Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Econometrics 129 (2005) 329–372
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Testing for structural change in regression with long memory processes

Štěpána Lazarová*

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

Abstract

The paper considers tests for structural change in time series regression models where both regressors and residuals may exhibit long range dependence. The limiting distribution of the test statistic depends on unknown parameters. While the unknown parameters can be consistently estimated and asymptotic critical values obtained by simulation, the paper proposes an alternative approach of approximating the distribution of the test statistic by a bootstrap procedure. The asymptotic validity of bootstrap is shown and the performance of the testing procedure is examined in a simple Monte Carlo experiment.

JEL classification: C22

Keywords: Structural change; Long memory; Bootstrap

1. Introduction

Parameter instability and structural change have been a subject of a large body of statistical and econometric literature. The maintained hypothesis of parameter stability has been tested both against specified and unspecified forms of alternative hypothesis. When employed as a model-diagnostic tool, stability tests are constructed against all possible functions describing the evolution of parameters over time. Such tests are based

*Tel.: +44 20 7882 5000; fax: +44 20 8983 3580.
E-mail address: s.lazarova@lse.ac.uk (Š. Lazarová).

doi:10.1016/j.jeconom.2004.09.011

on the behaviour of regression residuals, as in CUSUM tests of Brown et al. (1975) and Ploberger and Krämer (1990, 1992), or on the behaviour of parameter estimates, as in the fluctuation tests of Sen (1980) or Ploberger et al. (1989).

Alternatively, parameter stability tests can be designed against a specific alternative. Example of specific alternatives are one-time change in parameters as in the papers by Quandt (1960) or Andrews (1993), or parameters following random walk (Nyblom, 1989). Though constructed to detect specific parameter behaviour, these tests are usually shown to have power against a broader range of departures from the null of parameter constancy.

This paper considers tests for stability in slope coefficients in linear regression model where both regressors and errors are allowed to be long range dependent. The main contribution of the paper is twofold. First, the limiting distribution of the test statistics considered in the literature is typically a functional of Brownian motion. It is shown that this remains true for test statistics based on the slope coefficient estimator in linear model with stationary long memory series. Secondly, as an alternative to computing the critical values for the test statistic, a first-order bootstrap approximation of the distribution of the test statistic is proposed and the validity of the bootstrap procedure is shown.

The paper is organized as follows. Section 2 describes the model and the hypotheses of interest and states distributional results for the test statistic. Section 3 proposes a bootstrap approximation of the testing procedure and shows its validity. Section 4 offers a Monte Carlo study of the small sample performance of the bootstrap testing procedure. Section 5 concludes. The proofs of the results stated in the text are gathered in Section 6.

Throughout the paper, B denotes a p-dimensional vector of independent standard Brownian motions on $[0,1]$ or on a set $\Lambda \subset (0,1)$, $[\cdot]$ signifies integer part, $\overline{z}$ means the conjugate of a complex number z, $\|\cdot\|$ denotes the Euclidean norm of a matrix, $\mathbb{I}(\cdot)$ is the indicator function of a set, "$\Longrightarrow$" denotes weak convergence in the space $\mathrm{D}(\Lambda)^p$ of p-vectors of right-continuous functions with left-hand limits, endowed with the uniform topology. For any real numbers a, b, $a \vee b = \max\{a,b\}$ and $a \wedge b = \min\{a,b\}$. Starred notation in E^*, var*, cov* and similar refers to quantities conditional on data, taken with respect to the corresponding bootstrap probability measure. The statement $y_T \sim x_T$ is equivalent to the statement $y_T/x_T \to 1$ as $T \to \infty$. For generic functions f and g, $f_j = f(\lambda_j)$ where $\lambda_j = 2\pi j/T$, $j = 1,\ldots,T$ are Fourier frequencies, and $g_t = g(t/T)$ for $t = 1,\ldots,T$. For σ-algebras $\mathscr{F}$, $\mathscr{G}$, $\mathscr{F} \vee \mathscr{G}$ is their union, that is the smallest σ-algebra containing all elements of $\mathscr{F}$ and $\mathscr{G}$. Finally, C and D stand for generic constants.

