Article:
Iacone, Fabrizio orcid.org/0000-0002-2681-9036 and Lazarova, Stepana (2019) Semiparametric detection of changes in long range dependence. Journal of Time Series Analysis. 693–706. ISSN 1467-9892

https://doi.org/10.1111/jtsa.12448

Semiparametric detection of changes in long range dependence

Fabrizio Iacone*
Dipartimento di Economia, Management e Metodi Quantitativi,
Università degli Studi di Milano
and Department of Economics and Related Studies,
University of York

Štěpána Lazarová
Department of Economics, Queen Mary University of London
Mile End Road, London E1 4NS

14 August 2018

Abstract

We consider changes in the degree of persistence of a process when the degree of persistence is characterized as the order of integration of a strongly dependent process. To avoid the risk of incorrectly specifying the data generating process we employ local Whittle estimates which uses only frequencies local to zero. The limit distribution of the test statistic under the null is not standard but it is well known in the literature. A Monte Carlo study shows that this inference procedure performs well in finite samples. We demonstrate the practical utility of these results with an empirical example, where we analyse the inflation rate in Germany for the period 1986–2017.

Keywords: Long memory, persistence, break, local Whittle estimate.

JEL classification: C22.

*Financial support from the Economic and Social Research Council through grant R000239936 and from the Dennis Sargan Memorial Fund is gratefully aknowledged. We thank Peter M. Robinson and Robert Taylor and participants of the 14th Brazilian Time Series and Econometrics School and the Queen Mary University of London Econometrics Reading group for many helpful suggestions and discussions. We are grateful to two anonymous referees for useful comments. E-mail address: fabrizio.iacone@unimi.it, s.lazarova@qmul.ac.uk.

1 Introduction

We consider changes in the degree of persistence of a time series. We characterize the degree of persistence as the order of integration δ of a strongly dependent process. Changes in the order of integration have been documented in a number of macroeconomic and financial variables, such as output (De Long and Summers, 1988), the budget deficit (Hakkio and Rush, 1991), inflation (Halunga, Osborn, and Sensier, 2009, Kumar and Okimoto, 2007, Hassler and Meller, 2014). Financial studies include the analysis of financial market bubbles (Sollis, 2006, Frömmel and Kruse, 2012), international and sectoral bank equity index returns (Hassler, Rodrigues and Rubia, 2014), yield spreads of EMU government bonds (Sibbertsen, Wegener and Basse, 2014). Interest in the characterization of the degree of persistence and in its potential instability is particularly strong in the evaluation of macroeconomic policies such as inflation targeting because *ceteris paribus* a reduction of the order indicates a tighter control of the variable of interest (provided that the process is mean reverting, at least after the change). By the same argument, periods associated to $\delta = 1$ or larger indicate lack of control.

In early applied work it was assumed that δ was limited to integer numbers only (typically, $\delta = 0$ or $\delta = 1$). Tests to detect changes between these two states were developed by Kim (2000), Kim, Belaire-Franch and Badilli-Amador (2002), Busetti and Taylor (2004), Harvey, Leybourne and Taylor (2006), Leybourne, Taylor and Kim (2007) among others. In all these cases, the test statistics are based on ratios of partial sums and it is possible to detect a change in the order of integration because the limit distributions are well behaved under the null.

However, the assumption of integer δ seems particularly restrictive in the context of testing for a change in persistence because it leaves no alternative between fast reversion to the mean ($\delta = 0$) and no reversion at all ($\delta = 1$). Important variations in the long term dynamics may be represented with fractional changes in δ. Introducing a fractional δ allows to identify changes that could otherwise go unnoticed using a standard Dickey and Fuller type test, as for example the move from a mean-reverting but highly persistent fractional process to a unit root that was discussed by Frömmel and Kruse (2012). Moreover, the fractional nature of δ in this case also gives a measure of the size of the change. For example, in the work of Sibbertsen, Wegener and Basse (2014) one can see not only which countries were hit by the Euro area sovereign debt crisis, but also rank them to establish who was hit most heavily.

Testing for changes of non-integer δ was advocated by Beran and Terrin (1996), who recommended testing for a change in this parameter in the context of a fully parametric model. Horváth and Shao (1999) further developed this approach. Inference based on a fully parametric model is appealing because of good asymptotic properties of the maximum likelihood estimators. However, the requirement that the user specifies the correct model for the data generating process may be inconvenient, especially when a large number of parameters has to be considered, because the uncertainty about the model may adversely reflect on the result of the procedure. To avoid this risk, a number of semiparametric techniques for inference for δ were developed. The case for semiparametric estimation of δ is even more compelling in case the process is subject to a break. The uncertainty about the possibility of a break should make the researcher even less confident when formulating a fully parametric model, because the model selection procedure must be designed to deliver the correct model even under the alternative hypothesis that a break has indeed taken place.

A modified approach to testing for a change in persistence has been followed by Sibbertsen and Kruse (2009) who simulated appropriate critical values for the test statistic in Leybourne et al. (2007). Their critical values depend on δ. A non-parametric approach was adopted by Lavancier, Leipus, Philippe and Surgailis (2013) who proposed a modification of the test statistic of Kim (2000) and related statistics. Semiparametric detection of a break has been considered by Shimotsu (2006), who however assumed that the potential breakpoint is fixed in advance. Our choice is closer to the latter in the sense of being semiparametric but, like in the parametric test of Horváth and Shao (1999), we estimate δ before and after a potential break point, and compute a Wald type statistic for the difference between the two estimators. Since the potential break-point is in fact unknown we derive the limit distribution of the supremum of the Wald type statistic. However, unlike Horváth and Shao (1999), we estimate δ by local Whittle estimator, so our procedure does not require us to specify a complete parametric model and it is therefore robust to this type of misspecification. We find that under the null the limit distribution is well known and does not depend on δ.

The structure of the paper is as follows. In Section 2 we present the relevant asymptotic theory and in Section 3 we analyze the small sample properties with a Monte Carlo exercise. We present an application in Section 4 and we conclude in Section 5. The proofs of the theorems are to be found in the Appendix.

2 Testing for a change in the order of integration

To establish notation we first introduce the model for the case of a stationary process. Our model is similar to the model of Robinson (1995). For a stationary process x_t

with covariance $\gamma_s = E\left[(x_t - Ex_0)(x_{t+s} - E(x_0))\right]$ and spectral density $f(\lambda)$ such that $\gamma_s = \int_{-\pi}^{\pi} f(\lambda) e^{i\lambda s} d\lambda$, we consider a process as integrated of order δ, denoted $x_t \in I(\delta)$, if there is $\delta < 1/2$ and $G \in (0, \infty)$ such that

$$f(\lambda) \sim G\lambda^{-2\delta} \text{ as } \lambda \to 0^+, \tag{1}$$

where notation $a \sim b$ is used to indicate that the ratio a/b tends to 1. In model (1) the order of integration δ is usually the parameter of interest and G depends on all the other parameters of the spectral density.

For example, if x_t is ARFIMA(p, δ, q), $\Phi(L)\Delta^{\delta} x_t = \Theta(L)\varepsilon_t$, for ε_t independent and identically distributed with $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$, then $2\pi G = \Phi(1)^{-1}\Theta(1)\sigma_\varepsilon^2$. In comparison with such full parametric specification, the model in (1) is usually considered semiparametric.

We now introduce the local Whittle estimator and discuss how to use it to test for a change in δ when the process is subject to a break in the order of integration. For a generic time series x_t observed at times $t = 1, ..., T$, define the Fourier transform $w(\lambda) = \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{T} x_t e^{-i\lambda t}$ and the periodogram $I(\lambda) = |w(\lambda)|^2$. The local Whittle estimator is computed by minimizing with respect to d the loss function

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

where $\lambda_j = \frac{2\pi j}{T}$, for integers $j = 1, ..., m$, are Fourier frequencies and m is a user-chosen parameter. This loss function is discussed by Robinson (1995).

For a stationary process x_t, parameter δ in (1) does not depend on time. In practice, the persistence of a process may be subject to change over time. We consider a situation where the persistence measure δ can change at a certain point in time. Let $\lfloor x \rfloor$ denote the integer part of a real number x. We assume that there exists a break fraction τ^* with $0 < \tau^* < 1$ such that for $t < \lfloor \tau^* T \rfloor$, x_t is drawn from an $I(\delta_1)$ process, and for $t \geq \lfloor \tau^* T \rfloor$, x_t is a realization of an $I(\delta_2)$ process with $\delta_1 \neq \delta_2$. That is, at different points in time the series x_t is observed from two possibly different processes, x_{1t} which is $I(\delta_1)$ and x_{2t} which is $I(\delta_2)$, with $x_t = x_{1t}$ if $t < \lfloor \tau^* T \rfloor$ and $x_t = x_{2t}$ if $t \geq \lfloor \tau^* T \rfloor$. If $\delta_1 = \delta_2$, it is possible that x_{1t} and x_{2t} are generated by the same process. We wish to test the hypothesis of stability of the persistence. Our hypotheses of interest are therefore

$$H_0 : \delta_1 = \delta_2,$$
$$H_A : \delta_1 \neq \delta_2.$$

In order to test whether the parameter δ remained stable over the sample period, we estimate δ on two subsamples and compare the two estimators. For a time series sample x_t observed at times $t = 1$, ..., T, and an interval $[\sigma, \tau] \subset [0, 1]$, we define the Fourier transform and the periodogram of series $0, \ldots, 0, x_{\lfloor \sigma T \rfloor + 1}, \ldots, x_{\lfloor \tau T \rfloor}, 0, \ldots, 0$ as

$$w_{\sigma\tau}(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=\lfloor \sigma T \rfloor + 1}^{\lfloor \tau T \rfloor} x_t e^{-i\lambda t} \qquad \text{and} \qquad I_{\sigma\tau}(\lambda) = |w_{\sigma\tau}(\lambda)|^2$$

and the related local Whittle loss function as

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

We select τ in $(0, 1)$ and estimate parameter δ on for intervals $[0, \tau]$ and $[\tau, 1]$. Let

$$\widehat{\delta}_1(\tau) = \arg\min_{d \in [\Delta_1, \Delta_2] \subset (-1/2, 1/2)} R(d, I_{0\tau}), \tag{4}$$

$$\widehat{\delta}_2(\tau) = \arg\min_{d \in [\Delta_1, \Delta_2] \subset (-1/2, 1/2)} R(d, I_{\tau 1}), \tag{5}$$

so $\widehat{\delta}_1(\tau)$ and $\widehat{\delta}_2(\tau)$ are the estimators computed using only the first or the second part of the sample for a given τ. Given the estimators $\widehat{\delta}_1(\tau)$ and $\widehat{\delta}_2(\tau)$ of δ on the two subsamples, we can base a test statistic for the test of stability of δ on the normalized difference of the two estimators. We define the test statistic as

$$\widehat{t}(\tau) = \sqrt{4\tau(1-\tau)m}\left(\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)\right). \tag{6}$$

For any given $\tau \in (0, 1)$, it can be showed that under regularity conditions, as $T \to \infty$, test statistic $\widehat{t}(\tau)$ converges in distribution to a standard normal,

$$\widehat{t}(\tau) \to_d N(0, 1). \tag{7}$$

As the potential location $\lfloor \tau T \rfloor$ of the break is usually unknown, we consider $\widehat{t}(\tau)$ for all τ in a closed subset $[\tau_l, \tau_h]$ of $(0,1)$. Following Andrews (1993) we introduce the test statistic $\widehat{t}^2$ defined as

$$\widehat{t}^2 = \sup_{\tau \in [\tau_l, \tau_h] \subset (0,1)} \widehat{t}(\tau)^2 . \tag{8}$$

We establish weak convergence of $\widehat{t}(\tau)$ to a tight limit. This convergence together with the continuous mapping theorem then gives us the distribution of the $\widehat{t}^2$ test statistic under the null hypothesis.

Our analysis proceeds under the following assumptions.

Let $\mathcal{F}_t$ be the σ-algebras of events generated by ε_s, $s \leq t$.

Assumption 1 *The processes x_{1t} and x_{2t} have linear representation*

$$x_{\ell t} - E(x_{\ell t}) = \sum_{j=0}^{\infty} \alpha_{\ell j} \varepsilon_{t-j}, \qquad \ell = 1, 2,$$

where $\sum_{j=0}^{\infty} \alpha_{\ell j}^2 < \infty$ and, for $p = 1, ..., 8$,

$$E(\varepsilon_t^p | \mathcal{F}_{t-1}) = \omega_p < \infty \quad a.s., \qquad t = 0, \pm 1, \ldots,$$

$\omega_1 = 0$ and $\omega_2 = \sigma_\varepsilon^2$.

Assumption 2 *In a neighbourhood $(0, \varepsilon)$ of the origin, $A_\ell(\lambda) = \sum_{j=0}^{\infty} \alpha_{\ell j} e^{-ij\lambda}$ are differentiable for $\ell = 1, 2$ and*

$$\frac{d}{d\lambda} A_\ell(\lambda) = O\left(\frac{|A_\ell(\lambda)|}{\lambda}\right) \qquad as\ \lambda \to 0+ .$$

Assumption 3 *For some $\beta \in (0, 2]$, the spectral densities f_1 and f_2 satisfy*

$$f_1(\lambda) \sim G\lambda^{-2\delta_1}\left(1 + O\left(\lambda^\beta\right)\right) \qquad as\ \lambda \to 0+,$$
$$f_2(\lambda) \sim G\lambda^{-2\delta_2}\left(1 + O\left(\lambda^\beta\right)\right) \qquad as\ \lambda \to 0+,$$

where $G \in (0, \infty)$ and $\delta_1, \delta_2 \in [\Delta_1, \Delta_2] \subset [-1/2, 1/2]$.

Assumption 4 *As $T \to \infty$,*

$$\frac{1}{m} + \frac{m^{1+2\beta} \log^2 m}{T^{2\beta}} \to 0.$$

Let $B(\tau)$ be a standard Brownian motion process on $[0,1]$ and let "$\Rightarrow$" denote weak convergence in the Skorokhod topology. We obtain the following theorem.

Theorem 1 *Under Assumptions 1–4 and under the null hypothesis, for* $[\tau_l, \tau_h] \subset (0,1)$,

$$\widehat{t}^2 \Rightarrow \sup_{\tau \in [\tau_l, \tau_h]} \frac{(B(\tau) - \tau B(1))^2}{\tau (1-\tau)} \tag{9}$$

as $T \to \infty$.

Proof of Theorem 1 is provided in Section 6.2 of the Appendix.

The limit process $\sup_{\tau \in [\tau_l, \tau_h]} \frac{(B(\tau)-\tau B(1))^2}{4\tau(1-\tau)}$ is the supremum over $[\tau_l, \tau_h]$ of the square of a standardized tied down Bessel process. The distribution of the test statistic is identical to the distribution obtained by Andrews (1993), who also discusses what happens when $[\tau_l, \tau_h] = [0, 1]$. Andrews (1993) provides tables of various quantiles for the distribution. The upper 5% quantile is 8.85 when $[\tau_l, \tau_h] = [0.15, 0.85]$ and 9.31 when $[\tau_l, \tau_h] = [0.1, 0.9]$.

We can test $H_0 : \delta_1 = \delta_2$ against $H_A : \delta_1 \neq \delta_2$ at size α by computing the $\widehat{t}^2$ statistic and comparing its value with the upper α quantile. A value of the $\widehat{t}^2$ statistic in excess of the critical value leads to a rejection of H_0.

The following theorem shows that the test is consistent. With increasing sample size, the power of the test approaches 1 in probability.

Theorem 2 *Under Assumptions 1–4 and under the alternative hypothesis, for* $[\tau_l, \tau_h] \subset (0,1)$,

$$\widehat{t}^2 \xrightarrow{p} \infty$$

as $T \to \infty$.

Proofs of Theorem 2 can be found in Section 6.3 of the Appendix.

Remark 1. Assumptions 1–4 are based on the assumptions of Robinson (1995) who uses them to establish consistency and limit normality of the local Whittle estimator. The most notable difference is that in our case finite moments up to the eight order are needed instead of Robinson's fourth moments. This is because of the additional requirement of establishing tightness in the context of our problem of interest.

Remark 2. The statistic $\widehat{t}^2$ is related to the test statistic of Horváth and Shao (1999), where however δ is estimated within a fully parametric model.

