
\end{aligned}
$$

using (4.9) of Robinson (1995*a*) for the second equality. The term (56) is handled in exactly the same way yielding the same contribution. For the term (57) we can split it up in the same way as (58) and (59), and the contribution is the same.

Using summation by parts on (55) we find that, in the stationary case,

$$\sum_{j=1}^{k}\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)-g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)}\left[\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)-E\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)\right]$$

$$= \sum_{j=1}^{k-1}\left(\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)-g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)}-\frac{\tilde{g}_{j+1}(d_0,\boldsymbol{\theta}_0)-g_{j+1}(d_0,\boldsymbol{\theta}_0)}{g_{j+1}(d_0,\boldsymbol{\theta}_0)}\right)$$

$$\times\sum_{l=1}^{j}\left[\left(\frac{I_y(\lambda_l)}{\tilde{g}_l(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_l)\right)-E\left(\frac{I_y(\lambda_l)}{\tilde{g}_l(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_l)\right)\right]$$

$$+\frac{\tilde{g}_k(d_0,\boldsymbol{\theta}_0)-g_k(d_0,\boldsymbol{\theta}_0)}{g_k(d_0,\boldsymbol{\theta}_0)}\sum_{j=1}^{k}\left[\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)-E\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)\right]$$

$$= O_P\left((k/n)^{2d_0}(k^{1/3}(\log k)^{2/3}+k^{\varphi_y+1/2}n^{-\varphi_y}+k^{1/2}n^{-1/4})\right)$$

using (A.21) of AS. In the nonstationary case we use Lemma 1 of Velasco (1999) and get

$$\sum_{j=1}^{k}\frac{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)-g_j(d_0,\boldsymbol{\theta}_0)}{g_j(d_0,\boldsymbol{\theta}_0)}\left[\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)-E\left(\frac{I_y(\lambda_j)}{\tilde{g}_j(d_0,\boldsymbol{\theta}_0)}-2\pi I_\varepsilon(\lambda_j)\right)\right]$$

$$= O_P((k/n)^{2d_0}(k^{1/(5-4d_0)}(\log k)^{2/(5-4d_0)}+k^{\varphi_y+1/2}n^{-\varphi_y}$$
$$+k^{2d_0-1}(\log k)+n^{-1/2}k^{(1+d_0)/2}(\log n)^{5/4}+n^{-1/4}k^{d_0}(\log k)^{1/2})).$$

Finally the term (58) is handled in exactly the same way as the stationary case of (55) yielding the contribution $O_P\left((k/n)^{2d_0}(k^{1/3}(\log k)^{2/3}+k^{\varphi_w+1/2}n^{-\varphi_w}+k^{1/2}n^{-1/4})\right)$. ■

References

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modelling and forecasting realized volatility', *Econometrica* **71**, 579–625.

Andrews, D. W. K. & Guggenberger, P. (2003), 'A bias-reduced log-periodogram regression estimator for the long memory parameter', *Econometrica* **71**, 675–712.

Andrews, D. W. K. & Sun, Y. (2004), 'Adaptive local polynomial Whittle estimation of long-range dependence', *Econometrica* **72**, 569–614.

Arteche, J. (2004), 'Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models', *Journal of Econometrics* **119**, 131–154.

Baillie, R. T. (1996), 'Long memory processes and fractional integration in econometrics', *Journal of Econometrics* **73**, 5–59.

Baillie, R. T., Bollerslev, T. & Mikkelsen, H. O. (1996), 'Fractionally integrated generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics* **74**, 3–30.

Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman-Hall, New York.

Bollerslev, T. & Jubinski, D. (1999), 'Equity trading volume and volatility: Latent information arrivals and common long-run dependencies', *Journal of Business and Economic Statistics* **17**, 9–21.

Breidt, F. J., Crato, N. & de Lima, P. (1998), 'The detection and estimation of long memory in stochastic volatility', *Journal of Econometrics* **83**, 325–348.

Comte, F. & Renault, E. (1998), 'Long memory in continuous-time stochastic volatility models', *Mathematical Finance* **8**, 291–323.

Davies, R. B. & Harte, D. S. (1987), 'Tests for hurst effects', *Biometrika* **74**, 95–102.

Deo, R. S. & Hurvich, C. M. (2001), 'On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models', *Econometric Theory* **17**, 686–710.

Ding, Z., Granger, C. W. J. & Engle, R. F. (1993), 'A long memory property of stock returns and a new model', *Journal of Empirical Finance* **1**, 83–106.

Fox, R. & Taqqu, M. S. (1986), 'Large-sample properties of parameter estimates for strongly dependent stationary Gaussian series', *Journal of Time Series Analysis* **4**, 221–238.

Frederiksen, P. H. & Nielsen, M. Ø. (2007), 'Bias-reduced estimation of long memory stochastic volatility', *Working paper, Nordea Markets and Cornell University.*

Fuller, W. A. (1996), *Introduction to statistical time series*, Wiley, New York.

Geweke, J. & Porter-Hudak, S. (1983), 'The estimation and application of long-memory time series models', *Journal of Time Series Analysis* **4**, 221–238.

Harvey, A. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.

Harvey, A. (1998), Long memory in stochastic volatility, *in* J. Knight & S. Satchell, eds, 'Forecasting Volatility in Financial Markets', Butterworth-Heinemann, London, pp. 307–320.

Henry, M. & Zaffaroni, P. (2003), The long range paradigm for macroeconomics and finance, *in* P. Doukhan, G. Oppenheim & M. S. Taqqu, eds, 'Theory and Applications of Long-Range Dependence', Birkhäuser, Boston, pp. 417–438.

Hurvich, C. M., Moulines, E. & Soulier, P. (2005), 'Estimating long memory in volatility', *Econometrica* **73**, 1283–1328.

Hurvich, C. M. & Ray, B. K. (1995), 'Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes', *Journal of Time Series Analysis* **16**, 17–41.

Hurvich, C. M. & Ray, B. K. (2003), 'The local Whittle estimator of long-memory stochastic volatility', *Journal of Financial Econometrics* **1**, 445–470.

Künsch, H. R. (1987), Statistical aspects of self-similar processes, *in* Y. Prokhorov & V. V. Sazanov, eds, 'Proceedings of the First World Congress of the Bernoulli Society', VNU Science Press, Utrecht, pp. 67–74.

Liesenfeld, R. (2001), 'A generalized bivariate mixture model for stock price volatility and trading volume', *Journal of Econometrics* **104**, 141–178.

Marinucci, D. & Robinson, P. M. (1999), 'Alternative forms of fractional brownian motion', *Journal of Statistical Planning and Inference* **80**, 111–122.