2. Model and asymptotic results

We are interested in testing for structural change in regression models with processes that may possess long memory. We consider the model

$$y_t = \alpha + \beta_t' x_t + u_t, \tag{1}$$

where y_t is the observed dependent variable, α is an unknown intercept, β_t is a p-dimensional vector of unknown parameters, x_t is a p-dimensional vector of observations on the explanatory variables and u_t is an unobserved stochastic disturbance. Our hypothesis of interest is whether the parameter vector β_t stays constant,

$$\mathrm{H}_0: \quad \beta_t = \beta \quad \text{for some } \beta, \text{ for all } t = 1, \ldots, T.$$

The alternative is that of general parameter instability,

$$\mathrm{H}_1: \quad \beta_t \neq \beta_s \quad \text{for some } 1 < t, \; s < T.$$

Test procedures for the hypothesis of structural stability of general models are based on test statistics that can be written as

$$Z_T = \phi(E_T)$$

where E_T is a stochastic process on $[0, 1]$ or its subset with values in the space of right-continuous functions with left-hand limits and ϕ is a continuous functional. The process E_T is based on an estimator of parameters of a given model and its form reflects the choice of the testing principle. For example, if $\{e_t, p \leqslant t \leqslant T\}$ is the sequence of cumulative recursive residuals from the OLS estimates of the model (1) under the null as in the CUSUM test procedure of Brown et al. (1975), the stochastic process E_T can be defined as $E_T = \{E_T(\tau) = e_{[\tau T]}, \; p/T \leqslant \tau \leqslant 1\}$. Further examples of processes considered in the literature are Wald-, LM- and LR-like test statistic processes of Andrews (1993), CUSUM of squares process of Brown et al. (1975), OLS CUSUM process of Ploberger and Krämer (1992), OLS parameter estimates process of Ploberger et al. (1989) and Sen (1980) or MOSUM process of Chu et al. (1994).

The functional ϕ measures the excess fluctuation of the process E_T with respect to its hypothesised fluctuation. Depending on the belief about the form of the alternative, the functional ϕ can be chosen to obtain good power of the test. A functional widely used in literature is the supremum functional. The test statistic can also be based on the L_q-distance like Cramér–von Mises test statistic with $q = 2$. The range functional can have power advantage over the supremum functional in detecting smaller fluctuations of a process which changes its sign, as argued by Kuan and Hornik (1995). The average exponential functional of Andrews and Ploberger (1994) is shown to enjoy asymptotic optimality with respect to a weighted average power criterion.

In this paper, we base the test procedure on the OLS estimators of the coefficient δ in the model

$$y_t = \alpha + \beta' x_t + \delta' z_t + u_t \tag{2}$$

where

$$z_t = z_t(\tau) = \begin{cases} x_t, & t \leqslant [\tau T], \\ 0 & \text{otherwise}, \end{cases} \tag{3}$$

where δ is a p-dimensional vector of parameters and where τ lies in a subset Λ of $(0,1)$. In the interest of clarity, the explicit notation of dependence of z_t on τ is sometimes dropped in what follows. The choice $\Lambda = (0,1)$ appears natural but for technical reasons the set Λ needs to be restricted to have closure in $(0,1)$. The grounds for the restriction are discussed after stating Theorem 2 and its Corollary 1. In addition to technical reasons, there may be other motives for restricting the set Λ away from $(0,1)$. It may be suspected that the instability in question occurred in a specific subperiod of a given period. For example, if data for postwar productivity growth are examined, the attention might be focused on testing for an abrupt or gradual change in a period around the 1973 oil price shock.