Remark 3. If the location of the breakpoint is known in advance it seems natural to test for a break using the statistic $\widehat{t}^2(\tau^*)$ using critical values from the χ^2_1 distribution. When the potential breakpoint is not known, the statistic $\widehat{t}^2(\tau)$ for a user chosen point may still be considered. This is similar to the test advocated by Shimotsu (2006) who suggests to divide $[\tau_l, \tau_h]$ in equally spaced intervals. However, testing using the statistic $\widehat{t^2}(\tau)$ may result in low power when compared to testing using the $\widehat{t^2}$ statistic. To understand why, consider the case $\delta_1 > \delta_2$ and $\tau < \tau^*$. Then observations $x_1, \ldots, x_{\lfloor \tau T \rfloor}$ are obtained from a $I(\delta_1)$ process whereas a part of observations $x_{\lfloor \tau T \rfloor + 1}, \ldots, x_T$ comes from a $I(\delta_1)$ and a part from a $I(\delta_2)$ process. Therefore the periodogram of series $0, \ldots, 0, x_{\lfloor \tau T \rfloor + 1}, \ldots, x_T$ has features similar to those of the periodogram of a signal plus noise process with signal $I(\delta_1)$ and noise $I(\delta_2)$. Dalla, Giraitis and Hidalgo (2006) have shown that in this case $\widehat{\delta}_2(\tau) \to_p \delta_1$. In their Theorem 3, these authors discuss the conditions under which the estimator $\widehat{\delta}_2(\tau)$ of δ_1 may be subject to a lower order bias. This bias would at least warrant some power to a test using the $\widehat{t^2}(\tau)$ statistic.

3 A Monte Carlo exercise

The results of Section 2 are asymptotic. We therefore examine the performance of the proposed test procedure in finite samples.

In the first exercise, summarized in Table 1, we study the size of the test under a range of data generating processes (DGP) and bandwidths. As the test statistic is based on local Whittle estimates $\widehat{\delta}_1$ and $\widehat{\delta}_2$, we choose the DGP and bandwidths bearing in mind existing results for the local Whittle estimate.

For the DGP, we consider the model

$$(1 - \phi L)(1 + \theta L)^{-1} \Delta^{\delta} x_t = \varepsilon_t,$$

where ε_t is independently distributed with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = 1$.

Two sources of lower order bias can affect the local Whittle estimate of δ. The first source of bias is due to the approximation of the factor $\left|1 - e^{-i\lambda}\right|^{2\delta}$ in the spectral density of x_t by $\lambda^{2\delta}$ in the local Whittle loss function. The second one is due to the curvature of the spectrum of $\Delta^{\delta} x_t$ which in the local Whittle loss function is approximated as constant. Both approximations become less appropriate as we include frequencies further away from 0 in the local Whittle estimation stage. For example, if $\phi > 0$ and $\theta = 0$ then the local Whittle estimate of δ is subject to a positive bias which, for given m and T, is stronger the larger is ϕ. For given δ and ϕ, both sources of bias become stronger the larger m is for a given T, so simulations with larger bandwidths are potentially more at risk of size distortion.

We first consider five cases where our assumptions are satisfied. We assume that ε_t is normally distributed. First, $\phi = 0$, $\theta = 0$, $\delta = 0$, so that x_t is a normally independently distributed process. We consider this as a benchmark case, in the sense that this is the most favourable situation for local Whittle estimation and it should therefore have the best size properties for the test $\widehat{t^2}$ too. Second, $\phi = 0$, $\theta = 0$, $\delta = 0.4$, so x_t is a fractional noise. This is still a fairly favourable situation, as the spectral density of $\Delta^\delta x_t$ is constant, and therefore any size distortion should be primarily due to the approximation in the local Whittle loss function of the spectral density of x_t by $\lambda^{-2\delta}$. Third, $\phi = 0.5$, $\theta = 0$, $\delta = 0$, so that x_t is an AR(1) process. Fourth, $\phi = 0.8$, $\theta = 0$, $\delta = 0$, so that x_t is also an AR(1) process but with higher ϕ coefficient. Both cases introduce a different possible source of size distortion because the spectral density of x_t is not constant for frequencies not close to 0. Treating it as constant as in the local Whittle loss function may therefore generate size distortion in the estimates $\widehat{\delta}_1$ and $\widehat{\delta}_2$, especially when larger bandwidths are selected. The choice of two different parameters for ϕ is interesting as the larger is the parameter, the less appropriate it is to approximate the spectral density of $\Delta^\delta x_t$ as constant, therefore incurring a larger risk of size distortion for given bandwidth. Fifth, $\phi = 0$, $\theta = 0.8$, $\delta = 0$, so that x_t is an MA(1) process. In this case, we again consider a process with spectral density that is not constant. This may again generate size distortion, especially when larger bandwidths are used. As the spectral densities of the AR(1) and of the MA(1) processes are different, considering them both allows an interesting comparison.

In the remaining cases we explore situations that are of great practical interest but are not included in our theoretical framework. The sixth case that we consider, therefore, is a non-stationary fractional noise, with $\delta = 0.7$ with ε_t normally distributed. Such a process does not satisfy the assumptions of our model, but Velasco (1999) showed that results for the local Whittle estimate may be extended to include cases up to $\delta < 0.75$ without affecting the asymptotic properties of the estimate. As this requires a more extensive theoretical treatment, we have decided not to consider this range of δ formally. However, such values of the memory parameter δ may suit some empirical applications and it is therefore interesting to explore them in simulation. In the seventh case we set $\phi = 0$, $\theta = 0$, $\delta = 0$ but we generate ε_t as a t_5 distributed variate, so the moment condition from the assumptions is not met. Finally, in the last two exercises we take x_t to be normally independently distributed process but we do not assume that the process is observable anymore. Instead, we consider two cases for observables z_t defined as $z_t = \alpha_0 + \alpha_1 t + x_t$ or as $z_t = \alpha_0 + \alpha_1 DU(1/2)_t + x_t$, where $DU(1/2)_t = 0$ if $t < \lfloor 1/2T \rfloor$ and $DU(1/2)_t = 1$ if $t > \lfloor 1/2T \rfloor$. These two cases include two realistic situations in which either a trend or a change in the mean (at a known point) is fitted. Again, we did not allow for theses two cases theoretically in the interest of brevity. However, we note that Abadir, Distaso and Giraitis (2011, page 190), refer to sufficient conditions in Dalla, Giraitis and Hidalgo (2006) to see that asymptotic properties of the local Whittle estimate are not affected when regression residuals from a trend are used. The same argument could be used to justify using regression residuals on a broken mean.

Regarding the bandwidths, the choices $m = \lfloor T^{0.5} \rfloor$ and $m = \lfloor T^{0.65} \rfloor$ have emerged as popular for the local Whittle estimation. In particular, Abadir, Distaso and Giraitis (2007) have found that the latter bandwidth gives a good MSE performance in a range of situations. MSE-optimal bandwidths for the local Whittle estimate are of the type $m = \lfloor \alpha T^{0.8} \rfloor$ where α depends on the curvature of the spectrum of $\Delta^{\delta} x_t$, see for example Henry (2001). Assumption 4 requires that $m/T^{0.8} \to 0$ so $m = \lfloor T^{0.79} \rfloor$ is the largest bandwidth consistent with this assumption among the ones we considered. Our selection of candidate bandwidths is based partly on the choice already considered in the literature. It should however be noted that our problem is different. We are not interested in minimum MSE estimation of the local Whittle estimate of δ but rather in correct size when testing the null hypothesis of no break using the $\widehat{t}^2$ test. Minimum MSE bandwidth realizes a compromise between the lower order bias and the variance of the local Whittle estimate. These two factors, bias and variance, impact on tests in a different way. The former may cause size distortion, the latter loss of power. Thus, a good MSE performance may not be very relevant if we are interested in testing at the correct size. In our case of interest, that is testing using the $\widehat{t}^2$ statistic, the situation is further complicated because it is possible that, if both estimates $\widehat{\delta}_1$ and $\widehat{\delta}_2$ are subject to a lower order bias, these biases may partially offset each other in the test statistic. Thus larger bandwidths are potentially of interest here. We also refer to Shimotsu (2006) for a similar conjecture along these lines. To investigate this conjecture, the last bandwidth we are going to consider is $m = \lfloor T^{0.9} \rfloor$.

We simulate the fractional noise process using the Cholesky decomposition of the covariance matrix. For the $\delta = 0.7$ case, we simulate the Type 1 fractionally integrated process, generating a $I(-0.3)$ process $\Delta^{-0.3}\xi_t = \varepsilon_t$ and then integrating, $x_t = \sum_{s=1}^{t} \xi_s$.

We simulate the test statistic $\widehat{t}^2$ for $[\tau_l, \tau_h] = [0.15, 0.85]$ with two minor changes to the procedure. First, we do not restrict the optimization for the estimation of $\widehat{\delta}_1$ and $\widehat{\delta}_2$ to a compact subset of $(-1/2, 1/2)$ as we do not rule out applications with δ outside that range. Second, we replace m by $m^* = \sum_{j=1}^{m} \nu_j^2$, where $\nu_j = \ln j - \frac{1}{m}\sum_{k=1}^{m} \ln k$. Heuristic arguments as in Hurvich and Chen (2000, page 164), suggest that in finite samples the variance is better approximated as $4\sum_{j=1}^{m} \nu_j^2$, or as $4\sum_{j=1}^{m} \widetilde{\nu}_j^2$, where $\widetilde{\nu}_j = \ln(2\sin(\lambda_j/2)) - \frac{1}{m}\sum_{k=1}^{m} \ln(2\sin(\lambda_k/2))$ (noticing that both $4\sum_{j=1}^{m} \nu_j^2/m \to 1$ and $\sum_{j=1}^{m} \widetilde{\nu}_j^2/m \to 1$ as $m \to \infty$). Hurvich and Chen (2000) actually recommend $\sum_{j=1}^{m} \widetilde{\nu}_j^2$, but the two measures are close enough for our purposes. For example, when $T = 128$ and $m = \lfloor T^{0.5} \rfloor$, $\sum_{j=1}^{m} \nu_j^2/m = 0.5047$ and $\sum_{j=1}^{m} \widetilde{\nu}_j^2/m = 0.5000$. We also refer to Qu (2001) for a similar application of $\sum_{j=1}^{m} \nu_j^2$ instead of m in the formula for the variance. The numerical optimization is initialized using the log-periodogram regression estimates. We use $T = 128$, 256, 512, 1024 and consider four settings for m, $m = \lfloor T^{0.5} \rfloor$, $\lfloor T^{0.65} \rfloor$, $\lfloor T^{0.79} \rfloor$ and $\lfloor T^{0.9} \rfloor$. In cases in which m exceeds $T/2 - 1$, we set $m = T/2 - 1$. For example, when $T = 128$ then $m = \lfloor 128^{0.9} \rfloor = 78$ and we set $m = 63$. The empirical size of the tests is measured by how often the test statistic exceeds the 5% critical value, 8.85. For each case, we run 1000 repetitions.

In order to compare our test to the test based on the Whittle fully parametric model, we also perform the test in Horváth and Shao (1999), for the same $[\tau_u, \tau_h]$ interval and the same DGP and datasets. We always run the test under the assumption that x_t is a fractional noise, even when in fact x_t has a AR(1) or MA(1) component, so the parametric model is not correctly specified in those cases. For the sake of comparison, here too we compute the test statistic without restricting the optimization to a compact subset of $(-1/2, 1/2)$, and we use a finite T approximation of the factor $\pi^2/6$ that is used to standardize the test statistic.

We find that size performance is best for $m = \lfloor T^{0.65} \rfloor$. In all cases the empirical size converges to 5% as T gets larger and results are satisfactory for every model and sample size. We notice, in particular, that results seem satisfactory even in the cases not formally covered by our theoretical study, namely the cases in which a non-stationary fractional noise is used, or residuals from a regression, or when the available moments do not meet the requirement from our Assumption 4.

Results for $m = \lfloor T^{0.5} \rfloor$ are a bit more puzzling. For example, we did not expect to observe size distortion in the normally independently distributed case. We conjecture that this may be due to the small sample. However, this is a minor concern as, in view of the power study, we recommend $m = \lfloor T^{0.65} \rfloor$ over $m = \lfloor T^{0.5} \rfloor$ anyway. Results for $m = \lfloor T^{0.79} \rfloor$ and for $m = \lfloor T^{0.9} \rfloor$ are similarly characterized by relevant size distortion, especially in the latter case (recalling that this case was not covered by our Assumption 4 or by the corresponding assumption in Robinson, 1995). Here, the bandwidths are not recommended if relevant AR or MA components are expected and, for $m = \lfloor T^{0.9} \rfloor$, even in case of fractional noise with non-zero order of integration. Overall, it seems that a certain offsetting of the bias in the estimates $\widehat{\delta}_1$ and $\widehat{\delta}_2$ does indeed take place in the $\widehat{t}^2$ statistic, but this is not sufficient for recommending adopting longer bandwidths.

As for the test using the fully parametric estimates, we see that this is appropriately sized when the fractional noise model is correctly assumed but that it may be severely size distorted otherwise. It is particularly worrying, in this case, that the size is diverging as the sample is increased, suggesting spurious evidence of a break. This outcome is very important as it demonstrates empirically the advantage of using a semiparametric test against a fully parametric one, at least in this context.

In the second part of the Monte Carlo exercise we deal with detecting a break in δ, that is, we study the power of the test. The tests are based on $\widehat{t^2} = \sup_{\tau \in [0.15, 0.85]} \hat{t}(\tau)$ but for comparison, in some cases we also consider $\widehat{t}(\tau)^2$ for a values of τ that may or may not be equal to τ^*. Statistic $\widehat{t}(\tau)^2$ is interesting as this is the statistic we would use if we knew the point where the change in persistence has taken place. For example, in our application in Section 4 where we study inflation in Germany between 1986–2017, one may conjecture that a potential breakpoint is in January 1999 when the ECB took over from the Bundesbank the task of managing monetary policy. Heuristically, statistic $\widehat{t}(\tau)^2$ may be worth considering because it is χ_1^2 distributed and for given nominal size the critical value is less than the critical value of $\widehat{t^2}$. Thus, if we guessed the breakpoint correctly, we expect that $\widehat{t}(\tau)^2$ would have more power. However, situations in which the breakpoint can be identified in advance are fairly rare. We are therefore also interested in cases in which we used $\widehat{t}(\tau)^2$ despite having not identified τ correctly.

The null hypothesis of stability of δ is rejected if the test statistic exceeds the appropriate critical value with nominal size set at 5%. We use the same sample sizes as in the size exercise and we carry out 1000 repetitions for each experiment. We use bandwidths $m = \lfloor T^{0.5} \rfloor$, $\lfloor T^{0.65} \rfloor$ and $\lfloor T^{0.79} \rfloor$. We do not include $m = \lfloor T^{0.9} \rfloor$ as this case was not covered by the theoretical model and we have found from the size study that it suffered excessive size distortion. On the other hand, we retained $m = \lfloor T^{0.79} \rfloor$ even though this bandwidth too gave rise to excessive size distortion because by keeping it in the design we can better demonstrate the size-power trade-off that is associated with the bandwidth choice. To make sure that the power for bandwidth $m = \lfloor T^{0.79} \rfloor$ is genuine, and not the result of size distortion, we only focus on fractional noise models for x_t. In reality, however, we could not count on x_t being certainly a fractional noise, as this basically amounts to a precise parametric specification, so we would not use $m = \lfloor T^{0.79} \rfloor$ as it is susceptible to cause size distortion.

We consider several models. First, setting $\delta_1 = 0.4$, $\delta_2 = 0$, we look at breakpoints located at $\tau^* = 1/2$, $\tau^* = 1/3$ and $\tau^* = 2/3$ to investigate if the power is higher for break-points that are close to the middle of the sample period. For break-points that are not close to the middle of the sample, we also investigate if there is a relevant difference for break-points that are either at the beginning or at the end of the sample. Next, we consider $\delta_1 = 0.2$, $\delta_2 = 0$ with $\tau^* = 1/2$, to see if the power is higher for larger changes in δ, *ceteris paribus*. As with the size study, we also consider cases not covered by our theoretical results. In particular, we also consider $\delta_1 = 1$, $\delta_2 = 0$. Since changes from 0 to 1 or vice-versa are often considered in the empirical literature, power in this case seems important. We further consider a case in which $\delta_1 = 0.4$ and $\delta_2 = 0$ but subjecting the series also to a change in the mean at the same breakpoint. We kept $\tau^* = 1/2$, and the average moved from -1 in the first half of the sample to +1 in the second half. This situation seems to characterize at least some of the empirical studies that we discussed in the Introduction (see for example Sibbertsen, Wegener and Basse, 2014). In all cases, both before and after the break, the process is a fractional Gaussian noise, with different order of integration in the two subsamples.