Ray, B. K. & Tsay, R. (2000), 'Long-range dependence in daily stock volatility', *Journal of Business and Economic Statistics* **18**, 254–262.

Robinson, P. M. (1995*a*), 'Gaussian semiparametric estimation of long range dependence', *Annals of Statistics* **23**, 1630–1661.

Robinson, P. M. (1995*b*), 'Log-periodogram regression of time series with long range dependence', *Annals of Statistics* **23**, 1048–1072.

Robinson, P. M. (2003), Long-memory time series, *in* P. M. Robinson, ed., 'Time Series With Long Memory', Oxford University Press, Oxford, pp. 4–32.

Sun, Y. & Phillips, P. C. B. (2003), 'Nonlinear log-periodogram regression for perturbed fractional processes', *Journal of Econometrics* **115**, 355–389.

Velasco, C. (1999), 'Gaussian semiparametric estimation of non-stationary time series', *Journal of Time Series Analysis* **20**, 87–127.

Wright, J. H. (2002), 'Log periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns', *Econometric Reviews* **21**, 397–417.

Table 1: Simulation results for Model I

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.0109	0.2037	0.0184	0.2707	0.0124	0.2606	0.0135	0.2917
	4096	-0.0040	0.1280	0.0135	0.2093	-0.0029	0.1791	0.0085	0.2204
	8192	0.0061	0.0911	0.0128	0.1479	0.0026	0.1205	0.0103	0.1553
10	2048	0.0140	0.2639	0.0166	0.3106	0.0264	0.3123	0.0207	0.3271
	4096	0.0035	0.1793	0.0146	0.2462	0.0020	0.2268	0.0212	0.2606
	8192	0.0019	0.1194	0.0032	0.1570	0.0191	0.1901	0.0161	0.1947
20	2048	0.0004	0.3373	-0.0348	0.3391	-0.0097	0.3567	-0.0253	0.3524
	4096	-0.0005	0.2474	0.0003	0.2922	-0.0001	0.2840	0.0026	0.3023
	8192	-0.0047	0.2175	-0.0009	0.2392	0.0003	0.2380	-0.0001	0.2419
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.0002	0.1567	0.0002	0.2154	-0.0081	0.1953	-0.0139	0.2279
	4096	0.0015	0.0966	0.0076	0.1506	-0.0020	0.1233	0.0054	0.1759
	8192	0.0054	0.0706	0.0075	0.1025	0.0044	0.0907	0.0082	0.1244
10	2048	0.0057	0.2276	0.0056	0.2777	0.0094	0.2738	-0.0224	0.2735
	4096	0.0078	0.1399	0.0155	0.1930	0.0089	0.1774	-0.0095	0.1992
	8192	0.0047	0.0917	0.0125	0.1410	0.0034	0.1177	0.0002	0.1385
20	2048	-0.0152	0.3011	-0.0549	0.3058	-0.0294	0.3212	-0.1062	0.2997
	4096	0.0002	0.2201	-0.0055	0.2531	-0.0020	0.2518	-0.0445	0.2366
	8192	0.0073	0.1361	0.0146	0.1768	0.0073	0.1629	-0.0230	0.1629

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 2: Simulation results for Model II with $(\alpha_y, \beta_y) = (0, 0)$ and $(\alpha_x, \beta_x) = (0.5, 0)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	-0.0620	0.1977	0.0136	0.2678	-0.0045	0.2510	0.0042	0.2885
	4096	-0.0528	0.1346	0.0134	0.2116	-0.0131	0.1738	0.0080	0.2182
	8192	-0.0228	0.0931	0.0174	0.1523	0.0005	0.1250	0.0126	0.1595
10	2048	-0.0937	0.2484	0.0041	0.2968	0.0044	0.3052	0.0167	0.3263
	4096	-0.0661	0.1818	0.0170	0.2449	-0.0040	0.2235	0.0235	0.2653
	8192	-0.0416	0.1191	0.0144	0.1873	-0.0048	0.1600	0.0093	0.1906
20	2048	-0.1102	0.3165	-0.0479	0.3250	-0.0186	0.3525	-0.0271	0.3566
	4096	-0.1043	0.2482	-0.0071	0.2815	-0.0126	0.2819	0.0056	0.3004
	8192	-0.0613	0.1686	0.0102	0.2222	-0.0120	0.1966	0.0108	0.2248
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.1486	0.1869	-0.0524	0.2245	-0.0896	0.1881	-0.0381	0.2089
	4096	-0.1275	0.1519	-0.0116	0.1787	-0.0612	0.1319	-0.0023	0.1715
	8192	-0.1027	0.1206	-0.0049	0.1175	-0.0371	0.0944	0.0174	0.1309
10	2048	-0.2208	0.2570	-0.0709	0.2732	-0.1102	0.2537	-0.0693	0.2575
	4096	-0.1870	0.2138	-0.0149	0.2220	-0.0848	0.1736	-0.0260	0.1983
	8192	-0.1610	0.1806	-0.0100	0.1617	-0.0649	0.1284	0.0005	0.1525
20	2048	-0.2748	0.3104	-0.1417	0.2999	-0.1656	0.3127	-0.1423	0.2913
	4096	-0.2640	0.2919	-0.0583	0.2633	-0.1195	0.2556	-0.0728	0.2427
	8192	-0.2349	0.2554	-0.0155	0.2028	-0.0894	0.1800	-0.0252	0.1826

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 3: Simulation results for Model II with $(\alpha_y, \beta_y) = (0, 0)$ and $(\alpha_x, \beta_x) = (-0.8, 0)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.0133	0.2036	0.0205	0.2735	0.0138	0.2645	0.0126	0.2916
	4096	-0.0012	0.1285	0.0155	0.2072	0.0018	0.1825	0.0076	0.2189
	8192	0.0066	0.0904	0.0131	0.1501	0.0024	0.1214	0.0088	0.1564
10	2048	0.0149	0.2689	0.0156	0.3092	0.0241	0.3120	0.0179	0.3313
	4096	0.0046	0.1797	0.0144	0.2445	0.0012	0.2234	0.0227	0.2597
	8192	0.0037	0.1141	0.0076	0.1837	-0.0029	0.1575	0.0047	0.1899
20	2048	-0.0001	0.3340	-0.0329	0.3351	-0.0157	0.3509	-0.0213	0.3520
	4096	0.0032	0.2510	-0.0049	0.2858	-0.0025	0.2759	0.0048	0.2998
	8192	0.0067	0.1619	0.0095	0.2223	-0.0086	0.1973	0.0062	0.2258
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	0.0139	0.1595	-0.0034	0.2189	-0.0082	0.1978	-0.0137	0.2357
	4096	0.0114	0.0972	0.0055	0.1509	-0.0019	0.1256	0.0070	0.1782
	8192	0.0116	0.0716	0.0074	0.1028	0.0053	0.0903	0.0086	0.1257
10	2048	0.0291	0.2328	0.0065	0.2787	0.0092	0.2746	-0.0285	0.2753
	4096	0.0210	0.1404	0.0116	0.1936	0.0092	0.1799	-0.0048	0.2012
	8192	0.0046	0.0939	-0.0093	0.1324	-0.0053	0.1152	-0.0083	0.1398
20	2048	0.0060	0.3059	-0.0564	0.3084	-0.0331	0.3239	-0.1075	0.2998
	4096	0.0168	0.2232	-0.0132	0.2476	-0.0049	0.2478	-0.0493	0.2358
	8192	0.0119	0.1428	0.0045	0.1805	0.0005	0.1688	-0.0222	0.1704