For any fixed $\tau \in \Lambda$, the OLS estimator of the parameters β and δ in (2) is given by

$$\begin{pmatrix} \hat{\beta}(\tau) \\ \hat{\delta}(\tau) \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{T} (x_t - \overline{x})x_t' & \sum_{t=1}^{T} (x_t - \overline{x})z_t' \\ \sum_{t=1}^{T} (z_t - \overline{z})x_t' & \sum_{t=1}^{T} (z_t - \overline{z})z_t' \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^{T} (x_t - \overline{x})y_t \\ \sum_{t=1}^{T} (z_t - \overline{z})y_t \end{pmatrix}, \tag{4}$$

where $\overline{x} = T^{-1}\sum_{t=1}^{T} x_t$ and $\overline{z} = T^{-1}\sum_{t=1}^{[\tau T]} x_t$. Alternatively, model (2) can be translated into the frequency domain, becoming

$$w_y(\lambda_j) = \beta' w_x(\lambda_j) + \delta' w_z(\lambda_j) + w_u(\lambda_j), \quad j = 1, \ldots, T-1, \tag{5}$$

where

$$w_d(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} d_t \mathrm{e}^{\mathrm{i}t\lambda}$$

is the discrete Fourier transform of a sequence of p-dimensional vectors $d_1, \ldots, d_T$ and $\lambda_j = 2\pi j/T$ are the Fourier frequencies. Identifying $w_x(\lambda_j)$ and $w_z(\lambda_j)$ as regressors and $w_u(\lambda_j)$ as an error term, the OLS estimate of the parameters β and δ in (5) for $\tau \in \Lambda$ is given by

$$\begin{pmatrix} \hat{\beta}(\tau) \\ \hat{\delta}(\tau) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^{T-1} I_{xy}(\lambda_j) \\ \sum_{j=1}^{T-1} I_{zy}(\lambda_j) \end{pmatrix}, \tag{6}$$

where for any vector processes u_t, v_t,

$$I_{uv}(\lambda) = w_u(\lambda)\overline{w}_v'(\lambda)$$

is the cross-periodogram matrix. Leaving out the zero frequency from the frequency domain regression is equivalent to mean-correcting data before running the regression in the time domain. The estimators defined in (4) and (6) are therefore identical. Omission of the zero frequency permits inference on the slope parameters when the intercept is unknown. It is worth noting that due to the symmetry of the periodograms, (6) is equal to

$$\left(Re \begin{pmatrix} \sum_{j=1}^{[T/2]} I_{xx}(\lambda_j) & \sum_{j=1}^{[T/2]} I_{xz}(\lambda_j) \\ \sum_{j=1}^{[T/2]} I_{zx}(\lambda_j) & \sum_{j=1}^{[T/2]} I_{zz}(\lambda_j) \end{pmatrix} \right)^{-1} Re \begin{pmatrix} \sum_{j=1}^{[T/2]} I_{xy}(\lambda_j) \\ \sum_{j=1}^{[T/2]} I_{zy}(\lambda_j) \end{pmatrix} \tag{7}$$

for T odd while when T is even, (7) differs from (6) only by the order of $\mathrm{O_p}(1/T)$.

For each τ from a set $\Lambda \subset (0,1)$, an estimator $\hat{\delta}(\tau)$ of δ can be obtained from (6) and a process $\hat{\delta}$ can be defined as $\hat{\delta} = \{\hat{\delta}(\tau), \tau \in \Lambda\}$. For any T and any realization of processes $\{x_t\}$ and $\{u_t\}$ the function $\hat{\delta}$ is bounded and constant on the subintervals $[j/T, (j+1)/T) \cap \Lambda$, $j \in N$, and the process $\hat{\delta}$ is a random element of the space $D(\Lambda)^p$ of $p \times 1$ vectors of right-continuous functions on Λ with left-hand limits endowed with uniform metric.