The results of the power simulations are reported in Table 2. In all tests, for a given sample size and bandwidth rule we observe that the power is higher the larger is the break, and the closer the breakpoint is to the middle of the sample. We also observe that for a given bandwidth rule and model the power increases with the sample size, and that for a given sample size and model the power increases with the bandwidth. On the other hand, positioning the break-point at the beginning or at the end of the sample does not seem to alter the power of the test.

Finally, we find that a simultaneous break in the population average has a serious detrimental effect on the ability of our test to detect a change the memory parameter. We conjecture that this is due to the fact that the estimate of the memory parameter is inconsistent in the presence of a break in the mean, see for example Qu (2011).

As last exercise we compare the power of the tests using the statistics $\widehat{t}^2$ and $\widehat{t}(\tau)^2$. We find that if the potential breakpoint is chosen correctly in the $\widehat{t}(\tau)^2$, so that $\widehat{t}(\tau^*)^2$ is used, then $\widehat{t}(\tau^*)^2$ has indeed more power than $\widehat{t}^2$, as anticipated. Otherwise the power of the $\widehat{t}(\tau)^2$ test may be quite limited, especially when $\tau^* = 2/3$. We also notice that when $\tau^* = 2/3$ then the test based on $\widehat{t}(1/2)^2$ has more power than the test based on $\widehat{t}(1/3)^2$ and thus the larger error in choosing τ compared to τ^* is penalized with a more relevant loss of power.

Table 1: Empirical size in the case of no breaks

m	T	iid	AR $\phi = 0.5$	ARFIMA $(0, 0.4, 0)$	MA $\theta = 0.8$	AR $\phi = 0.8$	t distr 5 df	ARFIMA $(0, 0.7, 0)$	residuals from trend	residuals from break
	128	0.060	0.052	0.039	0.057	0.052	0.052	0.045	0.068	0.068
$\lfloor T^{0.5} \rfloor$	256	0.069	0.076	0.053	0.072	0.08	0.061	0.059	0.097	0.100
	512	0.088	0.086	0.064	0.083	0.083	0.087	0.054	0.106	0.092
	1024	0.079	0.075	0.048	0.077	0.07	0.077	0.062	0.111	0.084
	128	0.071	0.082	0.048	0.076	0.095	0.067	0.051	0.082	0.075
$\lfloor T^{0.65} \rfloor$	256	0.082	0.083	0.056	0.078	0.099	0.071	0.058	0.079	0.075
	512	0.071	0.069	0.053	0.068	0.101	0.066	0.051	0.075	0.074
	1024	0.050	0.052	0.034	0.047	0.074	0.051	0.055	0.053	0.053
	128	0.052	0.076	0.034	0.071	0.086	0.054	0.050	0.072	0.070
$\lfloor T^{0.79} \rfloor$	256	0.053	0.083	0.049	0.060	0.104	0.053	0.050	0.056	0.062
	512	0.055	0.084	0.051	0.055	0.123	0.045	0.064	0.078	0.061
	1024	0.049	0.075	0.042	0.056	0.152	0.042	0.081	0.042	0.049
	128	0.054	0.073	0.073	0.155	0.075	0.049	0.051	0.070	0.069
$\lfloor T^{0.90} \rfloor$	256	0.069	0.098	0.098	0.213	0.113	0.054	0.046	0.084	0.072
	512	0.045	0.103	0.103	0.257	0.119	0.04	0.074	0.063	0.047
	1024	0.049	0.103	0.103	0.253	0.122	0.044	0.156	0.048	0.055
	128	0.060	0.084	0.042	0.138	0.083	0.055	0.055	0.079	0.069
Parametric	256	0.071	0.112	0.059	0.185	0.135	0.060	0.055	0.085	0.073
	512	0.045	0.122	0.046	0.236	0.157	0.043	0.073	0.064	0.058
	1024	0.046	0.122	0.051	0.222	0.165	0.045	0.162	0.048	0.048

Table 2: Empirical power in the case of one break

m	T	$d_1 = 0.4$ $\tau^* = 1/3$ $\sup(t)$	$t = 1/3$	$t = 1/2$	$t = 2/3$	$d_1 = 0.4$ $\tau^* = 1/2$ $\sup(t)$	$d_1 = 0.4$ $\tau^* = 2/3$ $\sup(t)$	$t = 1/3$	$t = 1/2$	$t = 2/3$	$d_1 = 0.2$ $\tau^* = 1/2$ $\sup(t)$	$d_1 = 1$ $\tau^* = 1/2$ $\sup(t)$	Break $\sup(t)$
	128	0.142	0.197	0.209	0.161	0.152	0.146	0.044	0.070	0.184	0.071	0.483	0.016
$\lfloor T^{0.5} \rfloor$	256	0.227	0.296	0.310	0.257	0.245	0.224	0.066	0.108	0.288	0.100	0.804	0.027
	512	0.339	0.405	0.414	0.378	0.349	0.310	0.068	0.119	0.384	0.143	0.959	0.032
	1024	0.461	0.569	0.566	0.497	0.486	0.429	0.058	0.130	0.524	0.168	1	0.041
	128	0.282	0.432	0.360	0.295	0.292	0.244	0.075	0.105	0.399	0.095	0.930	0.016
$\lfloor T^{0.65} \rfloor$	256	0.459	0.622	0.565	0.444	0.467	0.423	0.095	0.160	0.572	0.140	0.996	0.014
	512	0.654	0.796	0.748	0.653	0.708	0.636	0.071	0.162	0.803	0.192	1	0.020
	1024	0.859	0.939	0.908	0.837	0.894	0.843	0.100	0.205	0.939	0.273	1	0.018
	128	0.473	0.678	0.542	0.409	0.523	0.450	0.104	0.214	0.656	0.116	1	0.031
$\lfloor T^{0.79} \rfloor$	256	0.756	0.908	0.796	0.628	0.798	0.739	0.179	0.309	0.882	0.237	1	0.044
	512	0.946	0.989	0.956	0.858	0.973	0.944	0.205	0.429	0.985	0.392	1	0.127
	1024	0.997	0.999	0.996	0.986	0.997	1	0.312	0.567	1	0.671	1	0.290

4 Empirical application: From the Bundesbank to the ECB

We use our semiparametric test for persistence stability to analyze the inflation rate in Germany for the period 1986–2017. Interest in inflation persistence is motivated by the fact that stabilizing inflation is a key monetary policy target. This is sometimes recognized explicitly in a formal inflation target, for example in Germany (until 1999) and the Euro area (after 1999), or in the United Kingdom, Canada, New Zealand, and other countries. Even in cases in which a formal inflation targeting commitment may be lacking, such as for the US, inflation stabilization is still relevant. In practice it is of course impossible to maintain inflation constantly on the target but it is at least important that deviations from the targets are not too extreme and not too strongly persistent because such deviations would signal long term imbalances.

The order of integration provides an intuitive and simple measure of persistence that can be given an easy economic interpretation. A low level persistence can be associated with tighter inflation control. Conversely, a large degree of the persistence index signals a situation in which the central bank does not or cannot control inflation. A test for a break, and possibly a comparative study of the estimators before and after the break, would also reveal if a structural change, either in the management of monetary policy, or in the structure of the economy, or both, has taken place. There is therefore a wide range of empirical work dedicated to the estimation of the order of integration and testing for a change in this order. In studies on US data, Kumar and Okimoto (2007), Sibbertsen and Kruse (2009) and Martins and Rodrigues (2014) have all found that inflation persistence declined since 1982. On the other hand, Hassler and Meller (2014) have found that inflation persistence has increased since 1973, a second break in 1980 not being significant. With integer orders only, Halunga, Osborn and Sensier (2009) have concluded that inflation persistence increased in the early 1970s and returned under control in the early 1980s.

Germany has received comparatively less attention, featuring occasionally in wider studies for a range of countries such as in a study by Martins and Rodrigues (2014). A dedicated study of the case of Germany seems of particular interest because of the history of its central bank's monetary policy. The Bundesbank was committed to the monetary policy target of price stability which was formally implemented with an intermediate target in form of monetary aggregate. However, the Bundesbank also announced an inflation projection for the medium term which was set as 2% since 1986 (with a band 1.5%–2% in 1997–1998). Although the Bundesbank was formally committed to a monetary target, Bernanke and Mihov (1996) showed that "the Bundesbank is much better described as an inflation targeter that as a money targeter". The same inflation target for monetary policy was officially adopted by the European Central Bank (ECB) for the Euro area, although with the slightly different statement of "below but close to 2%".

Broadly speaking, therefore, the ECB targeted the same inflation rate as the Bundesbank did. However as the ECB is a different institution from the Bundesbank and as its mandate is for the euro area, rather than for Germany only, it is important to check if the change in the monetary authority resulted in an increase or decrease of persistence. This experiment is particularly interesting because it is sometimes difficult to identify if a change in inflation persistence is due to a change in the structure of the economy, rather than to the attitude of the central bank: the introduction of the euro provides us with a natural experiment to compare the attitude of the ECB against the Bundesbank. Of course, the fact that the euro was introduced in January 1999 also provides us with an additional piece of information and we could also test for a break with known breakpoint. This would be advantageous because if the choice of a breakpoint is correct the test has more power. However, since the sample spans several other periods of interest, including the effects of the German reunification and the financial crisis, testing over the whole sample offers a wider picture of the inflation dynamics.

The case of Germany is also opportune because of additional institutional information available. As can be seen from the Monte Carlo exercise, neglecting a change in the mean may have adverse effect on the power of the test $\widehat{t}^2$. Over the period that we consider, we can be fairly confident that the average inflation in Germany was stable. We have already mentioned that the Bundesbank targeted inflation at 2%. We should also note that in our dataset, the sample average of inflation is approximately 1.985 for the period 1986–1999. In the second part of the sample, since the inflation of interest for the ECB refers to the whole euro-area, it is theoretically possible that the ECB met its target but still delivered a level of inflation significantly different from 2% for Germany. This would result in a change in the mean for German inflation. In fact, there is evidence that German inflation has been 2% on average even during the ECB mandate, see for example Hualde and Iacone (2017). Thus, there is a good factual argument to support the conjecture that the population average is constant over the whole time span.

In our empirical analysis, we use CPI data from Datastream, series code BDCONPRCF. The monthly time series spans the period from January 1986 to April 2017, for a total of 375 observations. We obtain inflation as $\log(cpi_t) - \log(cpi_{t-1})$. This is a monthly inflation rate. In Figure 1 we plot the annualized monthly inflation rate, $1200 \times (\log(cpi_t) - \log(cpi_{t-1}))$. In Figure 2 and 3 we also plot the autocorrelation function and the periodograms for a range of lags and frequencies only. We can easily detect there seasonality and some evidence of long memory in the lowest frequencies of the periodograms.

We first test for a constant mean using the robust test in Iacone, Leybourne and Taylor (2014) selecting the Bartlett kernel with bandwidth $b = 0.1$. The test statistic takes value 3.99 which is below any critical value because the critical value for $\delta = 0$ is 15.39 and the critical values are increasing in δ. We thus find further evidence corroborating the assumption that the average inflation did not change.

Next, we compute the test statistics $\widehat{t}^2$ for trimming region $[\tau_l, \tau_h] = [0.15, 0.85]$, and $\widehat{t}(\tau)^2$ with τ set so that $\lfloor \tau T \rfloor$ corresponds to January 1999, for bandwidth $m = \lfloor T^{0.65} \rfloor = 47$, and bandwidths $m = 27$ and $m = 29$. The last two bandwidths have been chosen to avoid the effects of seasonality which are concentrated around the frequency corresponding to $j = 31$. In contrast, the test performance for $m = \lfloor T^{0.65} \rfloor$ may be less reliable because of possible contamination from seasonality. The results are summarized in Table 3, where we also present estimates $\widehat{\delta}$, $\widehat{\delta}_1$ and $\widehat{\delta}_2$ for potential breakpoint in January 1999.

Table 3: German inflation rate, 1986–2017. Tests for a break and estimates of δ.

m	$\widehat{t}^2$	$\widehat{t}(\tau)^2$	$\widehat{\delta}$	$\widehat{\delta}_1$	$\widehat{\delta}_2$
27	1.47	0.93	0.39	0.48	0.26
29	1.18	0.28	0.33	0.35	0.24
$\lfloor T^{0.65} \rfloor$	1.05	0.59	0.08	0.09	-0.04

We find that the persistence across the two samples is slightly higher during the Bundesbank tenure than afterwards but not significantly so. No test leads to the rejection of the null hypothesis. Overall we interpret these results as evidence that inflation persistence for Germany did not increase with the change of the monetary authority and that at most it declined in the second part of the sample. On balance we conclude that the German inflation was not subject to major instability over these years.

5 Conclusions

We study the local Whittle estimator of the memory parameter in the presence of a structural break in the stochastic component. We find that when the location of the break is unknown the consistency of the test based on $\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)$ may rest on a lower order bias only and a test based on $\sup_\tau \left(\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)\right)$, $\tau \in [\tau_l, \tau_h] \subset (0,1)$, seems advisable. A Monte Carlo exercise supports this conjecture. We also find that in some circumstances the size of the test may be incorrect but that this effect is mitigated as the sample gets larger. We apply the test to study the persistence of inflation in Germany over the period 1986–2017. We find that the persistence did not change and that we can conclude that the transition from the Bundesbank to the Eurosystem did not deteriorate the measure of the inflation control.

6 Appendix

In this Appendix we present the technical results together with their proofs and auxiliary lemmas.

6.1 Consistency of estimators of δ

Proposition 1 *Under Assumptions 1–4 and under the null hypothesis with $\delta_1 = \delta_2 = \delta$,*

$$\sup_{\tau_\ell \le \tau \le 1} \left|\widehat{\delta}_1(\tau) - \delta\right| \xrightarrow{p} 0 \qquad \text{as } T \to \infty,$$
$$\sup_{0 \le \tau \le \tau_h} \left|\widehat{\delta}_2(\tau) - \delta\right| \xrightarrow{p} 0 \qquad \text{as } T \to \infty,$$

so that

$$\sup_{\tau_\ell \le \tau \le \tau_h} \left|\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)\right| \xrightarrow{p} 0 \qquad \text{as } T \to \infty.$$

Let $u_{\tau,j} = \frac{w_{0\tau}(\lambda_j)}{A_j}$, $v_{\tau,j} = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{\lfloor \tau T \rfloor} \varepsilon_t e^{-i\lambda_j t}$, $u_{\sigma\tau,j} = u_{\tau,j} - u_{\sigma,j}$ and $v_{\sigma\tau,j} = v_{\tau,j} - v_{\sigma,j}$.