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 4: Simulation results for Model III with $(\alpha_y, \beta_y) = (0, 0)$ and $(\alpha_x, \beta_x) = (0, 0.8)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	-0.0035	0.1535	0.0247	0.2290	0.0059	0.2064	0.0103	0.2503
	4096	-0.0083	0.0989	0.0095	0.1718	-0.0056	0.1410	0.0012	0.1858
	8192	0.0025	0.0719	0.0106	0.1180	0.0036	0.0990	0.0068	0.1340
10	2048	-0.0078	0.1848	0.0257	0.2653	0.0103	0.2462	0.0196	0.2828
	4096	-0.0064	0.1248	0.0114	0.1958	0.0008	0.1727	0.0097	0.2123
	8192	-0.0027	0.0867	0.0111	0.1453	0.0010	0.1200	0.0082	0.1564
20	2048	0.0006	0.2585	0.0118	0.3011	0.0196	0.3071	0.0194	0.3277
	4096	-0.0164	0.1666	0.0132	0.2412	-0.0000	0.2184	0.0123	0.2552
	8192	-0.0094	0.1141	0.0108	0.1776	-0.0034	0.1480	0.0093	0.1835
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.0396	0.1124	0.0051	0.1665	-0.0115	0.1366	-0.0084	0.1765
	4096	-0.0266	0.0758	0.0110	0.1151	-0.0017	0.0936	0.0056	0.1435
	8192	-0.0141	0.0547	0.0101	0.0808	0.0035	0.0701	0.0081	0.1042
10	2048	-0.0670	0.1585	0.0146	0.2235	-0.0071	0.1932	-0.0064	0.2174
	4096	-0.0371	0.0967	0.0226	0.1489	0.0004	0.1118	0.0020	0.1641
	8192	-0.0246	0.0695	0.0127	0.0998	0.0016	0.0830	0.0069	0.1217
20	2048	-0.0890	0.2183	-0.0099	0.2542	-0.0136	0.2503	-0.0406	0.2533
	4096	-0.0597	0.1479	0.0233	0.1955	0.0018	0.1701	-0.0029	0.2008
	8192	-0.0434	0.0935	0.0174	0.1341	-0.0022	0.1077	0.0032	0.1431

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 5: Simulation results for Model III with$(\alpha_y, \beta_y) = (0, 0)$ and $(\alpha_x, \beta_x) = (0, -0.8)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.0144	0.1427	0.0188	0.2184	0.0045	0.1921	0.0095	0.2345
	4096	0.0043	0.0917	0.0083	0.1621	-0.0036	0.1329	0.0034	0.1820
	8192	0.0079	0.0673	0.0085	0.1068	0.0035	0.0913	0.0074	0.1306
10	2048	0.0125	0.1713	0.0257	0.2513	0.0120	0.2291	0.0135	0.2626
	4096	0.0083	0.1159	0.0075	0.1831	-0.0031	0.1584	0.0056	0.1944
	8192	0.0065	0.0821	0.0085	0.1329	0.0010	0.1118	0.0046	0.1467
20	2048	0.0333	0.2349	0.0220	0.2936	0.0204	0.2831	0.0199	0.3066
	4096	0.0057	0.1470	0.0128	0.2208	0.0002	0.1940	0.0116	0.2380
	8192	0.0045	0.1041	0.0018	0.1508	-0.0044	0.1293	-0.0029	0.1585
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	0.0536	0.1144	-0.0099	0.1469	-0.0062	0.1293	0.0045	0.1891
	4096	0.0409	0.0785	-0.0031	0.0971	0.0028	0.0891	0.0067	0.1415
	8192	0.0319	0.0591	0.0029	0.0719	0.0074	0.0656	0.0074	0.0982
10	2048	0.0780	0.1555	-0.0036	0.1958	0.0005	0.1727	0.0042	0.2222
	4096	0.0602	0.1003	-0.0024	0.1208	0.0048	0.1044	0.0099	0.1582
	8192	0.0415	0.0715	-0.0044	0.0838	0.0024	0.0765	0.0064	0.1124
20	2048	0.1203	0.2265	-0.0163	0.2287	0.0076	0.2305	-0.0094	0.2436
	4096	0.0839	0.1455	-0.0021	0.1643	0.0092	0.1522	0.0088	0.1853
	8192	0.0529	0.0913	-0.0078	0.1031	0.0013	0.0952	-0.0003	0.1331

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 6: Simulation results for Model IV with $(\alpha_y, \beta_y) = (-0.8, 0)$ and $(\alpha_x, \beta_x) = (0.5, 0)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	-0.0579	0.1987	0.0167	0.2710	-0.0015	0.2515	0.0077	0.2867
	4096	-0.0498	0.1344	0.0150	0.2139	-0.0117	0.1766	0.0086	0.2186
	8192	-0.0207	0.0929	0.0177	0.1535	0.0007	0.1250	0.0067	0.1928
10	2048	-0.0900	0.2488	0.0029	0.2985	0.0051	0.3061	0.0175	0.3282
	4096	-0.0635	0.1812	0.0191	0.2492	-0.0021	0.2281	0.0242	0.2647
	8192	-0.0400	0.1186	0.0144	0.1859	-0.0047	0.1605	0.0106	0.1893
20	2048	-0.1076	0.3173	-0.0469	0.3274	-0.0186	0.3515	-0.0240	0.3569
	4096	-0.1017	0.2469	-0.0080	0.2827	-0.0108	0.2814	0.0038	0.3040
	8192	-0.0600	0.1685	0.0117	0.2239	-0.0121	0.1966	-0.0090	0.2250
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.1302	0.1759	-0.0543	0.2222	-0.0871	0.1894	-0.0439	0.2122
	4096	-0.1135	0.1414	-0.0153	0.1745	-0.0592	0.1324	-0.0006	0.1724
	8192	-0.0927	0.1125	-0.0062	0.1171	-0.0353	0.0942	0.0182	0.1316
10	2048	-0.2081	0.2491	-0.0619	0.2745	-0.1078	0.2542	-0.0682	0.2559
	4096	-0.1767	0.2052	-0.0157	0.2236	-0.0820	0.1731	-0.0259	0.1974
	8192	-0.1540	0.1747	-0.0110	0.1623	-0.0638	0.1275	-0.0001	0.1538
20	2048	-0.2699	0.3086	-0.1392	0.3017	-0.1656	0.3142	-0.1452	0.2919
	4096	-0.2587	0.2882	-0.0602	0.2637	-0.1175	0.2540	-0.0716	0.2416
	8192	-0.2297	0.2507	-0.0188	0.2041	-0.0873	0.1820	-0.0225	0.1838