The test statistic based on the process $\hat{\delta}$ is then $Z_T = \phi(\sqrt{T}\hat{\delta})$ for any continuous functional $\phi : D(\Lambda)^p \mapsto R$. For example, the Kolmogorov–Smirnov (or Bartlett) test statistic is defined as

$$\mathrm{KS}_T = \sup_{\tau \in \Lambda} \sqrt{T} \|\hat{\delta}(\tau)\|$$

and the Cramér–von Mises statistic is given by

$$\mathrm{CvM}_T = \int_{\Lambda} T \|\hat{\delta}(\tau)\|^2 \, \mathrm{d}\tau$$

where $\|\cdot\|$ is the Euclidean norm. Under the null hypothesis, the additional regressor z_t has no explanatory power and the process $\hat{\delta}$ is uniformly close to zero, whereas under the alternative, $\hat{\delta}$ can be expected to differ significantly from zero on a set $\Lambda_1 \subset \Lambda$ of Lebesgue measure greater than zero. The norm functionals like KS and CvM constitute one-tailed tests, rejecting H_0 for large values of the test statistic. In principle, two-tailed tests can be constructed for functionals whose range includes both positive and negative values.

It can be expected that the test procedure based on model (2) has power mainly against one-time break alternatives of the form

$$\mathrm{H}_1: \quad \beta_t = \begin{cases} \beta + \delta, & t = 1, \ldots, [\tau_0 T], \\ \beta, & t = [\tau_0 T] + 1, \ldots, T \end{cases} \tag{8}$$

for some $\tau_0 \in \Lambda$ and some constants $\delta \neq 0$, β, but we show that our test procedure has power under a broader range of alternatives.

Our analysis proceeds under the following assumptions. It is assumed that $\{x_t\}$ and $\{u_t\}$ are covariance stationary linear processes that satisfy Conditions 1–5:

Condition 1.

$$x_t = \sum_{j=0}^{\infty} a_j \xi_{t-j}, \quad \sum_{j=0}^{\infty} ||a_j||^2 < \infty, \quad a_0 = I,$$

$$u_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \quad b_0 = 1.$$

Let $\mathscr{F}_t$ and $\mathscr{G}_t$ be the σ-algebras of events generated by ξ_s, $s \leqslant t$, and ε_s, $s \leqslant t$, respectively.

Condition 2. $\{\xi_t\}$ *is a stochastic process that satisfies*

1. $\mathrm{E}(\xi_t|\mathscr{F}_{t-1} \vee \mathscr{G}_t) = 0$ *a.s.*,
2. $\mathrm{E}(\xi_t\xi_t'|\mathscr{F}_{t-1} \vee \mathscr{G}_t) = \mathrm{E}(\xi_t\xi_t') = \Xi$ *a.s.*, *and*
3. *the joint fourth cumulants of* $\xi_{t_ij_i}, j_i = 1, \ldots, p$ *and* $i = 1, \ldots, 4$, *where* ξ_{tj} *denotes the* *jth component of the vector* ξ_t, *satisfy*

$$cum(\xi_{t_1j_1}, \xi_{t_2j_2}, \xi_{t_3j_3}, \xi_{t_4j_4}|\mathscr{G}_T) = \begin{cases} \kappa_{\xi,j_1,j_2,j_3,j_4} \ a.s. & t_1 = t_2 = t_3 = t_4, \\ 0 \ a.s. & otherwise, \end{cases}$$

with $|\kappa_\xi| = \max_{j_i=1,\ldots,p,i=1,\ldots,4}|\kappa_{\xi,j_1,j_2,j_3,j_4}| < \infty$.

Condition 3. $\{\varepsilon_t\}$ *is a stochastic process that satisfies*

1. $\mathrm{E}(\varepsilon_t|\mathscr{F}_t \vee \mathscr{G}_{t-1}) = 0$ *a.s.*,
2. $\mathrm{E}(\varepsilon_t^2|\mathscr{F}_t \vee \mathscr{G}_{t-1}) = \mathrm{E}(\varepsilon_t^2) = \sigma_\varepsilon^2$ *a.s.*, *and*
3. *the joint fourth cumulant of* ε_{t_i}, $i = 1, \ldots, 4$ *satisfies*

$$cum(\varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3}, \varepsilon_{t_4}|\mathscr{F}_T) = \begin{cases} \kappa \ a.s. & t_1 = t_2 = t_3 = t_4, \\ 0 \ a.s. & otherwise \end{cases}$$

with $|\kappa| < \infty$.