Lemma 1 *Under Assumptions 1–4, for any integers $1 \le |j_s| \le m$ with $1 \le s \le p$ where $p = 2, \ldots, 6$ and 8, and for σ and τ such that $[\sigma, \tau] \subset [0, \tau^*]$ or $[\sigma, \tau] \subset [\tau^*, 1]$, there is a finite constant C such that*

$$\left|cum\left(u_{\sigma\tau,j_1} - v_{\sigma\tau,j_1}, \ldots, u_{\sigma\tau,j_p} - v_{\sigma\tau,j_p}\right)\right| \le C(\tau - \sigma)^{\frac{p}{2}} \left|j_1 \cdots j_p\right|^{-\frac{1}{2}}. \tag{10}$$

Proof of Lemma 1. Let $A_j = A_1(\lambda_j)$ when $[\sigma, \tau] \subset [0, \tau^*]$ and $A_j = A_2(\lambda_j)$ when $[\sigma, \tau] \subset [\tau^*, 1]$. When $p = 8$, using formulas (2.6.3) and (2.10.3) of Brillinger (1981), the cumulant on the left of (10) can be written as

$$
\begin{aligned}
&\frac{\kappa_8}{T^4}\frac{1}{(2\pi)^9}\underbrace{\int\dots\int_{-\pi}^{\pi}}_{7\times}\left(\frac{A(\omega_1)}{A(\lambda_{j_1})}-1\right)\left(\frac{A(\omega_2)}{A(\lambda_{j_2})}-1\right)\left(\frac{A(\omega_3)}{A(\lambda_{j_3})}-1\right)\\
&\times\left(\frac{A(\omega_4)}{A(\lambda_{j_4})}-1\right)\left(\frac{A(\omega_5)}{A(\lambda_{j_5})}-1\right)\left(\frac{A(\omega_6)}{A(\lambda_{j_6})}-1\right)\left(\frac{A(\omega_7)}{A(\lambda_{j_7})}-1\right)\\
&\times\left(\frac{A(-\omega_1-\ldots-\omega_7)}{A(\lambda_{j_8})}-1\right)D_{\sigma\tau}(\omega_1-\lambda_{j_1})\,D_{\sigma\tau}(\omega_2-\lambda_{j_2})\,D_{\sigma\tau}(\omega_3-\lambda_{j_3})\\
&\times D_{\sigma\tau}(\omega_4-\lambda_{j_4})\,D_{\sigma\tau}(\omega_5-\lambda_{j_5})\,D_{\sigma\tau}(\omega_6-\lambda_{j_6})\,D_{\sigma\tau}(\omega_7-\lambda_{j_7})\\
&\times D_{\sigma\tau}(-\omega_1-\ldots-\omega_7-\lambda_{j_8})\,d\omega_1\ldots d\omega_7,
\end{aligned}
$$

where $\kappa_8 = cum(\varepsilon_t,\ldots,\varepsilon_t)$ is the eighth cumulant of ε_t and

$$D_{\sigma\tau}(\lambda)=\sum_{t=\lfloor\sigma T\rfloor+1}^{\lfloor\tau T\rfloor}e^{it\lambda}.$$

It follows from the Schwarz inequality and periodicity that this is bounded by

$$\frac{\kappa_8}{(2\pi)^2}\left(P_{\sigma\tau,j_1}\cdots P_{\sigma\tau,j_8}\right)^{\frac{1}{2}}$$

where

$$P_{\sigma\tau,j}=\int_{-\pi}^{\pi}\left|\frac{A(\omega)}{A(\lambda_j)}-1\right|^2 K_{\sigma\tau}(\omega-\lambda_j)\,d\omega$$

and

$$K_{\sigma\tau}(\lambda)=K_{\sigma\tau,T}(\lambda)=\frac{|D_{\sigma\tau}(\lambda)|^2}{2\pi T}.$$

Noting that $K_{\sigma\tau,T}(\lambda)=\frac{\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}{T}K_{01,\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}(\lambda)$, we write

$$P_{\sigma\tau,j}=\frac{\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}{T}\int_{-\pi}^{\pi}\left|\frac{A(\omega)}{A(\lambda_j)}-1\right|^2 K_{01,\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}(\omega-\lambda_j)\,d\omega. \tag{11}$$

The kernel $K_{01,\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}$ has the property

$$K_{01,\lfloor\tau T\rfloor-\lfloor\sigma T\rfloor}(\lambda)\le\frac{C}{|\lambda|},\qquad 0<\lambda\le\pi,\quad T\ge 1$$

by Lemma 1 of Lazarová (2005). Using the arguments in the proof of Lemma 3 of Robinson (1995), the integral in (11) can be seen to be $O\left(|j|^{-1}\right)$ uniformly over integers $1\le|j|\le m$. Therefore

$$P_{\sigma\tau,j}\le C(\tau-\sigma)|j|^{-1} \tag{12}$$

uniformly over integers $1\le|j|\le m$ and bound (10) holds for $p=8$. A similar approach yields proof of bound (10) for $p=2,\ldots,6$. ■

Lemma 2 *Under Assumptions 1–4, for $1 \leq j \leq m$, and for σ and τ such that $\tau - \sigma \geq 1/T$ and $[\sigma, \tau] \subset [0, \tau^*]$ or $[\sigma, \tau] \subset [\tau^*, 1]$, there is a finite constant C such that*

$$(a) \qquad E\left|u_{\sigma\tau,j} - v_{\sigma\tau,j}\right|^8 \leq C\left(\tau - \sigma\right)^4 j^{-4},$$

$$(b) \qquad E\left|v_{\sigma\tau,j}\right|^8 \leq C\left(\tau - \sigma\right)^4.$$

Proof of Lemma 2. (a) Using formula (2.8) of McCullagh (1987), we have that for random variables $Y_1, \ldots, Y_r$,

$$E\left(Y_1 \cdots Y_r\right) = \sum_{\pi} \prod_{B \in \pi} cum\left(Y_i : i \in B\right), \tag{13}$$

where π runs through the list of all partitions of $\{1, \ldots, r\}$ and B runs through the list of all blocks of the partition π. Since $E\left(u_{\sigma\tau,j} - v_{\sigma\tau,j}\right) = 0$, part (a) is implied by Lemma 1.

(b) We have

$$\begin{aligned} E\left|v_{\sigma\tau,j}\right|^8 &= \frac{1}{\left(2\pi T\right)^4} \sum_{t,s,r,v,t_1,s_1,r_1,v_1 = \lfloor \sigma T \rfloor + 1}^{\lfloor \tau T \rfloor} E\varepsilon_t\varepsilon_s\varepsilon_r\varepsilon_v\varepsilon_{t_1}\varepsilon_{s_1}\varepsilon_{r_1}\varepsilon_{v_1} \\ &\quad \times e^{i(t-s)\lambda_j} e^{i(r-v)\lambda_j} e^{i(t_1-s_1)\lambda_j} e^{i(r_1-v_1)\lambda_j} \\ &\leq \frac{1}{\left(2\pi T\right)^4} \sum_{t,s,r,v,t_1,s_1,r_1,v_1 = \lfloor \sigma T \rfloor + 1}^{\lfloor \tau T \rfloor} \left|E\varepsilon_t\varepsilon_s\varepsilon_r\varepsilon_v\varepsilon_{t_1}\varepsilon_{s_1}\varepsilon_{r_1}\varepsilon_{v_1}\right|. \end{aligned}$$

Using (13) it can be seen that $E\left|v_{\sigma\tau,j}\right|^8$ is bounded by

$$\frac{C}{T^4}\left(\sum_{t,s,r,v} \kappa_2^4 + \sum_{t,s,r}\left(\kappa_2^2\kappa_4 + \kappa_2\kappa_3^2\right) + \sum_{t,s}\left(\kappa_2\kappa_6 + \kappa_3\kappa_5 + \kappa_4^2\right) + \sum_t \kappa_8\right),$$

where $\kappa_p = cum\left(\varepsilon_t, \ldots, \varepsilon_t\right)$ is the p-th cumulant of ε_t and where the sums run from $\lfloor \sigma T \rfloor + 1$ to $\lfloor \tau T \rfloor$. This is bounded by

$$C\left(\frac{\left(\tau - \sigma\right)^4 T^4}{T^4} + \frac{\left(\tau - \sigma\right)^3 T^3}{T^4} + \frac{\left(\tau - \sigma\right)^2 T^2}{T^4} + \frac{\left(\tau - \sigma\right) T}{T^4}\right) \leq C\left(\tau - \sigma\right)^4$$

since $\frac{1}{T} \leq \tau - \sigma$. ■

Let

$$\nu_j = \log j - \frac{1}{m}\sum_{k=1}^{m} \log k = \log\left(\frac{j}{m}\right) - \frac{1}{m}\sum_{k=1}^{m} \log\left(\frac{k}{m}\right). \tag{14}$$

Let $D_k\left(\lambda\right) = \sum_{t=1}^{k} e^{it\lambda}$ and $K_k = \frac{1}{2\pi T}\left|D_k\left(\lambda\right)\right|^2$.

Lemma 3 *For k such that $k/T \to a$ with $0 < a \leq 1$ as $T \to \infty$, we have*

$$(a) \qquad \frac{2\pi}{mT}\sum_{j=1}^{m}\sum_{\ell=1}^{m}\nu_j\nu_\ell K_k\left(\lambda_j-\lambda_\ell\right)=\frac{k}{T}+o\left(1\right),$$

$$(b) \qquad \frac{2\pi}{mT}\sum_{j=1}^{m}\sum_{\ell=1}^{m}\nu_j\nu_\ell K_k\left(\lambda_j+\lambda_\ell\right)=o\left(1\right).$$

Proof of Lemma 3. (a) We have

$$\begin{aligned}
&\frac{2\pi}{mT}\sum_{j=1}^{m}\sum_{\ell=1}^{m}\nu_j\nu_\ell K_k\left(\lambda_j-\lambda_\ell\right)\\
&=\frac{2\pi}{mT}\sum_{j=1}^{m}\nu_j^2\sum_{\ell=1}^{T}K_k\left(\lambda_j-\lambda_\ell\right)-\frac{2\pi}{mT}\sum_{j=1}^{m}\nu_j^2\sum_{\ell=m+1}^{T}K_k\left(\lambda_j-\lambda_\ell\right)\\
&+\frac{2\pi}{mT}\sum_{j=1}^{m}\nu_j\sum_{\ell=1}^{m}\left(\nu_\ell-\nu_j\right)K_k\left(\lambda_j-\lambda_\ell\right). \qquad (15)
\end{aligned}$$

The first term on the right of (15) is equal to

$$\frac{k}{T}\frac{1}{m}\sum_{j=1}^{m}\nu_j^2=\frac{k}{T}\left(1+o\left(1\right)\right)$$

because $m^{-1}\sum_{j=1}^{m}\nu_j^2=1+O\left(m^{-1}\log^2 m\right)$. Kernel K has the following properties:

$$K_k\left(\lambda\right)\leq\frac{k^2}{2\pi T} \qquad \lambda\in\left[0,2\pi\right], \qquad (16)$$

$$K_k\left(\lambda\right)\leq\frac{\pi}{2T\lambda^2} \qquad \lambda\in\left(0,\pi\right]. \qquad (17)$$

For $|\ell|\leq T/\left(2k\right)$ the first bound for $K_k\left(\lambda_\ell\right)$ is at least as good as the second bound.

For large enough T, $T/2>m+1$ and the second term on the right of (15) is bounded by

$$\begin{aligned}
&\frac{C\log^2 m}{mT}\sum_{j=1}^{m}\sum_{\ell=m-j+1}^{T-j}K_k\left(\lambda_\ell\right)\leq\frac{C\log^2 m}{mT}\sum_{j=1}^{m}\left(\sum_{\ell=m-j+1}^{T/2}+\sum_{\ell=j}^{T/2}\right)\frac{1}{T\lambda_\ell^2}\\
&=\frac{C\log^2 m}{m}\sum_{j=1}^{m}\left(\sum_{\ell=m-j+1}^{T/2}+\sum_{\ell=j}^{T/2}\right)\frac{1}{\ell^2}\leq\frac{C\log^2 m}{m}\sum_{j=1}^{m}\left(\frac{1}{m-j+1}+\frac{1}{j}\right)\\
&\leq\frac{C\log^3 m}{m}
\end{aligned}$$

because $\nu_j=O\left(\log m\right)$, because $\sum_{j=a}^{b}j^{-2}\leq Ca^{-1}$ and because kernel K is symmetric.

Let $a_k = \lfloor T/(2k) \rfloor$. For sufficiently large T it is $m > a_k$ and the third term on the right of (15) is bounded in absolute value by

$$\begin{aligned} &\frac{C\log m}{mT}\sum_{\ell=1}^{m-1}\sum_{j=\ell+1}^{m}\left|\nu_{j-\ell}-\nu_j\right| K_k\left(\lambda_\ell\right) \\ &\leq \frac{C\log m}{mT}\left(\sum_{\ell=1}^{a_k}\sum_{j=\ell+1}^{m}\left|\nu_{j-\ell}-\nu_j\right|\frac{k^2}{T}+\sum_{\ell=a_k+1}^{m-1}\sum_{j=\ell+1}^{m}\left|\nu_{j-\ell}-\nu_j\right|\frac{T}{\ell^2}\right). \end{aligned} \tag{18}$$

By the Taylor theorem,

$$\nu_{j-\ell}-\nu_j = \log(j-\ell)-\log j = -\frac{\ell}{\xi} \qquad j-\ell \leq \xi \leq j,$$

so that

$$\left|\nu_{j-\ell}-\nu_j\right| \leq \frac{\ell}{j-\ell} \qquad 1 \leq \ell \leq m,\ \ell+1 \leq j \leq m.$$

Therefore (18) is bounded by

$$\begin{aligned} &\frac{C\log m}{mT}\left(\frac{k^2}{T}\sum_{\ell=1}^{a_k}\ell\sum_{j=\ell+1}^{m}\frac{1}{j-\ell}+T\sum_{\ell=a_k+1}^{m-1}\sum_{j=\ell+1}^{m}\frac{1}{(j-\ell)\,\ell}\right) \\ &= \frac{C\log m}{m}\left(\sum_{\ell=1}^{a_k}\ell\sum_{j=1}^{m-\ell}\frac{1}{j}+\sum_{\ell=a_k+1}^{m-1}\frac{1}{\ell}\sum_{j=1}^{m-\ell}\frac{1}{j}\right) \\ &\leq \frac{C\log m}{m}\left(\log m\sum_{\ell=1}^{a_k}\ell+\log^2 m\right) \leq \frac{C\log^3 m}{m} = o(1). \end{aligned}$$

Gathering results, it can be seen that part (a) holds.

(b) We have that

$$\frac{2\pi}{mT}\sum_{j=1}^{m}\sum_{\ell=1}^{m}\nu_j\nu_\ell K_k\left(\lambda_j+\lambda_\ell\right) = \frac{2\pi}{mT}\sum_{j=1}^{m}\sum_{\ell=1+j}^{m+j}\nu_j\nu_{\ell-j}K_k\left(\lambda_\ell\right)$$

which is bounded in absolute value by

$$\frac{C\log^2 m}{mT}\sum_{j=1}^{m}\sum_{\ell=1+j}^{m+j}\frac{T}{\ell^2} \leq \frac{C\log^2 m}{m}\sum_{j=1}^{m}\frac{1}{j+1} \leq C\frac{\log^3 m}{m} = o(1)$$

which shows that part (b) holds true. ∎

For any real number a, let $|a|_+ = \max\{1, |a|\}$.

Lemma 4 *The following inequalities hold:*

$$\text{(a)} \qquad \sum_{j=1}^{m}\sum_{k=1}^{m}\frac{1}{j+k} \leq Cm\log m,$$

$$\text{(b)} \qquad \sum_{j=1}^{m}\sum_{k=1}^{m}\frac{1}{|j-k|_+} \leq Cm\log m,$$

$$\text{(c)} \qquad \sum_{l=1}^{m}\frac{1}{j+l}\frac{1}{k+l} \leq \frac{C\log m}{j+k}, \qquad 1\leq j,k\leq m,$$

$$\text{(d)} \qquad \sum_{l=1}^{m}\frac{1}{|j-l|_+}\frac{1}{l+k} \leq \frac{C\log m}{j+k}, \qquad 1\leq j,k\leq m.$$

$$\text{(e)} \qquad \sum_{l=1}^{m}\frac{1}{|j-l|_+\,|k-l|_+} \leq \frac{C\log m}{|j-k|_+}, \qquad 1\leq j,k\leq m,$$

Proof. (a)

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\frac{1}{j+k} \leq \sum_{j=1}^{m}\sum_{k=1}^{m}\frac{1}{k} \leq Cm\log m.$$

(b)

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\frac{1}{|j-k|_+} = \sum_{j=1}^{m}1 + 2\sum_{d=1}^{m-1}\sum_{j=d+1}^{m}\frac{1}{d} \leq m + 2m\sum_{d=1}^{m-1}\frac{1}{d} \leq Cm\log m.$$

(c)

$$\sum_{l=1}^{m}\frac{1}{j+l}\frac{1}{k+l} \leq \sum_{l=1}^{m}\frac{1}{l}\frac{1}{j+k} \leq \frac{C\log m}{j+k}.$$

(d)

$$\begin{aligned}
\sum_{l=1}^{m}\frac{1}{|j-l|_+}\frac{1}{l+k} &= \sum_{l=1}^{j-1}\frac{1}{j-l}\frac{1}{l+k} + \frac{1}{j+k} + \sum_{l=j+1}^{m}\frac{1}{l-j}\frac{1}{l+k} \\
&= \sum_{l=1}^{j-1}\frac{1}{l}\frac{1}{j+k-l} + \frac{1}{j+k} + \sum_{l=1}^{m-j}\frac{1}{l}\frac{1}{l+(j+k)} \\
&\leq \sum_{l=1}^{j+k-1}\frac{1}{l}\frac{1}{j+k-l} + \frac{1}{j+k} + \sum_{l=1}^{m-j}\frac{1}{l}\frac{1}{j+k} \\
&\leq \frac{C\log m}{j+k} + \frac{1}{j+k} + \frac{C\log m}{j+k} \leq \frac{C\log m}{j+k},
\end{aligned}$$

where the term $\sum_{l=1}^{j+k-1}\frac{1}{l}\frac{1}{j+k-l}$ is bounded using inequality $\sum_{j=1}^{a-1}\frac{1}{j}\frac{1}{a-j} \leq \frac{C\log a}{a}$ for $a = j+k$.