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 7: Simulation results for Model IV with $(\alpha_y, \beta_y) = (-0.8, 0)$ and $(\alpha_x, \beta_x) = (-0.8, 0)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.0175	0.2051	0.0226	0.2725	0.0156	0.2640	0.0122	0.2922
	4096	0.0019	0.1289	0.0158	0.2095	0.0006	0.1838	0.0073	0.2222
	8192	0.0086	0.0911	0.0112	0.1455	0.0033	0.1206	0.0112	0.1568
10	2048	0.0184	0.2689	0.0156	0.3090	0.0243	0.3121	0.0188	0.3292
	4096	0.0068	0.1803	0.0149	0.2459	0.0025	0.2209	0.0232	0.2617
	8192	0.0052	0.1444	0.0090	0.1866	-0.0024	0.1582	0.0068	0.1874
20	2048	0.0021	0.3350	-0.0341	0.3342	-0.0170	0.3487	-0.0242	0.3523
	4096	0.0052	0.2520	-0.0047	0.2861	0.0016	0.2797	0.0051	0.3011
	8192	0.0077	0.1622	0.0046	0.2261	-0.0093	0.1947	0.0095	0.2292
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	0.0329	0.1646	-0.0054	0.2180	-0.0058	0.2001	-0.0116	0.2314
	4096	0.0248	0.1004	0.0032	0.1486	-0.0006	0.1259	0.0090	0.1805
	8192	0.0205	0.0740	0.0069	0.1028	0.0063	0.0904	0.0091	0.1262
10	2048	0.0439	0.2369	0.0062	0.2805	0.0105	0.2754	-0.0237	0.2722
	4096	0.0308	0.1431	0.0098	0.1916	0.0093	0.1790	-0.0036	0.1974
	8192	0.0133	0.0946	-0.0012	0.1327	-0.0043	0.1168	-0.0063	0.1415
20	2048	0.0116	0.3081	-0.0567	0.3087	-0.0321	0.3241	-0.1091	0.2989
	4096	0.0218	0.2246	-0.0142	0.2478	-0.0053	0.2485	-0.0472	0.2377
	8192	0.0170	0.1435	0.0031	0.1710	0.0029	0.1714	-0.0221	0.1710

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 8: Simulation results for Model V with $(\alpha_y, \beta_y) = (0, -0.8)$ and $(\alpha_x, \beta_x) = (0, 0.8)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.3067	0.3790	0.0179	0.2211	0.0883	0.2694	0.0270	0.2585
	4096	0.2261	0.2608	0.0037	0.1550	0.0492	0.1735	0.0133	0.1955
	8192	0.1655	0.1864	0.0130	0.1087	0.0561	0.1255	0.0117	0.1382
10	2048	0.2710	0.3675	0.0130	0.2572	0.0701	0.2928	0.0250	0.2833
	4096	0.2040	0.2525	0.0119	0.1846	0.0470	0.1894	0.0183	0.2168
	8192	0.1357	0.1681	0.0113	0.1307	0.0406	0.1319	0.0110	0.1542
20	2048	0.1837	0.3620	0.0203	0.3127	0.0539	0.3298	0.0355	0.3263
	4096	0.1635	0.2571	0.0219	0.2473	0.0408	0.2378	0.0286	0.2658
	8192	0.1057	0.1636	0.0074	0.1640	0.0222	0.1541	0.0151	0.1790
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.2596	0.4143	-0.0444	0.1989	0.1301	0.3058	0.0208	0.2102
	4096	0.2995*	0.4583*	0.0112	0.1083	0.1513	0.2153	0.0200	0.1476
	8192	0.3611	0.4097	0.0151	0.0686	0.1221	0.1514	0.0124	0.1089
10	2048	-0.2123	0.4204	-0.0190	0.1988	0.1460	0.3243	0.0265	0.2330
	4096	0.2645*	0.4359*	0.0167	0.1044	0.1464	0.2112	0.0208	0.1529
	8192	0.3123	0.3751	0.0136	0.0784	0.1095	0.1473	0.0156	0.1140
20	2048	-0.1827	0.4107	-0.0063	0.2145	0.1267	0.3235	-0.0173	0.2569
	4096	0.1994*	0.4026*	0.0168	0.1477	0.1245	0.2348	0.0164	0.1942
	8192	0.2554	0.3329	0.0175	0.1033	0.0990	0.1659	0.0215	0.1405

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 9: Simulation results for Model V with $(\alpha_y, \beta_y) = (0, -0.8)$ and $(\alpha_x, \beta_x) = (0, -0.8)$

		LWN		LPWN(1,0)		LPWN(0,1)		LPWN(1,1)	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	2048	0.3311	0.3899	0.0103	0.2054	0.0860	0.2616	0.0248	0.2486
	4096	0.2452	0.2754	0.0074	0.1487	0.0597	0.1721	0.0108	0.1887
	8192	0.1687	0.1876	0.0086	0.1016	0.0574	0.1213	0.0059	0.1305
10	2048	0.2993	0.3815	0.0073	0.2410	0.0736	0.2789	0.0266	0.2744
	4096	0.2229	0.2628	0.0036	0.1714	0.0440	0.1835	0.0209	0.2062
	8192	0.1579	0.1822	0.0144	0.1186	0.0504	0.1278	0.0167	0.1410
20	2048	0.2224	0.3741	0.0138	0.2809	0.0684	0.3174	0.0321	0.3066
	4096	0.1989	0.2618	0.0214	0.2085	0.0532	0.2115	0.0298	0.2339
	8192	0.1388	0.1764	0.0142	0.1524	0.0380	0.1461	0.0186	0.1655
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	2048	-0.2917	0.4077	-0.0846	0.2273	0.0918	0.3189	0.0174	0.2027
	4096	0.2863*	0.4677*	-0.0021	0.1165	0.1491	0.2215	0.0212	0.1431
	8192	0.3540	0.4284	0.0080	0.0645	0.1234	0.1510	0.0081	0.1032
10	2048	-0.2717	0.4115	-0.0677	0.2208	0.1099	0.3330	0.0313	0.2311
	4096	0.2509*	0.4533*	-0.0015	0.1216	0.1394	0.2185	0.0174	0.1553
	8192	0.3026	0.4056	0.0066	0.0696	0.1139	0.1471	0.0098	0.1123
20	2048	-0.2992	0.4025	-0.0528	0.2180	0.1012	0.3235	0.0004	0.2427
	4096	0.1423*	0.4261*	-0.0104	0.1397	0.1020	0.2170	-0.0166	0.1587
	8192	0.2319	0.3805	0.0014	0.0813	0.1010	0.1509	0.0136	0.1201

Note: The polynomial approximation used under the heading "LPWN(R_y, R_w)" is (R_y, R_w).

Table 10: Local Whittle estimation of long memory in volatility of DJIA stocks

	$m = \lfloor n^{0.6} \rfloor$			$m = \lfloor n^{0.7} \rfloor$			$m = \lfloor n^{0.8} \rfloor$		
Ticker Symbol	LW	LPW	LWN	LW	LPW	LWN	LW	LPW	LWN
AA	0.3002	0.3718	0.5292	0.2019	0.2956	0.5916	0.1379	0.1977	0.6063
	(0.0395)	(0.0592)	(0.0768)	(0.0258)	(0.0387)	(0.0476)	(0.0169)	(0.0253)	(0.0308)
AIG	0.3696	0.5017	0.6793	0.2938	0.3941	0.6202	0.2042	0.2990	0.6471
	(0.0395)	(0.0592)	(0.0686)	(0.0258)	(0.0387)	(0.0466)	(0.0169)	(0.0253)	(0.0299)
AXP	0.3691	0.4860	0.8225	0.3260	0.3928	0.5552	0.2115	0.3088	0.6514
	(0.0395)	(0.0592)	(0.0635)	(0.0258)	(0.0387)	(0.0490)	(0.0169)	(0.0253)	(0.0298)
BA	0.2593	0.4087	0.6809	0.2094	0.2484	0.5870	0.1509	0.2065	0.5336
	(0.0395)	(0.0592)	(0.0685)	(0.0258)	(0.0387)	(0.0478)	(0.0169)	(0.0253)	(0.0327)
C	0.3853	0.4789	0.7273	0.2908	0.3855	0.6677	0.2141	0.2992	0.6309
	(0.0395)	(0.0592)	(0.0666)	(0.0258)	(0.0387)	(0.0451)	(0.0169)	(0.0253)	(0.0302)
CAT	0.2477	0.3364	0.6247	0.1915	0.3053	0.5522	0.1280	0.2022	0.5788
	(0.0395)	(0.0592)	(0.0711)	(0.0258)	(0.0387)	(0.0491)	(0.0169)	(0.0253)	(0.0315)
DD	0.1425	0.1738	0.4238	0.1008	0.1292	0.4195	0.0810	0.0956	0.3366
	(0.0395)	(0.0592)	(0.0861)	(0.0258)	(0.0387)	(0.0565)	(0.0169)	(0.0253)	(0.0420)
DIS	0.3033	0.4448	0.9074	0.2361	0.3155	0.7582	0.1744	0.2134	0.6824
	(0.0395)	(0.0592)	(0.0613)	(0.0258)	(0.0387)	(0.0428)	(0.0169)	(0.0253)	(0.0292)
GE	0.3615	0.5044	0.7528	0.2497	0.3659	0.7796	0.1807	0.2580	0.7546
	(0.0395)	(0.0592)	(0.0657)	(0.0258)	(0.0387)	(0.0423)	(0.0169)	(0.0253)	(0.0281)
GM	0.2567	0.3489	0.4949	0.1987	0.2606	0.4890	0.1603	0.2027	0.4091