(e) Without loss of generality, we assume that $j \le k$. We write

$$\begin{aligned}
&\sum_{l=1}^{m} \frac{1}{|j-l|_+} \frac{1}{|k-l|_+} \\
&= \sum_{l=1}^{j-1} \frac{1}{j-l} \frac{1}{k-l} + \sum_{l=j+1}^{k-1} \frac{1}{l-j} \frac{1}{k-l} + \sum_{l=k+1}^{m} \frac{1}{l-j} \frac{1}{l-k} + \frac{2}{|j-k|_+}. \\
&= \sum_{l=1}^{j-1} \frac{1}{l} \frac{1}{k-j+l} + \sum_{l=1}^{k-j-1} \frac{1}{l} \frac{1}{k-j-l} + \sum_{l=1}^{m-k} \frac{1}{l+k-j} \frac{1}{l} + \frac{2}{|j-k|_+} \\
&\le \sum_{l=1}^{j-1} \frac{1}{l} \frac{1}{|j-k|_+} + \sum_{l=1}^{k-j-1} \frac{1}{l} \frac{1}{k-j-l} + \sum_{l=1}^{m} \frac{1}{|j-k|_+} \frac{1}{l} + \frac{2}{|j-k|_+} \\
&\le \frac{C \log m}{|j-k|_+}.
\end{aligned}$$

The term $\sum_{l=1}^{k-j-1} \frac{1}{l} \frac{1}{k-j-l}$ is equal to zero when $j = k$ and is bounded using inequality $\sum_{j=1}^{a-1} \frac{1}{j} \frac{1}{a-j} \le \frac{C \log a}{a}$ for $a = k - j$ when $j < k$. ■

For a triangular array $\{\mu_{j,m}(a), 1 \le j \le m\}_{m=1}^{\infty}$, let $\mu(a) = \lim_{m\to\infty} \frac{1}{m} \sum_{j=1}^{m} \mu_{j,m}(a)$. For simplicity, we drop the reference to m and a in what follows and write μ_j for $\mu_{j,m}(a)$ and μ for $\mu(a)$. We are in particularly concerned with the cases where $\mu_j = \mu_{j,m}$ assumes either of the following values for all j:

$$\mu_j = \left(\frac{j}{m}\right)^a \qquad \text{with } \mu = \frac{1}{1+a}, \qquad a > -1, \tag{19}$$

$$\mu_j = \log\left(\frac{j}{m}\right)\left(\frac{j}{m}\right)^a \nu_j \qquad \text{with } \mu = \frac{1-a}{(1+a)^3}, \qquad a > -1. \tag{20}$$

Lemma 5 *Under Assumptions 1–4,*

(a) $$\frac{1}{m} \sum_{j=1}^{m} \mu_j \frac{I_{0\tau,j}}{G\lambda_j^{-2\delta_1}} - \tau\mu \Longrightarrow 0 \qquad \tau \in [0, \tau^*],$$

(b) $$\frac{1}{m} \sum_{j=1}^{m} \mu_j \frac{I_{\tau^*\tau,j}}{G\lambda_j^{-2\delta_2}} - (\tau - \tau^*)\mu \Longrightarrow 0 \qquad \tau \in [\tau^*, 1],$$

(c) $$\frac{1}{m} \sum_{j=1}^{m} \mu_j \frac{I_{\tau\tau^*,j}}{G\lambda_j^{-2\delta_1}} - (\tau^* - \tau)\mu \Longrightarrow 0 \qquad \tau \in [0, \tau^*],$$

(d) $$\frac{1}{m} \sum_{j=1}^{m} \mu_j \frac{I_{\tau 1,j}}{G\lambda_j^{-2\delta_2}} - (1-\tau)\mu \Longrightarrow 0 \qquad \tau \in [\tau^*, 1],$$

(e) $$\frac{1}{m}\sum_{j=1}^{m}\mu_j \frac{w_{0\tau^*,j}}{G^{\frac{1}{2}}\lambda_j^{-\delta_1}}\frac{\overline{w}_{\tau^*\tau,j}}{G^{\frac{1}{2}}\lambda_j^{-\delta_2}} \Longrightarrow 0 \qquad \tau \in [\tau^*, 1],$$

(f) $$\frac{1}{m}\sum_{j=1}^{m}\mu_j \frac{w_{\tau\tau^*,j}}{G^{\frac{1}{2}}\lambda_j^{-\delta_1}}\frac{\overline{w}_{\tau^*1,j}}{G^{\frac{1}{2}}\lambda_j^{-\delta_2}} \Longrightarrow 0 \qquad \tau \in [0, \tau^*],$$

where $\mu_j = \mu_{j,m}$ assumes either the value defined in (19) or (20). Moreover, for any $\varepsilon > 0$ and $D < \infty$, the convergence in parts (a) to (f) holds uniformly over $-1+\varepsilon \le a \le D$ in the sense that

$$\sup_{-1+\varepsilon\le a\le D}\left|\frac{1}{m}\sum_{j=1}^{m}\mu_j \frac{I_{0\tau,j}}{G\lambda_j^{-2\delta_1}} - \tau\mu\right| \Longrightarrow 0 \qquad \tau \in [0,\tau^*]$$

in part (a) and similarly in parts (b)-(f).

Proof of Lemma 5. (a) Denote $g_j = G\lambda_j^{-2\delta_1}$ and let

$$Y(\tau) = \frac{1}{m}\sum_{j=1}^{m}\mu_j \frac{I_{0\tau,j}}{g_j} - \tau\mu.$$

We need to prove that $Y(\tau) = o_p(1)$ for any $\tau \in [0,\tau^*]$ and that the process Y is tight. Denote $f_j = f_1(\lambda_j)$, $I_{\varepsilon 0\tau,j} = |v_{\tau,j}|^2 = \left|\frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{\lfloor\tau T\rfloor}\varepsilon_t e^{-i\lambda_j t}\right|^2$ and write $Y(\tau)$ as

$$\frac{1}{m}\sum_{j=1}^{m}\mu_j\left(1-\frac{g_j}{f_j}\right)\frac{I_{0\tau,j}}{g_j} + \frac{1}{m}\sum_{j=1}^{m}\mu_j\left(\frac{I_{0\tau,j}}{f_j} - \frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)$$
$$+\frac{1}{m}\sum_{j=1}^{m}\mu_j\left(\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j} - \frac{\lfloor\tau T\rfloor}{T}\right) + \left(\frac{\lfloor\tau T\rfloor}{T}\frac{1}{m}\sum_{j=1}^{m}\mu_j - \tau\mu\right). \qquad (21)$$

The first moment of the absolute value the first term of (21) is bounded by

$$\left(\max_{1\le j\le m}\left|1-\frac{g_j}{f_j}\right|\right)\frac{1}{m}\sum_{j=1}^{m}|\mu_j|\, E\left|\frac{I_{0\tau,j}}{g_j}\right|$$

which is $o(1)$ as $T\to\infty$ because

$$\max_{1\le j\le m}\left|1-\frac{g_j}{f_j}\right| = O\left(\lambda_j^{\beta}\right) = o\left(m^{-1/2}\right)$$

by Assumptions 3 and 4, because

$$E\left|\frac{I_{0\tau,j}}{g_j}\right| \le C \qquad j = 1,\ldots,m$$

for T sufficiently large by Assumptions 1–4 and by Lemma 3 of Lazarová (2005), and because for μ_j defined in (19) and (20), $m^{-1}\sum_{j=1}^{m}|\mu_j| < C\log^2 m$.

By summation by parts, the expectation of the absolute value of the second term of (21) can be bounded by

$$\frac{1}{m}\sum_{k=1}^{m-1}|\mu_k-\mu_{k+1}|\,E\left|\sum_{j=1}^{k}\left(\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right|+|\mu_m|\,E\left|\frac{1}{m}\sum_{j=1}^{m}\left(\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right|.$$

Proceeding as Robinson (1995) did in bounding his expression (3.17), p. 1637, and employing Lemma 3 of Lazarová (2005), we obtain

$$E\left|\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right|\leq C\frac{\log^{\frac{1}{2}} j}{j^{\frac{1}{2}}},$$

so

$$E\left|\sum_{j=1}^{k}\left(\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right|\leq Ck^{\frac{1}{2}}\log^{\frac{1}{2}} k.$$

To bound $|\mu_j-\mu_{j+1}|$ for $1\leq j\leq m-1$, we first consider $\mu_j=\log\left(\frac{j}{m}\right)\left(\frac{j}{m}\right)^a\nu_j$. Denoting $b_j=\log\left(\frac{j}{m}\right)\left(\frac{j}{m}\right)^a$, we can write

$$|\mu_j-\mu_{j+1}|\leq|b_j|\,|\nu_j-\nu_{j+1}|+|b_j-b_{j+1}|\,|\nu_{j+1}|\,.$$

It can be easily seen that $|b_j|\leq j^a m^{-a}\log m$ and $|\nu_{j+1}|\leq C\log m$. By the mean value theorem, $|\nu_j-\nu_{j+1}|=\log(j+1)-\log j=(1+\xi)^{-1}\leq j^{-1}$, where $j-1\leq\xi\leq j$. Also by the mean value theorem, $|b_j-b_{j+1}|\leq m^{-a}\xi^{a-1}(1+|a|\,|\log(\xi/m)|)$ where $j\leq\xi\leq j+1$, so that $|b_j-b_{j+1}|\leq Cm^{-a}j^{a-1}\log m$ because $\xi^{a-1}\leq Cj^{a-1}$. These considerations imply that

$$|\mu_j-\mu_{j+1}|\leq Cm^{-a}j^{a-1}\log^2 m. \tag{22}$$

Similarly, when $\mu_j=\left(\frac{j}{m}\right)^a$, we obtain $|\mu_j-\mu_{j+1}|\leq Cj^{a-1}m^{-a}$. Since $|\mu_j|\leq 1$ for cases (19) or (20), the first absolute moment of the second term of (21) is bounded by

$$Cm^{-1-a}\log^{5/2} m\sum_{k=1}^{m-1}k^{a-1/2}+m^{1/2}\log^{1/2} m=o(1)\,.$$

Using summation by parts, the third term of (21) can be bounded by

$$\begin{aligned}&\frac{1}{m}\sum_{k=1}^{m-1}|\mu_k-\mu_{k+1}|\left|\sum_{j=1}^{k}\left(\frac{1}{T\sigma_\varepsilon^2}\sum_{t=1}^{\lfloor\tau T\rfloor}\sum_{s=1}^{\lfloor\tau T\rfloor}\varepsilon_t\varepsilon_s e^{-i(t-s)\lambda_j}-\frac{\lfloor\tau T\rfloor}{T}\right)\right|\\&+\frac{|\mu_m|}{m}\left|\sum_{j=1}^{m}\left(\frac{1}{T\sigma_\varepsilon^2}\sum_{t=1}^{\lfloor\tau T\rfloor}\sum_{s=1}^{\lfloor\tau T\rfloor}\varepsilon_t\varepsilon_s e^{-i(t-s)\lambda_j}-\frac{\lfloor\tau T\rfloor}{T}\right)\right|.\end{aligned}\tag{23}$$

By Assumption 3,

$$
\begin{aligned}
& E\left|\sum_{j=1}^{k}\left(\frac{1}{T\sigma_{\varepsilon}^{2}}\sum_{t=1}^{\lfloor\tau T\rfloor}\sum_{s=1}^{\lfloor\tau T\rfloor}\varepsilon_t\varepsilon_s e^{i(t-s)\lambda_j}-\frac{\lfloor\tau T\rfloor}{T}\right)\right|^2 \\
&=\frac{\kappa_4}{\sigma_{\varepsilon}^4}\frac{\lfloor\tau T\rfloor}{T}\frac{k^2}{T}+\frac{1}{T^2}\sum_{t=1}^{\lfloor\tau T\rfloor}\sum_{s=1}^{\lfloor\tau T\rfloor}\sum_{j=1}^{k}\sum_{\ell=1}^{k}\left(e^{i(t-s)(\lambda_j-\lambda_\ell)}+e^{i(t-s)(\lambda_j+\lambda_\ell)}\right) \\
&=O\left(\frac{k^2}{T}\right)+\frac{2\pi}{T}\sum_{j=1}^{k}\sum_{\ell=1}^{k}\left(K_{\lfloor\tau T\rfloor}\left(\lambda_j-\lambda_\ell\right)+K_{\lfloor\tau T\rfloor}\left(\lambda_j+\lambda_\ell\right)\right),
\end{aligned}
$$

where $\kappa_4 = cum\left(\varepsilon_t, \varepsilon_t, \varepsilon_t, \varepsilon_t\right)$, where the second term is $O\left(\log^2 k\right)$ using inequalities in Lemma 4 and inequalities (16) and (17). The first term of (23) is therefore

$$
\begin{aligned}
& O_p\left(m^{-1}\sum_{k=1}^{m-1}k^{a-1}m^{-a}\log m\left(kT^{-1/2}+k^{1/2}\right)\right) \\
&=O_p\left(T^{-1/2}\log^2 m+m^{-1}\log^3 m\right)=o_p\left(1\right).
\end{aligned}
$$

In a similar way, the second term of (23) is

$$
O_p\left(m^{-1}\log^2 m\left(mT^{-1/2}+\log m\right)\right)=o_p\left(1\right).
$$

Finally, the last term of (21) is $o\left(1\right)$ by the definition of μ. Gathering results and using the Markov inequality, we can see that $Y\left(\tau\right)=o_p\left(1\right)$ for any $\tau\in\left[0,\tau^*\right]$.