	(0.0483)	(0.0592)	(0.0794)	(0.0258)	(0.0387)	(0.0522)	(0.0169)	(0.0253)	(0.0375)
HD	0.3703	0.4581	0.6782	0.2489	0.3670	0.7321	0.1723	0.2432	0.7401
	(0.0395)	(0.0592)	(0.0686)	(0.0258)	(0.0387)	(0.0434)	(0.0169)	(0.0253)	(0.0283)
HON	0.2614	0.3354	0.9898	0.2323	0.2447	0.5859	0.1787	0.2253	0.4242
	(0.0395)	(0.0592)	(0.0594)	(0.0258)	(0.0387)	(0.0478)	(0.0169)	(0.0253)	(0.0368)
HPQ	0.3503	0.4688	0.8591	0.2366	0.3049	0.9061	0.1845	0.2290	0.7583
	(0.0395)	(0.0592)	(0.0625)	(0.0258)	(0.0387)	(0.0400)	(0.0169)	(0.0253)	(0.0280)
IBM	0.3417	0.4778	0.7626	0.2653	0.3295	0.6922	0.1931	0.2638	0.6359
	(0.0395)	(0.0592)	(0.0654)	(0.0258)	(0.0387)	(0.0444)	(0.0169)	(0.0253)	(0.0301)
INTC	0.3467	0.4755	0.7436	0.2396	0.3325	0.7685	0.1807	0.2532	0.6894
	(0.0395)	(0.0592)	(0.0661)	(0.0258)	(0.0387)	(0.0426)	(0.0169)	(0.0253)	(0.0291)
JNJ	0.3734	0.4400	0.6639	0.2601	0.3750	0.6850	0.1940	0.2641	0.6394
	(0.0395)	(0.0592)	(0.0692)	(0.0258)	(0.0387)	(0.0446)	(0.0169)	(0.0253)	(0.0301)
JPM	0.3603	0.5424	0.7173	0.3032	0.3741	0.6029	0.2174	0.2865	0.6058
	(0.0395)	(0.0592)	(0.0670)	(0.0258)	(0.0387)	(0.0472)	(0.0169)	(0.0253)	(0.0308)
KO	0.3677	0.5028	0.8104	0.2584	0.3653	0.8065	0.1833	0.2506	0.7923
	(0.0395)	(0.0592)	(0.0639)	(0.0258)	(0.0387)	(0.0418)	(0.0169)	(0.0253)	(0.0275)
MCD	0.2640	0.4591	0.6632	0.1798	0.2513	0.6936	0.1170	0.1701	0.7116
	(0.0395)	(0.0592)	(0.0693)	(0.0258)	(0.0387)	(0.0444)	(0.0169)	(0.0253)	(0.0287)
MMM	0.2635	0.3792	0.9891	0.2016	0.2744	0.9899	0.1430	0.1944	0.8712
	(0.0395)	(0.0592)	(0.0595)	(0.0258)	(0.0387)	(0.0388)	(0.0169)	(0.0253)	(0.0266)
MO	0.3041	0.4106	0.7409	0.2531	0.3152	0.5484	0.1879	0.2505	0.5163
	(0.0395)	(0.0592)	(0.0662)	(0.0258)	(0.0387)	(0.0493)	(0.0169)	(0.0253)	(0.0332)
MRK	0.2612	0.3504	0.5687	0.2063	0.2535	0.5034	0.1540	0.1930	0.4599
	(0.0395)	(0.0592)	(0.0742)	(0.0258)	(0.0387)	(0.0514)	(0.0169)	(0.0253)	(0.0352)
MSFT	0.3421	0.4756	0.8192	0.2908	0.3507	0.6156	0.2023	0.2883	0.6223
	(0.0395)	(0.0592)	(0.0636)	(0.0258)	(0.0387)	(0.0467)	(0.0169)	(0.0253)	(0.0304)
PFE	0.3354	0.3740	0.6473	0.2407	0.3168	0.6237	0.1644	0.2407	0.6324
	(0.0395)	(0.0592)	(0.0700)	(0.0258)	(0.0387)	(0.0465)	(0.0169)	(0.0253)	(0.0302)
PG	0.3262	0.4378	0.7656	0.2274	0.3433	0.7514	0.1944	0.2435	0.5525
	(0.0395)	(0.0592)	(0.0653)	(0.0258)	(0.0387)	(0.0429)	(0.0169)	(0.0253)	(0.0321)
SBC	0.3411	0.4017	0.7310	0.2692	0.3410	0.5545	0.1866	0.2581	0.5784
	(0.0395)	(0.0592)	(0.0665)	(0.0258)	(0.0387)	(0.0491)	(0.0169)	(0.0253)	(0.0315)
UTX	0.3435	0.4700	0.6515	0.2413	0.3426	0.6751	0.1650	0.2417	0.6892
	(0.0395)	(0.0592)	(0.0698)	(0.0258)	(0.0387)	(0.0449)	(0.0169)	(0.0253)	(0.0291)
VZ	0.3317	0.4262	0.8357	0.2578	0.3458	0.6661	0.1866	0.2642	0.6138
	(0.0395)	(0.0592)	(0.0631)	(0.0258)	(0.0387)	(0.0452)	(0.0169)	(0.0253)	(0.0306)
WMT	0.3728	0.4847	0.7582	0.2570	0.3477	0.7860	0.1668	0.2573	0.8247
	(0.0395)	(0.0592)	(0.0655)	(0.0258)	(0.0387)	(0.0422)	(0.0169)	(0.0253)	(0.0271)
XOM	0.2498	0.3889	0.6534	0.2271	0.2390	0.4419	0.1507	0.2254	0.5013
	(0.0395)	(0.0592)	(0.0697)	(0.0258)	(0.0387)	(0.0550)	(0.0169)	(0.0253)	(0.0337)