To prove tightness of process Y, we write $Y\left(\tau\right)$ as

$$
Y\left(\tau\right)=Y_1\left(\tau\right)+Y_2\left(\tau\right)-\tau\mu,
$$

where

$$
\begin{aligned}
Y_1\left(\tau\right)&=\frac{1}{m}\sum_{j=1}^{m}\mu_j\frac{f_j}{g_j}\left(\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_{\varepsilon}^2}I_{\varepsilon 0\tau,j}\right), \\
Y_2\left(\tau\right)&=\frac{2\pi}{\sigma_{\varepsilon}^2}\frac{1}{m}\sum_{j=1}^{m}\mu_j\frac{f_j}{g_j}I_{\varepsilon 0\tau,j}
\end{aligned}
$$

with $I_{\varepsilon 0\tau,j}$ as defined above equation (21). Tightness of the processes Y_i is implied by the moment condition of Billingsley (1999, Theorem 13.5, p. 142),

$$
E\left|Y_i\left(\tau\right)-Y_i\left(\rho\right)\right|^2\left|Y_i\left(\rho\right)-Y_i\left(\sigma\right)\right|^2\leq\left(F\left(\tau\right)-F\left(\sigma\right)\right)^{2\alpha}, \qquad i=1,2,
$$

where $\alpha>\frac{1}{2}$, $\sigma\leq\rho\leq\tau$ and F is a nondecreasing, continuous function on $[0,1]$. Denoting

$$
\pi_j=\mu_j f_j/g_j,
$$

we obtain for $0 \leq \sigma \leq \tau \leq \tau^*$ that

$$Y_1(\tau) - Y_1(\sigma) = \frac{2\pi}{\sigma_\varepsilon^2} \frac{1}{m} \sum_{j=1}^{m} \pi_j (a_{1j} + \ldots + a_{9j})$$

where

$$\begin{aligned}
&a_{1j} = |u_{\sigma\tau,j} - v_{\sigma\tau,j}|^2, \qquad a_{2j} = (u_{\sigma\tau,j} - v_{\sigma\tau,j}) \bar{v}_{\sigma\tau,j}, \qquad a_{3j} = \bar{a}_{2j}, \\
&a_{4j} = (u_{\sigma\tau,j} - v_{\sigma\tau,j}) \bar{v}_{\sigma,j}, \qquad a_{5j} = \bar{a}_{4j}, \\
&a_{6j} = (\bar{u}_{\sigma\tau,j} - \bar{v}_{\sigma\tau,j})(u_{\sigma,j} - v_{\sigma,j}), \qquad a_{7j} = \bar{a}_{6j}, \\
&a_{8j} = (u_{\sigma,j} - v_{\sigma,j}) \bar{v}_{\sigma\tau,j}, \qquad a_{9j} = \bar{a}_{8j}
\end{aligned}$$

and where $u_{\tau,j}$, $v_{\tau,j}$, $u_{\sigma\tau,j}$ and $v_{\sigma\tau,j}$ were defined at the beginning of this section. The fourth moment of the difference $Y_1(\tau) - Y_1(\sigma)$ is given by

$$\begin{aligned}
E|Y_1(\tau) - Y_1(\sigma)|^4 &\leq CE \left| \frac{1}{m} \sum_{j=1}^{m} \pi_j a_{1j} \right|^4 + \ldots + CE \left| \frac{1}{m} \sum_{j=1}^{m} \pi_j a_{9j} \right|^4 \\
&= \frac{C}{m^4} \sum_{r=1}^{9} \sum_{j,k,\ell,p=1}^{m} \pi_j \pi_k \pi_\ell \pi_p E a_{rj} \bar{a}_{rk} a_{r\ell} \bar{a}_{rp} \\
&\leq \frac{C}{m^4} \sum_{r=1}^{9} \sum_{j,k,\ell,p=1}^{m} |\pi_j \pi_k \pi_\ell \pi_p| \left(E|a_{rj}|^4 E|a_{rk}|^4 E|a_{r\ell}|^4 E|a_{rp}|^4 \right)^{\frac{1}{4}},
\end{aligned}$$

where the last inequality follows from the Schwarz inequality. When $r = 9$,

$$E|a_{9j}|^4 = E|(\bar{u}_{\sigma,j} - \bar{v}_{\sigma,j}) v_{\sigma\tau,j}|^4 \leq \left(E|\bar{u}_{\sigma,j} - \bar{v}_{\sigma,j}|^8 E|v_{\sigma\tau,j}|^8 \right)^{\frac{1}{2}}.$$

By Lemma 2, the last displayed expression is $O\left((\tau - \sigma)^2 j^{-2}\right)$. It can be shown in a similar way that for $1 \leq r \leq 8$, $E|a_{rj}|^4$ is also $O\left((\tau - \sigma)^2 j^{-2}\right)$. Therefore

$$E|Y_1(\tau) - Y_1(\sigma)|^4 \leq C(\tau - \sigma)^2 \left(\frac{1}{m} \sum_{j=1}^{m} |\mu_j| \frac{f_j}{g_j} j^{-\frac{1}{2}} \right)^4.$$

Since

$$\frac{1}{m} \sum_{j=1}^{m} |\mu_j| \leq -\frac{C}{m} \log m \sum_{j=1}^{m} \log\left(\frac{j}{m}\right) \left(\frac{j}{m}\right)^{-1+\varepsilon} \leq C$$

as can be seen by bounding the sum by an integral, and since $\max_{1 \leq j \leq m} \frac{f_j}{g_j} = O(1)$ by Assumption 1, the sum $m^{-1} \sum_{j=1}^{m} |\mu_j| f_j g_j^{-1} j^{-\frac{1}{2}}$ is bounded. It follows that process $Y_1(\tau)$ is tight.

Regarding process Y_2, we note that

$$I_{\varepsilon 0\tau,j} - I_{\varepsilon 0\sigma,j} = |v_{\sigma\tau,j}|^2 + v_{\sigma\tau,j} \bar{v}_{\sigma,j} + v_{\sigma,j} \bar{v}_{\sigma\tau,j}$$

and obtain bound

$$E\left|Y_2(\tau)-Y_2(\sigma)\right|^4 \le CE\left|\frac{1}{m}\sum_{j=1}^{m}\pi_j\left|v_{\sigma\tau,j}\right|^2\right|^4 + 2CE\left|\frac{1}{m}\sum_{j=1}^{m}\pi_j v_{\sigma\tau,j}\bar{v}_{\sigma,j}\right|^4.$$

The first term on the right is

$$\begin{aligned}
&\frac{C}{m^4}\sum_{j,k,\ell,p=1}^{m}\pi_j\pi_k\pi_\ell\pi_p E\left|v_{\sigma\tau,j}v_{\sigma\tau,k}v_{\sigma\tau,\ell}v_{\sigma\tau,p}\right|^2\\
&\le \frac{C}{m^4}\sum_{j,k,\ell,p=1}^{m}\left|\pi_j\pi_k\pi_\ell\pi_p\right|\left(E\left|v_{\sigma\tau,j}\right|^8 E\left|v_{\sigma\tau,k}\right|^8 E\left|v_{\sigma\tau,\ell}\right|^8 E\left|v_{\sigma\tau,p}\right|^8\right)^{\frac{1}{4}}\\
&= C\left(\frac{1}{m}\sum_{j=1}^{m}\left|\pi_j\right|\left(E\left|v_{\sigma\tau,j}\right|^8\right)^{\frac{1}{4}}\right)^4 \le C(\tau-\sigma)^4\left(\frac{1}{m}\sum_{j=1}^{m}\left|\pi_j\right|\right)^4.
\end{aligned}$$

The second term on the right is

$$\begin{aligned}
&\frac{C}{m^4}\sum_{j,k,\ell,p=1}^{m}\pi_j\pi_k\pi_\ell\pi_p E v_{\sigma\tau,j}\bar{v}_{\sigma,j}\bar{v}_{\sigma\tau,k}v_{\sigma,k}v_{\sigma\tau,\ell}\bar{v}_{\sigma,\ell}\bar{v}_{\sigma\tau,p}v_{\sigma,p}\\
&\le \frac{C}{m^4}\sum_{j,k,\ell,p=1}^{m}\left|\pi_j\pi_k\pi_\ell\pi_p\right|\\
&\times\left(E\left|v_{\sigma\tau,j}\right|^8 E\left|\bar{v}_{\sigma,j}\right|^8 E\left|\bar{v}_{\sigma\tau,k}\right|^8 E\left|v_{\sigma,k}\right|^8 E\left|v_{\sigma\tau,\ell}\right|^8 E\left|\bar{v}_{\sigma,\ell}\right|^8 E\left|\bar{v}_{\sigma\tau,p}\right|^8 E\left|v_{\sigma,p}\right|^8\right)^{\frac{1}{8}}\\
&\le C(\tau-\sigma)^2\left(\frac{1}{m}\sum_{j=1}^{m}\left|\pi_j\right|\right)^4.
\end{aligned}$$

Proceeding as when bounding $E\left|Y_1(\tau)-Y_1(\sigma)\right|^4$, it can be shown that $m^{-1}\sum_{j=1}^{m}\left|\pi_j\right| \le C$, so $Y_2(\tau)$ is indeed tight. It follows that process Y is tight and that part (a) holds. The proof of parts (b), (c) and (d) is similar.

Examining the proofs, it can be seen that the convergence holds uniformly over $-1+\varepsilon \le a \le D$.

Parta (b)–(f) can be proved in a similar way. ■

Proof of Proposition 1. To prove that $\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| \xrightarrow{p} 0$ under the null, it is sufficient to prove that for any $\zeta>0$ there exists $\eta>0$ such that

$$P\left(\inf_{\tau\in[\tau_\ell,1]}\inf_{d\in\left[-\frac{1}{2},\frac{1}{2}\right],|d-\delta|\ge\eta}\left(R(d,I_{0\tau})-R(\delta,I_{0\tau})\right)\ge\zeta\right)\to 1 \qquad \text{as } T\to\infty, \tag{24}$$

where

$$R\left(d, I_{0\tau}\right)=\log \left(\frac{1}{m} \sum_{j=1}^{m} I_{0\tau, j}\left(\frac{j}{m}\right)^{2d}\right)-\frac{2d}{m} \sum_{j=1}^{m} \log \left(\frac{j}{m}\right).$$

Define

$$\ell\left(d, I_{0\tau}\right)=\frac{1}{m} \sum_{j=1}^{m} \frac{I_{0\tau}\left(\lambda_{j}\right)}{G \lambda_{j}^{-2\delta}}\left(\frac{j}{m}\right)^{2(d-\delta)}$$

and write

$$R\left(d, I_{0\tau}\right)-R\left(\delta, I_{0\tau}\right)=\log \ell\left(d, I_{0\tau}\right)-\log \ell\left(\delta, I_{0\tau}\right)-2(d-\delta) \frac{1}{m} \sum_{j=1}^{m} \log \left(\frac{j}{m}\right).$$

By Lemma 5 with $\mu_{j}=(j/m)^{2(d-\delta)}$, for any $0<\varepsilon \leq 1$,

$$\sup_{\delta-\frac{1}{2}+\frac{\varepsilon}{2} \leq d \leq \frac{1}{2}}\left|\ell\left(d, I_{0\tau}\right)-\frac{\tau}{1+2(d-\delta)}\right| \Longrightarrow 0.$$

Therefore uniformly in $\tau \in\left[\tau_{\ell}, 1\right]$ and $d \in\left[\delta-\frac{1}{2}+\frac{\varepsilon}{2}, \frac{1}{2}\right]$, $|d-\delta| \geq \eta$, as $T \rightarrow \infty$,

$$\begin{aligned}
& R\left(d, I_{0\tau}\right)-R\left(\delta, I_{0\tau}\right) \\
& =\log \left(\frac{\tau}{1+2(d-\delta)}\right)-\log (\tau)-2(d-\delta) \frac{1}{m} \sum_{j=1}^{m} \log \left(\frac{j}{m}\right)+o_{p}(1) \\
& =-\log (1+2(d-\delta))+2(d-\delta)+o_{p}(1) \geq c+o_{p}(1) \qquad (25)
\end{aligned}$$

where $c>0$, because $m^{-1} \sum_{j=1}^{m} \log (j/m)=-1+o(1)$ and because $\log (1+x)<x$ for all $|x|>0$.

On the other hand, uniformly in $\tau \in\left[\tau_{\ell}, 1\right]$ and $d \in\left[-\frac{1}{2}, \delta-\frac{1}{2}+\frac{\varepsilon}{2}\right]$, as $T \rightarrow \infty$,

$$\begin{aligned}
& R\left(d, I_{0\tau}\right)-R\left(\delta, I_{0\tau}\right) \\
& \geq \log \ell\left(\delta-\frac{1}{2}+\frac{\varepsilon}{2}, I_{0\tau}\right)-\log \ell\left(\delta, I_{0\tau}\right)+2(d-\delta)+o(1) \\
& =\log \left(\frac{\tau}{1+(-1+\varepsilon)}\right)-\log (\tau)+o_{p}(1)+2(d-\delta)+o(1) \\
& =-\log \varepsilon+2(d-\delta)+o_{p}(1) \geq-\log \varepsilon-2+o_{p}(1) \geq c+o_{p}(1) \qquad (26)
\end{aligned}$$

when ε is small. Bounds (25) and (26) imply that condition (24) is satisfied.

The proof that $\sup_{\tau \in\left[0, \tau_{h}\right]}\left|\widehat{\delta}_{2}(\tau)-\delta\right| \xrightarrow{p} 0$ is similar. Finally,

$$\begin{aligned}
& \sup_{\tau \in\left[\tau_{l}, \tau_{h}\right]}\left|\widehat{\delta}_{1}(\tau)-\widehat{\delta}_{2}(\tau)\right|=\sup_{\tau \in\left[\tau_{l}, \tau_{h}\right]}\left|\widehat{\delta}_{1}(\tau)-\delta+\delta-\widehat{\delta}_{2}(\tau)\right| \\
& \leq \sup_{\tau \in\left[\tau_{l}, 1\right]}\left|\widehat{\delta}_{1}(\tau)-\delta\right|+\sup_{\tau \in\left[0, \tau_{h}\right]}\left|\widehat{\delta}_{2}(\tau)-\delta\right| \rightarrow_{p} 0.
\end{aligned}$$

■

6.2 Asymptotic distribution of test statistic under the null

Proposition 2 *Under Assumptions 1–4 and under the null with $\delta_1 = \delta_2 = \delta$,*

$$2\sqrt{m}\begin{pmatrix} \widehat{\delta}_1(\tau) - \delta \\ \widehat{\delta}_2(\tau) - \delta \end{pmatrix} \Longrightarrow \begin{pmatrix} \frac{1}{\tau} B(\tau) \\ \frac{1}{1-\tau}\left(B(1) - B(\tau)\right) \end{pmatrix}$$

on $\tau \in [\tau_\ell, \tau_h]$, so that

$$2\sqrt{m}\left(\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)\right) \Longrightarrow \frac{B(\tau) - \tau B(1)}{\tau(1-\tau)}$$

and

$$4m\tau(1-\tau)\left(\widehat{\delta}_1(\tau) - \widehat{\delta}_2(\tau)\right)^2 \Longrightarrow \frac{\left(B(\tau) - \tau B(1)\right)^2}{\tau(1-\tau)}.$$

Lemma 6 *Under Assumptions 1–4 and under the null with $\delta_1 = \delta_2 = \delta$,*

(a)
$$\widehat{\delta}_1(\tau) - \delta = -\left(1 + o_p(1)\right)\frac{1}{2\tau}\frac{1}{m}\sum_{j=1}^{m}\nu_j \frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}},$$

where $o_p(1)$ is uniform over $\tau \in [\tau_\ell, 1]$,

(b)
$$\widehat{\delta}_2(\tau) - \delta = -\left(1 + o_p(1)\right)\frac{1}{2(1-\tau)}\frac{1}{m}\sum_{j=1}^{m}\nu_j \frac{I_{\tau 1,j}}{G\lambda_j^{-2\delta}},$$

where $o_p(1)$ is uniform over $\tau \in [0, \tau_h]$.

Proof of Lemma 6. (a) The proof is an extension of the proof of bound (11) in Theorem 1 of Dalla et al. (2006). Write

$$\frac{\partial R(d, I_{0\tau})}{\partial d} = \frac{T(d, I_{0\tau})}{V(d, I_{0\tau})}$$

where

$$T(d, I_{0\tau}) = \frac{2}{m}\sum_{j=1}^{m}\frac{I_{0\tau}(\lambda_j)}{\lambda_j^{-2\delta}}\left(\frac{j}{m}\right)^{2(d-\delta)}\nu_j,$$

$$V(d, I_{0\tau}) = \frac{1}{m}\sum_{j=1}^{m}\frac{I_{0\tau}(\lambda_j)}{\lambda_j^{-2\delta}}\left(\frac{j}{m}\right)^{2(d-\delta)},$$

where v_j is defined in (14). By the mean value theorem,

$$\widehat{\delta}_1(\tau) - \delta = \left(\frac{\partial T\left(\widetilde{\delta}(\tau), I_{0\tau}\right)}{\partial d} \right)^{-1} \left(T\left(\widehat{\delta}_1(\tau), I_{0\tau}\right) - T(\delta, I_{0\tau}) \right),$$

where $\widetilde{\delta}_1(\tau)$ is an intermediate point between δ and $\widehat{\delta}_1(\tau)$. Since $\mathbb{I}\left(\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| > \varepsilon\right) = o_p(1)$ for any $\varepsilon > 0$ by Proposition 1, we have

$$\widehat{\delta}_1(\tau) - \delta = \left(\widehat{\delta}_1(\tau) - \delta\right) \mathbb{I}\left(\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| \le \varepsilon\right) + o_p(1).$$

Let $0 < \varepsilon < \min\left\{\frac{1}{2}-\delta, \frac{1}{2}+\delta\right\}$. When $\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| \le \varepsilon$, Lemma 5 implies that

$$V\left(\widehat{\delta}_1(\tau), I_{0\tau}\right) \ge \frac{G}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2\varepsilon}\frac{I_{0\tau}(\lambda_j)}{G\lambda_j^{-2\delta}} \Longrightarrow \frac{G\tau}{1+2\varepsilon} > 0 \qquad \text{for all } \tau_\ell \le \tau \le 1.$$

We also have $\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| \in (-1/2, 1/2)$, therefore $\frac{\partial R}{\partial d}\left(\widehat{\delta}_1(\tau), I_{0\tau}\right) = 0$, $T\left(\widehat{\delta}_1(\tau), I_{0\tau}\right) = 0$ and

$$\left(\widehat{\delta}_1(\tau) - \delta\right) \mathbb{I}\left(\sup_{\tau\in[\tau_\ell,1]}\left|\widehat{\delta}_1(\tau)-\delta\right| \le \varepsilon\right) = -\left(\frac{\partial T\left(\widetilde{\delta}(\tau), I_{0\tau}\right)}{\partial d} \right)^{-1} T(\delta, I_{0\tau}). \qquad (27)$$

From Lemma 5 with $\mu_j = \log\left(\frac{j}{m}\right)\left(\frac{j}{m}\right)^a \nu_j$ and from Proposition 1 it follows that

$$\frac{\partial T\left(\widetilde{\delta}_1(\tau), I_{0\tau}\right)}{\partial d} = \frac{4G}{m}\sum_{j=1}^{m}\log\left(\frac{j}{m}\right)\left(\frac{j}{m}\right)^{2\left(\widetilde{\delta}_1(\tau)-\delta\right)}\nu_j\frac{I_{0\tau}(\lambda_j)}{G\lambda_j^{-2\delta}} \Longrightarrow 4G\tau.$$

because $\left|\widetilde{\delta}_1(\tau)-\delta\right| \le \left|\hat{\delta}_1(\tau)-\delta\right|$. Therefore the right-hand side of equation (27) is equal to

$$-\left(4G\tau\left(1+o_p(1)\right)\right)^{-1} T(\delta, I_{0\tau}) = -\frac{T(\delta, I_{0\tau})}{4G\tau} + o_p(1)$$

uniformly over $[\tau_l, 1]$ and part (a) is established.

Part (b) is proved similarly. ■

Lemma 7 *Under Assumptions 1–4,*

$$\text{(a)} \qquad \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j} \Longrightarrow B(\tau) \qquad \tau\in[0,1], \qquad (28)$$

(b) $$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{2\pi}{\sigma_\varepsilon^2}w_{\varepsilon\tau\tau^*,j}\overline{w}_{\varepsilon\tau^*1,j}\Longrightarrow 0 \qquad \tau\in[0,\tau^*],$$

(c) $$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{2\pi}{\sigma_\varepsilon^2}w_{\varepsilon 0\tau^*,j}\overline{w}_{\varepsilon\tau^*\tau,j}\Longrightarrow 0 \qquad \tau\in[\tau^*,1].$$

Proof of Lemma 7. (a) The left-hand side of (28) can be written as

$$Y_T(\tau)=\sum_{t=2}^{\lfloor\tau T\rfloor}\varepsilon_t\left(\sum_{s=1}^{t-1}\varepsilon_s c_{t-s}\right)=\sum_{t=2}^{\lfloor\tau T\rfloor}\varepsilon_t d_t,$$

where $d_t=\sum_{s=1}^{t-1}\varepsilon_s c_{t-s}$ for $t\geq 2$ and $d_1=0$, with

$$c_s=\frac{2}{\sigma_\varepsilon^2\sqrt{m}T}\sum_{j=1}^{m}\nu_j\cos(s\lambda_j)$$

and $\nu_j=\ln j-\frac{1}{m}\sum_{k=1}^{m}\ln k$. The realizations of the process Y_T belong to the space $D[0,1]$ of real functions which are right continuous with left-hand limits. The sequence $\{\varepsilon_t d_t|\mathcal{F}_t, 1\leq t\leq T\}$ is a martingale difference sequence with respect to the σ-algebras $\mathcal{F}_t$ of events generated by ε_s, $s\leq t$. The first two moments of the process Y_T are $EY_T(\tau)=0$ and

$$\begin{aligned}E|Y_T(\tau)|^2&=\sigma_\varepsilon^4\sum_{t=2}^{\lfloor\tau T\rfloor}\sum_{s=1}^{t-1}c_{t-s}^2\\&=\frac{1}{mT^2}\sum_{j,l=1}^{m}\nu_j\nu_\ell\sum_{t,s=1}^{\lfloor\tau T\rfloor}\left(e^{i(t-s)(\lambda_j-\lambda_k)}+e^{i(t-s)(\lambda_j+\lambda_k)}\right)\\&=\frac{2\pi}{mT}\sum_{j,l=1}^{m}\nu_j\nu_\ell\left(K_{\lfloor\tau T\rfloor}(\lambda_j-\lambda_k)+K_{\lfloor\tau T\rfloor}(\lambda_j+\lambda_k)\right).\end{aligned}$$

By Lemma 3,

$$E|Y(\tau)|^2=\sigma_\varepsilon^4\sum_{t=2}^{\lfloor\tau T\rfloor}\sum_{s=1}^{t-1}c_{t-s}^2\to\tau \quad \text{for } \tau\in[0,1]. \tag{29}$$

The variance of the process Y_T therefore increases asymptotically linearly in t and the weak convergence of the process Y_T in (28) holds if the following two conditions of Theorem 2 of Scott (1973) are satisfied:

(A) $\sum_{t=1}^{\lfloor\tau T\rfloor}E(\varepsilon_t^2 d_t^2|\mathcal{F}_{t-1})\xrightarrow{p}\tau$ as $T\to\infty$, $0<\tau\leq 1$ and
(B) $\sum_{t=1}^{T}E(\varepsilon_t^2 d_t^2\mathbb{I}(\varepsilon_t^2 d_t^2\geq\delta)|\mathcal{F}_{t-1})\xrightarrow{p}0$ as $T\to\infty$, for any $\delta>0$,

where $\mathbb{I}(A)$ denotes the indicator function of the set A.

We first prove that condition (A) holds. By Assumption 1,

$$\sum_{t=1}^{\lfloor\tau T\rfloor} E\left(\varepsilon_t^2 d_t^2 | \mathcal{F}_{t-1}\right) = \sigma_\varepsilon^2 \sum_{t=1}^{\lfloor\tau T\rfloor} d_t^2 = \sigma_\varepsilon^2 \sum_{t=1}^{\lfloor\tau T\rfloor} \sum_{s,r=1}^{t-1} \varepsilon_s \varepsilon_r c_{t-s} c_{t-r} \tag{30}$$

which has expectation $\sigma_\varepsilon^4 \sum_{t=1}^{\lfloor\tau T\rfloor} \sum_{s=1}^{t-1} c_{t-s}^2$. This expression converges to τ for $0 < \tau \leq 1$ by (29). The second moment of the right-hand side of (30) is

$$\left(\sigma_\varepsilon^4 \sum_{t=1}^{\lfloor\tau T\rfloor} \sum_{s=1}^{t-1} c_{t-s}^2\right)^2 + \sigma_\varepsilon^4 \kappa_4 \sum_{t,s=1}^{\lfloor\tau T\rfloor} \sum_{s=1}^{t\wedge r-1} c_{t-s}^2 c_{r-s}^2 + 2\sigma_\varepsilon^8 \sum_{t,s=1}^{\lfloor\tau T\rfloor} \left(\sum_{s=1}^{t\wedge r-1} c_{t-s} c_{r-s}\right)^2, \tag{31}$$

where $\kappa_4 = cum\left(\varepsilon_t, \varepsilon_t, \varepsilon_t, \varepsilon_t\right)$ and where $a \wedge b$ denotes $\min\{a, b\}$. The first term of (31) converges to τ^2 as implied by (29). The second term of (31) is bounded by

$$C \sum_{t,r,s=1}^{T} c_{t-s}^2 c_{r-s}^2 = \frac{C}{m^2 T^4} \sum_{\delta_1,\ldots,\delta_4} \sum_{t,s,r=1}^{T} \sum_{j,k,l,q=1}^{m} \nu_j \nu_k \nu_l \nu_q e^{i(\delta_1\lambda_j + \delta_2\lambda_k)(t-s)} e^{i(\delta_3\lambda_l + \delta_4\lambda_q)(r-s)}$$

$$\begin{aligned} &= \frac{C}{m^2 T^4} \sum_{\delta_1,\ldots,\delta_4} \sum_{j,k,l,q=1}^{m} \nu_j \nu_k \nu_l \nu_q \sum_{s=1}^{T} e^{i(\delta_1\lambda_j + \delta_2\lambda_k + \delta_3\lambda_l + \delta_4\lambda_q)s} \\ &\times \sum_{t=1}^{T} e^{i(\delta_1\lambda_j + \delta_2\lambda_k)t} \sum_{r=1}^{T} e^{i(\delta_3\lambda_l + \delta_4\lambda_q)r}, \end{aligned}$$

where the sum is over all $(\delta_1, \delta_2, \delta_3, \delta_4) \in \{1, -1\}^4$. The last displayed expression is bounded in absolute value by

$$\frac{C \log^4 m}{m^2 T^3} \sum_{\delta_1,\ldots,\delta_4} \sum_{j,k,l,q=1}^{m} \left|D_T\left(\delta_1\lambda_j + \delta_2\lambda_k\right)\right| \left|D_T\left(\delta_3\lambda_l + \delta_4\lambda_q\right)\right|.$$

because $|\nu_j| \leq C \log m$. Since $\left|D_T(\lambda_j)\right| \leq CTj^{-1}$ and $|D_T(0)| \leq T$ (see the inequalities (16) and (17)), the last displayed expression is bounded by

$$\begin{aligned} &\frac{C \log^4 m}{m^2 T^3} \sum_{\delta_1,\ldots,\delta_4} \sum_{j,k,l,q=1}^{m} \frac{1}{\left|\delta_1\lambda_j + \delta_2\lambda_k\right|_+} \frac{1}{\left|\delta_3\lambda_l + \delta_4\lambda_q\right|_+} \\ &\leq \frac{C \log^4 m}{m^2 T} \sum_{\delta_1,\ldots,\delta_4} \sum_{j,k,l,q=1}^{m} \frac{1}{\left|\delta_1 j + \delta_2 k\right|_+} \frac{1}{\left|\delta_3 l + \delta_4 q\right|_+}, \end{aligned}$$

where for any real number a, we denote $|a|_+ = \max\{1, |a|\}$. Inequalities in Lemma 4 and Assumption 4 imply that the second term of (31) is bounded by

$$\frac{C\log^4 m}{m^2T}m^2\log^2 m = \frac{C\log^6 m}{T} = o\left(1\right).$$

Due to the symmetry of c_t, the third term of (31) is bounded by $C\sum_{t,s,r,p=1}^{\lfloor\tau T\rfloor} c_{t-s}c_{r-s}c_{t-p}c_{r-p}$ which is equal to

$$\begin{aligned}
&\frac{C}{m^2T^4}\sum_{\delta_1,\ldots,\delta_4}\sum_{t,s,r,p=1}^{\lfloor\tau T\rfloor}\sum_{j,k,l,q=1}^{m}\nu_j\nu_k\nu_l\nu_q e^{i(\delta_1\lambda_j+\delta_3\lambda_l)t}e^{i(\delta_2\lambda_k+\delta_4\lambda_q)r}\\
&\times e^{-i(\delta_1\lambda_j+\delta_2\lambda_k)s}e^{-i(\delta_3\lambda_l+\delta_4\lambda_q)p}\\
&=\frac{C\log^4 m}{m^2T^4}\sum_{\delta_1,\ldots,\delta_4}\sum_{j,k,l,q=1}^{m}\nu_j\nu_k\nu_l\nu D_{\lfloor\tau T\rfloor}\left(-\delta_1\lambda_j-\delta_2\lambda_k\right)D_{\lfloor\tau T\rfloor}\left(-\delta_3\lambda_l-\delta_4\lambda_q\right)\\
&\times D_{\lfloor\tau T\rfloor}\left(\delta_1\lambda_j+\delta_3\lambda_l\right)D_{\lfloor\tau T\rfloor}\left(\delta_2\lambda_k+\delta_4\lambda_q\right)\\
&\leq\frac{C\log^4 m}{m^2}\sum_{\delta_1,\ldots,\delta_4}\sum_{j,k,l,q=1}^{m}\frac{1}{|\delta_1 j+\delta_2 k|_+}\frac{1}{|\delta_3 l+\delta_4 q|_+}\frac{1}{|\delta_1 j+\delta_3 l|_+}\frac{1}{|\delta_2 k+\delta_4 q|_+}.
\end{aligned}$$

Using inequalities in Lemma 4, the last expression is bounded by

$$\frac{C\log^4 m}{m^2}\log^4 m = o\left(1\right).$$

Gathering results, we conclude that for arbitrary $0<\tau\leq 1$, the first and second moment of $\sum_{t=1}^{\lfloor\tau T\rfloor}E\left(\varepsilon_t^2 d_t^2|\mathcal{F}_{t-1}\right)$ converge in probability to τ and τ^2, respectively, and that condition (A) is satisfied.

Next we prove that condition (B) holds true. We write

$$\sum_{t=1}^{T}d_t^2E\left(\varepsilon_t^2\mathbb{I}\left(\varepsilon_t^2d_t^2\geq\delta\right)|\mathcal{F}_{t-1}\right)\leq\max_{1\leq t\leq T}E\left(\varepsilon_t^2\mathbb{I}\left(\varepsilon_t^2d_t^2\geq\delta\right)|\mathcal{F}_{t-1}\right)\sum_{t=1}^{T}d_t^2.$$

The sum $\sum_{t=1}^{T}d_t^2$ is $O_p\left(1\right)$ because $\sigma_\varepsilon^2E\sum_{t=1}^{T}d_t^2=\sigma_\varepsilon^4\sum_{t=1}^{T}\sum_{s=1}^{t-1}c_{t-s}^2\to 1$ by (29). To bound the maximum in the above inequality, we first bound $\max_{1\leq t\leq T}d_t^2$. We note that

$$\max_{1\leq t\leq T}d_t^2=\left(\max_{1\leq t\leq T}d_t^4\right)^{1/2}\leq\left(\sum_{t=1}^{T}d_t^4\right)^{1/2}.$$

The expectation of $\sum_{t=1}^{T}d_t^4$ equals

$$\kappa_4\sum_{t=1}^{T}\sum_{s=1}^{t-1}c_{t-s}^4+3\sigma_\varepsilon^4\sum_{t=1}^{T}\sum_{s=1}^{t-1}\sum_{r=1}^{t-1}c_{t-s}^2c_{t-r}^2. \tag{32}$$

The first term of (32) is bounded by

$$C\sum_{t,s=1}^{T}c_{t-s}^{4}=\frac{C}{m^{2}T^{4}}\sum_{\delta_1,\ldots,\delta_4}\sum_{j,k,l,q=1}^{m}\nu_j\nu_k\nu_l\nu_q\left|\sum_{t=1}^{T}e^{i(\delta_1\lambda_j+\delta_2\lambda_k+\delta_3\lambda_l+\delta_4\lambda_q)t}\right|^{2}$$
$$\leq\frac{C\log^4 m}{m^2T^4}\sum_{j,k,l,q=1}^{m}T^2\leq\frac{C\log^4 m}{m^2T^2}m^4=\frac{Cm^2\log^4 m}{T^2}=o\left(1\right)$$

by Assumption 4.

The second term of (32) is bounded by $C\sum_{t,s,r=1}^{T}c_{t-s}^{2}c_{t-r}^{2}$, which, like the second term of (31), is $o_p\left(1\right)$. We deduce that $E\sum_{t=1}^{T}d_t^4=o\left(1\right)$ and so $\max_{1\leq t\leq T}d_t^2=o_p\left(1\right)$. Assumption 1 implies that ε_t^2 are uniformly integrable (a sufficient condition for the uniform integrability is that $E\left|\varepsilon_t\right|^{2+\alpha}<C$ for some $0<\alpha,C<\infty$). The uniform integrability together with the fact that $\max_{1\leq t\leq T}d_t^2=o_p\left(1\right)$ imply that $\max_{1\leq t\leq T}E\left(\varepsilon_t^2\mathbb{I}\left(\varepsilon_t^2d_t^2\geq\delta\right)|\mathcal{F}_{t-1}\right)=o_p\left(1\right)$ for any $\delta>0$ and therefore that the condition (B) is satisfied.