Note: Asymptotic standard errors in parentheses.

Table 11: LPWN estimation of long memory in volatility of DJIA stocks

	$m = \lfloor n^{0.6} \rfloor$			$m = \lfloor n^{0.7} \rfloor$			$m = \lfloor n^{0.8} \rfloor$		
Ticker Symbol	$(1,0)$	$(0,1)$	$(1,1)$	$(1,0)$	$(0,1)$	$(1,1)$	$(1,0)$	$(0,1)$	$(1,1)$
AA	0.5427	0.5525	0.5448	0.5196	0.5394	0.5247	0.5869	0.6027	0.5115
	(0.1139)	(0.1538)	(0.2315)	(0.0759)	(0.1012)	(0.1528)	(0.0469)	(0.0649)	(0.1007)
AIG	0.5872	0.6785	0.5887	0.6156	0.6068	0.6107	0.5833	0.6006	0.6209
	(0.1097)	(0.1487)	(0.2279)	(0.0701)	(0.0990)	(0.1483)	(0.0470)	(0.0650)	(0.0969)
AXP	0.9538	0.8216	0.9612	0.9719	0.6755	0.8239	0.5877	0.6057	0.7991
	(0.0903)	(0.1465)	(0.2198)	(0.0586)	(0.0975)	(0.1441)	(0.0469)	(0.0649)	(0.0946)
BA	0.5071	0.5606	0.5096	0.7064	0.6994	0.7474	0.6396	0.5770	0.7329
	(0.1177)	(0.1534)	(0.2351)	(0.0661)	(0.0971)	(0.1448)	(0.0451)	(0.0654)	(0.0951)
C	0.8962	0.8305	0.8982	0.7761	0.7036	0.7396	0.6577	0.6307	0.8016
	(0.0923)	(0.1465)	(0.2195)	(0.0636)	(0.0970)	(0.1449)	(0.0446)	(0.0645)	(0.0946)
CAT	0.6275	0.6235	0.6073	0.3601	0.3838	0.6125	0.4861	0.5134	0.4698
	(0.1065)	(0.1504)	(0.2266)	(0.0924)	(0.1118)	(0.1482)	(0.0514)	(0.0671)	(0.1030)
DD	0.4049	0.4667	0.4548	0.4507	0.4513	0.4495	0.4479	0.4397	0.3991
	(0.1325)	(0.1604)	(0.2424)	(0.0816)	(0.1060)	(0.1592)	(0.0536)	(0.0700)	(0.1084)
DIS	0.8803	0.9059	0.8788	0.9092	0.9898	0.9131	0.7966	0.8451	0.7870
	(0.0929)	(0.1463)	(0.2195)	(0.0600)	(0.0963)	(0.1441)	(0.0412)	(0.0630)	(0.0947)
GE	0.7370	0.7521	0.6984	0.7141	0.7467	0.7184	0.9888	0.7671	0.7187
	(0.0995)	(0.1472)	(0.2223)	(0.0658)	(0.0965)	(0.1453)	(0.0381)	(0.0632)	(0.0952)
GM	0.3761	0.4049	0.3799	0.4737	0.4881	0.4790	0.4820	0.4459	0.5354
	(0.1381)	(0.1677)	(0.2574)	(0.0796)	(0.1037)	(0.1564)	(0.0516)	(0.0697)	(0.0997)
HD	0.7040	0.6777	0.7078	0.6591	0.6800	0.6633	0.9888	0.7769	0.7015
	(0.1013)	(0.1487)	(0.2220)	(0.0681)	(0.0974)	(0.1465)	(0.0381)	(0.0631)	(0.0954)
HON	0.9893	0.9898	0.9890	0.7424	0.9854	0.9841	0.9418	0.5515	0.7446
	(0.0892)	(0.1467)	(0.2201)	(0.0648)	(0.0963)	(0.1445)	(0.0388)	(0.0660)	(0.0950)
HPQ	0.8452	0.8583	0.9214	0.9882	0.8807	0.9890	0.8530	0.9114	0.9206
	(0.0943)	(0.1464)	(0.2196)	(0.0583)	(0.0960)	(0.1445)	(0.0402)	(0.0630)	(0.0945)
IBM	0.7089	0.7619	0.7099	0.7851	0.7875	0.8044	0.7361	0.6765	0.7777
	(0.1011)	(0.1471)	(0.2219)	(0.0633)	(0.0962)	(0.1442)	(0.0425)	(0.0639)	(0.0947)
INTC	0.7274	0.7428	0.7266	0.7571	0.7736	0.7606	0.9879	0.7136	0.7686
	(0.1000)	(0.1473)	(0.2215)	(0.0643)	(0.0963)	(0.1447)	(0.0381)	(0.0635)	(0.0948)
JNJ	0.8592	0.8002	0.8639	0.6528	0.6438	0.7825	0.6879	0.6691	0.6793
	(0.0937)	(0.1467)	(0.2195)	(0.0683)	(0.0981)	(0.1444)	(0.0437)	(0.0639)	(0.0958)
JPM	0.4847	0.5400	0.4904	0.7293	0.6681	0.7001	0.6627	0.6469	0.6157
	(0.1204)	(0.1546)	(0.2374)	(0.0652)	(0.0976)	(0.1456)	(0.0444)	(0.0642)	(0.0971)
KO	0.9505	0.8090	0.9498	0.8450	0.8158	0.8256	0.9896	0.8388	0.7714
	(0.0904)	(0.1466)	(0.2198)	(0.0616)	(0.0961)	(0.1441)	(0.0381)	(0.0630)	(0.0948)
MCD	0.3414	0.4715	0.3472	0.6579	0.6890	0.6678	0.9896	0.7150	0.6721
	(0.1461)	(0.1599)	(0.2665)	(0.0681)	(0.0972)	(0.1464)	(0.0381)	(0.0635)	(0.0959)
MMM	0.9874	0.9893	0.9879	0.9880	0.9008	0.9894	0.9845	0.9869	0.9647
	(0.0893)	(0.1467)	(0.2201)	(0.0583)	(0.0960)	(0.1445)	(0.0382)	(0.0632)	(0.0947)
MO	0.9526	0.7394	0.9335	0.9565	0.6506	0.7847	0.5946	0.5469	0.7824
	(0.0904)	(0.1474)	(0.2197)	(0.0589)	(0.0979)	(0.1444)	(0.0466)	(0.0661)	(0.0947)
MRK	0.5558	0.5643	0.5546	0.6002	0.5832	0.5917	0.5732	0.5396	0.5713
	(0.1126)	(0.1532)	(0.2306)	(0.0709)	(0.0997)	(0.1491)	(0.0474)	(0.0663)	(0.0983)
MSFT	0.8010	0.8184	0.8003	0.9278	0.7643	0.8381	0.6497	0.6093	0.7937
	(0.0963)	(0.1465)	(0.2200)	(0.0596)	(0.0964)	(0.1441)	(0.0448)	(0.0648)	(0.0946)
PFE	0.8072	0.9896	0.8145	0.7372	0.6778	0.7131	0.6184	0.6149	0.7040
	(0.0960)	(0.1467)	(0.2199)	(0.0649)	(0.0974)	(0.1454)	(0.0458)	(0.0647)	(0.0954)
PG	0.9260	0.7646	0.9320	0.6881	0.7028	0.6932	0.9086	0.6644	0.7960
	(0.0913)	(0.1470)	(0.2196)	(0.0668)	(0.0970)	(0.1458)	(0.0393)	(0.0640)	(0.0946)
SBC	0.9711	0.8535	0.9734	0.9480	0.6398	0.8381	0.6196	0.5861	0.7954
	(0.0898)	(0.1464)	(0.2200)	(0.0591)	(0.0982)	(0.1441)	(0.0458)	(0.0652)	(0.0946)
UTX	0.5862	0.5983	0.5865	0.6373	0.6438	0.6395	0.6574	0.6684	0.6621
	(0.1098)	(0.1515)	(0.2281)	(0.0691)	(0.0981)	(0.1473)	(0.0446)	(0.0640)	(0.0961)
VZ	0.9228	0.8350	0.8787	0.8688	0.7483	0.8166	0.7044	0.6157	0.7801
	(0.0914)	(0.1464)	(0.2195)	(0.0610)	(0.0965)	(0.1442)	(0.0433)	(0.0647)	(0.0947)
WMT	0.8249	0.7575	0.8242	0.8152	0.8149	0.8149	0.8994	0.7736	0.9865
	(0.0952)	(0.1471)	(0.2198)	(0.0624)	(0.0961)	(0.1442)	(0.0394)	(0.0631)	(0.0948)
XOM	0.4872	0.6519	0.4900	0.8676	0.6669	0.7009	0.4411	0.4415	0.6900
	(0.1201)	(0.1494)	(0.2374)	(0.0610)	(0.0976)	(0.1456)	(0.0540)	(0.0699)	(0.0956)

Note: The heading "(R_y, R_w)" indicates the LPWN(R_y, R_w) estimator. Asymptotic standard errors in parentheses.

Table 12: Parametric Whittle estimation of long memory in volatility of DJIA stocks

Ticker Symbol	$\hat{d}$	$\hat{\alpha}_y$	$\hat{\beta}_y$	$\hat{\sigma}^2_\eta$	$\hat{\alpha}_x$	$\hat{\beta}_x$	$\hat{\sigma}^2_\varepsilon$
AA	0.5777 (0.0943)	−0.5431 (0.1864)	–	0.1634 (0.1122)	0.0362 (0.0206)	–	3.3266 (0.1360)
AIG	0.6377 (0.0748)	−0.7946 (0.1447)	–	0.2771 (0.1629)	−0.8357 (0.0925)	0.9185 (0.3044)	2.8112 (0.7084)
AXP	0.5915 (0.0695)	–	−0.7416 (0.0863)	1.9245 (0.2775)	0.2959 (0.1167)	–	1.2799 (0.3486)
BA	0.5532 (0.1019)	−0.9275 (0.2558)	0.6585 (0.2779)	0.1152 (0.0844)	–	0.0584 (0.0215)	3.2629 (0.1125)
C	0.6201 (0.0671)	–	–	0.0913 (0.0482)	–	–	3.1170 (0.0866)
CAT	0.4444 (0.1982)	–	–	0.2211 (0.4432)	0.6814 (0.1389)	−0.7236 (0.1823)	3.2299 (0.5543)
DD	0.2478 (0.0978)	−0.7859 (0.3492)	0.8867 (0.3360)	0.7682 (0.8560)	–	–	4.1923 (0.7411)
DIS	0.7555 (0.1331)	–	–	0.0125 (0.0161)	–	0.0604 (0.0160)	3.3340 (0.0758)
GE	0.7509 (0.1177)	0.6409 (0.2936)	−0.8718 (0.0901)	0.1393 (0.1774)	–	–	3.1386 (0.1740)
GM	0.5040 (0.1438)	−0.9730 (0.0440)	0.9371 (0.1043)	0.1624 (0.2038)	0.6010 (0.2992)	−0.5739 (0.2931)	3.2918 (0.2419)
HD	0.6211 (0.1200)	0.4356 (0.0676)	−0.8670 (0.0453)	1.5042 (0.5819)	–	–	1.8557 (0.5603)
HON	0.4166 (0.0672)	−0.3490 (0.3161)	–	0.6309 (0.4581)	−0.6202 (0.3682)	0.6475 (0.3093)	2.7333 (0.4812)
HPQ	0.9298 (0.1684)	−0.9010 (0.1339)	–	0.0085 (0.0133)	0.6844 (0.1604)	−0.6567 (0.1633)	3.3724 (0.0762)
IBM	0.6775 (0.0978)	−0.6652 (0.1972)	–	0.1063 (0.0743)	0.0325 (0.0196)	–	3.1645 (0.1037)
INTC	0.7168 (0.0865)	−0.8816 (0.1020)	–	0.0712 (0.0557)	−0.9341 (0.0319)	0.9686 (0.0836)	3.0632 (0.3243)
JNJ	0.5824 (0.1038)	0.3924 (0.0799)	−0.7923 (0.0630)	0.9809 (0.7210)	–	–	2.4001 (0.6923)
JPM	0.5798 (0.0624)	–	–	0.1461 (0.0688)	–	–	3.1921 (0.1005)
KO	0.8234 (0.1214)	−0.7461 (0.2712)	–	0.0249 (0.0267)	–	0.0423 (0.0166)	3.2834 (0.0802)
MCD	0.6211 (0.1290)	0.5379 (0.1005)	−0.8949 (0.0414)	0.8105 (0.4203)	–	–	2.6296 (0.4025)
MMM	0.7032 (0.1748)	–	–	0.0128 (0.0236)	–	–	3.5654 (0.0871)
MO	0.5410 (0.0760)	0.6217 (0.3652)	–	0.0185 (0.0339)	–	0.0349 (0.0194)	3.2044 (0.0876)
MRK	0.4903 (0.0764)	–	–	0.1430 (0.0873)	–	–	3.2380 (0.1142)
MSFT	0.5987 (0.0725)	−0.7656 (0.1451)	–	0.2990 (0.1846)	−0.8105 (0.0887)	0.8940 (0.2491)	2.8675 (0.6008)
PFE	0.6093 (0.0850)	–	–	0.0777 (0.0428)	–	–	3.3014 (0.0872)
PG	0.5724 (0.0741)	–	–	0.0901 (0.0565)	–	–	3.1889 (0.0942)
SBC	0.5518 (0.0657)	–	–	0.1294 (0.0681)	–	–	3.2578 (0.1018)
UTX	0.6159 (0.1041)	0.5142 (0.0943)	−0.8598 (0.0475)	0.8469 (0.4559)	–	–	2.5056 (0.4314)
VZ	0.6778 (0.1022)	−0.5928 (0.3156)	–	0.1039 (0.0747)	–	0.0714 (0.0240)	3.2410 (0.1074)
WMT	0.8328 (0.1229)	–	–	0.0078 (0.0087)	–	0.0363 (0.0154)	3.4027 (0.0733)
XOM	0.4962 (0.0698)	–	–	0.1532 (0.0839)	–	–	3.2217 (0.1109)

Note: Asymptotic standard errors (evaluated as the inverse of the negative Hessian) in parentheses.

Research Papers 2008

CREATES
Center for Research in Econometric Analysis of Time Series

2008-16: Jie Zhu: FIEGARCH-M and and International Crises: A Cross-Country Analysis

2008-17: Almut E. D. Veraart: Inference for the jump part of quadratic variation of Itô semimartingales

2008-18: Michael Sørensen: Parametric inference for discretely sampled stochastic differential equations

2008-19: Anne Péguin-Feissolle, Birgit Strikholm and Timo Teräsvirta: Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form

2008-20: Stefan Holst Bache, Christian M. Dahl and Johannes Tang Kristensen: Determinants of Birthweight Outcomes: Quantile Regressions Based on Panel Data

2008-21: Ole E. Barndorff-Nielsen, José Manuel Corcuera, Mark Podolskij and Jeannette H.C. Woerner: Bipower variation for Gaussian processes with stationary increments

2008-22: Mark Podolskij and Daniel Ziggel: A Range-Based Test for the Parametric Form of the Volatility in Diffusion Models

2008-23: Silja Kinnebrock and Mark Podolskij: An Econometric Analysis of Modulated Realised Covariance, Regression and Correlation in Noisy Diffusion Models

2008-24: Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Small Bandwidth Asymptotics for Density-Weighted Average Derivatives

2008-25: Mark Podolskij and Mathias Vetter: Bipower-type estimation in a noisy diffusion setting

2008-26: Martin Møller Andreasen: Ensuring the Validity of the Micro Foundation in DSGE Models

2008-27: Tom Engsted and Thomas Q. Pedersen: Return predictability and intertemporal asset allocation: Evidence from a bias-adjusted VAR model

2008-28: Frank S. Nielsen: Local polynomial Whittle estimation covering non-stationary fractional processes

2008-29: Per Frederiksen, Frank S. Nielsen and Morten Ørregaard Nielsen: Local polynomial Whittle estimation of perturbed fractional processes