The proof for parts (b) and (c) follows in a similar way. ■

Lemma 8 *Under Assumptions 1–4 and under the null hypothesis with* $\delta=\delta_1=\delta_2$,

$$\left(\begin{array}{c}\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}\\ \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{I_{\tau1,j}}{G\lambda_j^{-2\delta}}\end{array}\right)\Longrightarrow\left(\begin{array}{c}B\left(\tau\right)\\ B\left(1\right)-B\left(\tau\right)\end{array}\right)\qquad\tau\in\left[\tau_l,\tau_h\right].$$

Proof of Lemma 8. By Lemma 7, it is sufficient to prove that for $\tau\in\left[\tau_l,\tau^*\right]$,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon0\tau,j}\right)\Longrightarrow0,\tag{33}$$

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{\tau\tau^*,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon\tau\tau^*,j}\right)\Longrightarrow0,\tag{34}$$

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{w_{\tau\tau^*,j}}{G^{\frac12}\lambda_j^{-\delta}}\frac{\overline{w}_{\tau^*1,j}}{G^{\frac12}\lambda_j^{-\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}w_{\varepsilon\tau\tau^*,j}\overline{w}_{\varepsilon\tau^*1,j}\right)\Longrightarrow0,\tag{35}$$

and that for $\tau\in\left[\tau^*,\tau_h\right]$,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{\tau^*\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon\tau^*\tau,j}\right)\Longrightarrow0,\tag{36}$$

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{\tau1,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon\tau1,j}\right)\Longrightarrow0,\tag{37}$$

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{w_{0\tau^*,j}}{G^{\frac12}\lambda_j^{-\delta}}\frac{\overline{w}_{\tau^*\tau,j}}{G^{\frac12}\lambda_j^{-\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}w_{\varepsilon0\tau^*,j}\overline{w}_{\tau^*\tau1,j}\right)\Longrightarrow0.\tag{38}$$

We prove the convergence in (33), the convergence in (34)–(38) can be shown using similar arguments. Proceeding as in the proof of (4.8) of Robinson (1995) while employing Lemma 3 of Lazarová (2005) and referring to (12), we obtain that

$$\sum_{j=1}^{k}\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)=O_p\left(k^{\frac{1}{3}}\log^{\frac{2}{3}}k+k^{\beta+1}T^{-\beta}+k^{\frac{1}{2}}T^{-\frac{1}{4}}\right)$$

uniformly over $\tau\in[\tau_\ell,\tau^*]$ and $1\le k\le m$. Using summation by parts, we get

$$\begin{aligned}\left|\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right| &\le \frac{1}{\sqrt{m}}\sum_{k=1}^{m-1}\left|\nu_k-\nu_{k+1}\right|\left|\sum_{j=1}^{k}\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right|\\ &\quad+\left|\nu_m\right|\left|\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\right|.\end{aligned}$$

The first term is

$$\frac{1}{\sqrt{m}}\sum_{k=1}^{m-1}\frac{1}{k}O_p\left(k^{\frac{1}{3}}\log^{\frac{2}{3}}k+k^{\beta+1}T^{-\beta}+k^{\frac{1}{2}}T^{-\frac{1}{4}}\right)=o_p(1)$$

and the second term is

$$m^{-\frac{1}{2}}\log m O_p\left(m^{\frac{1}{3}}\log^{\frac{2}{3}}m+m^{\beta+1}T^{-\beta}+m^{\frac{1}{2}}T^{-\frac{1}{4}}\right)=o_p(1).$$

Therefore

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)\xrightarrow{p}0$$

for $\tau\in[\tau_l,\tau^*]$.

Next we prove tightness of the process on the left hand side of (33). Write

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{I_{0\tau,j}}{G\lambda_j^{-2\delta}}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right)=Y_1(\tau)+Y_2(\tau), \tag{39}$$

where

$$\begin{aligned}Y_1(\tau)&=\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{f_j}{g_j}\left(\frac{I_{0\tau,j}}{f_j}-\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}\right),\\ Y_2(\tau)&=\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left(\frac{f_j}{g_j}-1\right)\frac{2\pi}{\sigma_\varepsilon^2}I_{\varepsilon 0\tau,j}.\end{aligned}$$

From the proof of Lemma 5 we can see that

$$E\left|Y_1(\tau)-Y_1(\sigma)\right|^4 \leq C(\tau-\sigma)^2\left(m^{-\frac{1}{2}}\sum_{j=1}^{m}\left|\nu_j\right|\frac{f_j}{g_j}j^{-\frac{1}{2}}\right)^4.$$

The expression in the last bracket is bounded,

$$m^{-1}\sum_{j=1}^{m}\left|\nu_j\right|\frac{f_j}{g_j}\left(\frac{j}{m}\right)^{-\frac{1}{2}} \leq \max_{1\leq j\leq m}\frac{f_j}{g_j}m^{-1}\sum_{j=1}^{m}\left|\nu_j\right|\left(\frac{j}{m}\right)^{-\frac{1}{2}} \leq C,$$

because $|\nu_j| \leq C\log m$ and $\max_{1\leq j\leq m}\frac{f_j}{g_j} \leq C$, therefore

$$E\left|Y_1(\tau)-Y_1(\sigma)\right|^4 \leq C(\tau-\sigma)^2$$

and $Y_1(\tau)$ is tight. Further, proceeding as in the proof of Lemma 5 we obtain

$$E\left|Y_2(\tau)-Y_2(\sigma)\right|^4 \leq C(\tau-\sigma)^2\left(m^{-\frac{1}{2}}\sum_{j=1}^{m}\left|\nu_j\right|\left|\frac{f_j}{g_j}-1\right|\right)^4.$$

Now $|f_j/g_j-1| \leq C(j/T)^{\beta}$ by Assumption 3, therefore

$$m^{-\frac{1}{2}}\sum_{j=1}^{m}\left|\nu_j\right|\left|\frac{f_j}{g_j}-1\right| \leq Cm^{-\frac{1}{2}}T^{-\beta}\log m\sum_{j=1}^{m}j^{\beta} = O\left(\frac{m^{\beta+\frac{1}{2}}}{T^{\beta}}\log m\right) = o(1)$$

by Assumption 4, so $Y_2(\tau)$ is tight and (33) holds. ■

Proof of Proposition 2. The proposition follows from Lemma 6 and Lemma 8. ■
Proof of Theorem 1. The theorem follows from Proposition 2 and from the continuous mapping theorem. ■

6.3 Power of test

Lemma 9 *Under Assumptions 1–4 and under the null hypothesis with $\delta_1 \neq \delta_2$, as $T \to \infty$,*

$$\widehat{\delta}_1(\tau^*) \xrightarrow{p} \delta_1 \qquad \text{and} \qquad \widehat{\delta}_2(\tau^*) \xrightarrow{p} \delta_2,$$

so that

$$\widehat{\delta}_1(\tau^*) - \widehat{\delta}_2(\tau^*) \xrightarrow{p} \delta_1 - \delta_2.$$

Proof of Lemma 9. The lemma can be proved using the same strategy as in the proof of Lemma 5. ■

Proof of Theorem 2. The theorem follows from Lemma 9. ■

References

[1] Abadir, K.M., W. Distaso and L. Giraitis, 2007. Non-stationarity extended local Whittle estimation, *Journal of Econometrics* 141, 1353–1384.

[2] Abadir, K.M., W. Distaso and L. Giraitis, 2011. An I(d) model with trend and cycles, *Journal of Econometrics* 163, 186-199.

[3] Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point, *Econometrica* 61, 821–856.

[4] Beran, J., and N. Terrin, 1996. Testing for a change of the long-memory parameter, *Biometrika* 83, 627–638.

[5] Bernanke, B.S., and I. Mihov, 1996. What does the Bundesbank target?, *NBER Working Paper* No. 5764.

[6] Billingsley, P., 1999. *Convergence of probability measures*, second ed., Wiley, New York.

[7] Brillinger, D.R., 1981. *Time series: data analysis and theory*, Holden-Day, San Francisco.

[8] Busetti, F., and A.M.R. Taylor, 2004. Tests of stationarity against a change in persistence, *Journal of Econometrics* 123, 33–66.

[9] Dalla, V., L. Giraitis and J. Hidalgo, 2006. Consistent estimation of the memory parameter for nonlinear time series, *Journal of Time Series Analysis* 27, 211–251.

[10] De Long, J.B., and L.H. Summers, 1988. How does macroeconomic policy affect output?, *Brookings Papers on Economic Activity* 2, 433–494.

[11] Frömmel, M., and R. Kruse, 2012. Testing for a rational bubble under long memory, *Quantitative Finance* 12, 1723–1732.

[12] Hakkio, C.S., and M. Rush, 1991. Is the budget deficit "too large"?, *Economic Inquiry* 29, 429–445.

[13] Halunga, A.G., D.R. Osborn, and M. Sensier, 2009. Changes in the order of integration of US and UK inflation, *Economics Letters* 102, 30–32.

[14] Harvey, D., S.J. Leybourne and A.M.R. Taylor, 2006. Modified tests for a change in persistence, *Journal of Econometrics* 134, 441–469.

[15] Hassler, U., and B. Meller, 2014. Detecting multiple breaks in long memory: the case of U.S. inflation, *Empirical Economics* 46, 653–680.

[16] Hassler, U., P.M.M. Rodrigues, and A. Rubia, 2014. Persistence in the banking industry: fractional integration and breaks in memory? *Journal of Empirical Finance* 29, 95–112.

[17] Henry, M., 2001. Robust automatic bandwidth for long memory, *Journal of Time Series Analysis* 22, 293–316.

[18] Horváth, L., and Q.M. Shao, 1999. Limit theorems for quadratic forms with applications to Whittle's estimates, *Journal of Applied Probability* 9, 146–187.

[19] Hurvich, C.M., and W.W. Chen, 2000. An efficient taper for potentially overdifferenced time series, *Journal of Time Series Analysis* 21, 155-180.

[20] Hualde, J., and F. Iacone, 2017. Revisiting inflation in the euro area allowing for long memory, *Economics Letters* 156, 145–150.

[21] Iacone, F., S. Leybourne and A.M.R. Taylor, 2014. A fixed-b test for a break in level at an unknown time under fractional integration, *Journal of Time Series Analysis* 35, 40–54.

[22] Kim, J.Y., 2000. Detection of a change in persistence in a linear time series, *Journal of Econometrics* 95, 97–116.

[23] Kim, J.Y., J. Belaire-Franch, and R. Badilli-Amador, 2002. Corrigendum to 'Detection of a change in persistence in a linear time series', *Journal of Econometrics* 109, 389–392.

[24] Kumar, M. S. and T. Okimoto, 2007. Dynamics of persistence in international inflation rates, *Journal of Money, Credit and Banking* 39, 1457–1479.

[25] Lavancier, F., R. Leipus, A. Philippe, and D. Surgailis, 2013. Detection of nonconstant long memory parameter, *Econometric Theory* 29, 1009–1056.

[26] Lazarová, Š., 2005. Testing for structural change in regression with long memory processes, *Journal of Econometrics* 129, 329–372.

[27] Leybourne, S., A.M.R. Taylor, and T. Kim, 2007. CUSUM of squares-based tests for a change in persistence, *Journal of Time Series Analysis* 28, 408–33.

[28] Martins, L.F., and P.M.M. Rodrigues, 2014. Testing for persistence change in fractionally integrated models: an application to world inflation rates, *Computational Statistics and Data Analysis* 76, 502-522.

[29] McCullagh, P., 1987. *Tensor Methods in Statistics*, Chapman and Hall, London.

[30] Qu, Z., 2011. A test against spurious long memory, *Journal of Business and Economic Statistics* 29, 423-438.

[31] Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence, *Annals of Statistics* 23, 1630–1661.

[32] Scott, D.J.,1973. Central limit theorems for martingales and for processes with stationary increments using a Skorokhod representation approach, *Advances in Applied Probability* 5, 119–137.

[33] Shimotsu, K., 2006. Simple (but effective) tests of long memory versus structural breaks, Queen's Economics Department, Working Paper 1101.

[34] Sibbertsen, P., and R. Kruse, 2009. Testing for a break in persistence under long-range dependencies, *Journal of Time Series Analysis* 30, 263–285.

[35] Sibbertsen, P., C. Wegener, and T. Basse, 2014. Testing for a break in the persistence in yield spreads of EMU government bonds, *Journal of Banking and Finance* 41, 109–118.

[36] Sollis, R., 2006. Testing for bubbles: an application of tests for change in persistence, *Applied Financial Economics* 16, 491–498.

[37] Velasco, C., 1999. Gaussian semiparametric estimation of non-stationary time series, *Journal of Time Series Analysis* 20, 87-127.

Figure 1. German monthly inflation, at annual rate

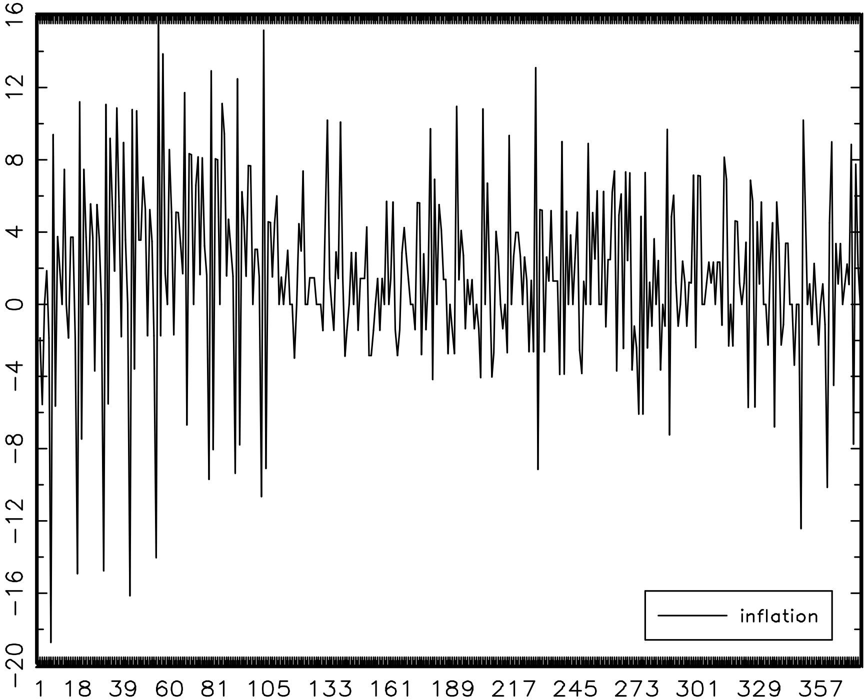

Figure 2. German inflation, sample autocorrelation

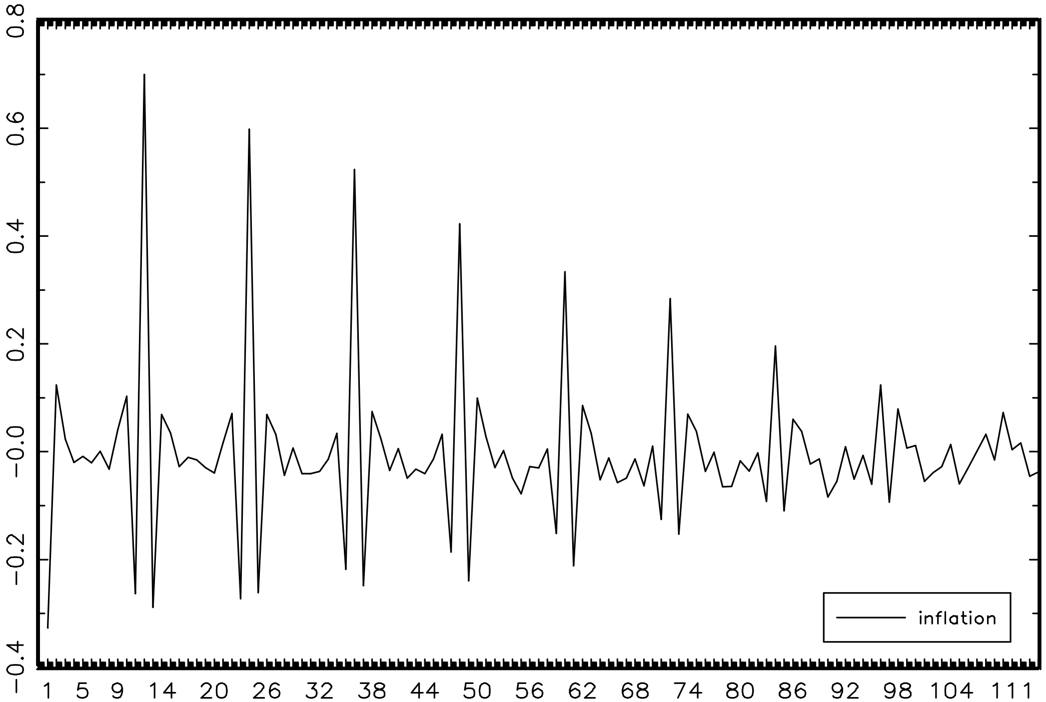

Figure 3. German inflation, periodogram