Condition 4. *The functions*

$$A(\lambda) = \sum_{j=0}^{\infty} a_j \mathrm{e}^{\mathrm{i}j\lambda} \quad and \quad B(\lambda) = \sum_{j=0}^{\infty} b_j \mathrm{e}^{\mathrm{i}j\lambda}$$

satisfy the following assumptions:

1. *there exist constants* $0 < C_{x,k}, C_u < \infty$ *and* $d_{x,k}, d_u \in [0, \frac{1}{2})$, $k = 1, 2, \ldots, p$, *such that* $|A_{kk}(\lambda)| \sim C_{x,k}\lambda^{-d_{x,k}}$, $|B(\lambda)| \sim C_u\lambda^{-d_u}$ *as* $\lambda \to 0+$,
2. $A(\lambda)$ *and* $B(\lambda)$ *are differentiable on* $(0, \pi]$ *and* $\|\mathrm{d}A(\lambda)/\mathrm{d}\lambda\| = \mathrm{O}(\|A(\lambda)\|/\lambda)$, $|\mathrm{d}B(\lambda)/\mathrm{d}\lambda| = \mathrm{O}(|B(\lambda)|/\lambda)$ *uniformly over* $(0, \pi]$ *and*
3. $\|A(\lambda)\| > 0$ *and* $|B(\lambda)| > 0$ *for* $\lambda \in (0, \pi]$.

Condition 5.

$$\int_{-\pi}^{\pi} \|f_{xx}(\lambda)f_{uu}(\lambda)||\,\mathrm{d}\lambda < \infty, \qquad \mathrm{E}(x_tx_t') > 0,$$

where $f_{xx}(\lambda)$ *and* $f_{uu}(\lambda)$ *are spectral densities of processes* x_t *and* u_t, *respectively*.

The conditions are similar to those used by Robinson (1995a, b, 1998) and Hidalgo (2003). A further remark is that while the fourth moments are assumed constant, the third moments are free to vary and so only second order stationarity is required.

Conditions 1–3 imply homoskedasticity of regressors and errors. This assumption could presumably be relaxed to allow for a certain degree of heterogeneity. Conditions 1–3 also imply that x_t and u_s are uncorrelated for all t and s and that $\mathrm{E}(x_tu_tx_su_s) = \mathrm{E}(x_tx_s)\mathrm{E}(u_tu_s)$ for all t and s and therefore that the spectral density of

$x_t u_t$ at frequency zero is $2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda)\,\mathrm{d}\lambda$ if Condition 5 holds. One of the reasons for imposing the condition $\mathrm{E}(x_t u_t x_s u_s) = \mathrm{E}(x_t x_s)\mathrm{E}(u_t u_s)$ is that it allows us to use

$$\frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx}(\lambda_j) I_{uu}(\lambda_j)$$

of Robinson (1998) to consistently estimate $2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda)\,\mathrm{d}\lambda$ without having to select a bandwidth. If the condition $\mathrm{E}(x_t u_t x_s u_s) = \mathrm{E}(x_t x_s)\mathrm{E}(u_t u_s)$ is not valid, the long run variance of $x_t u_t$ has an additional component which is a function of the fourth cumulants and which is not estimated by $\hat{\Omega}$. When x_t and u_t are short memory processes, the results of Taniguchi (1982) and Keenan (1987) can be used to estimate the additional component of variance, but no estimation methods are available for long memory time series. Relaxing condition $\mathrm{E}(x_t u_t x_s u_s) = \mathrm{E}(x_t x_s)\mathrm{E}(u_t u_s)$ would thus come at a price of a considerable amount of technical work. Therefore, though assumption of no correlation between regressors and errors is admittedly somewhat restrictive and excludes for example some cases of interest studied by cointegration literature, we do not attempt to relax this assumption.

Condition 4 allows for a possible singularity at the zero frequency but the results of this paper could be generalized to the case of a singularity at a nonzero frequency or of more than one singularity. The validity of the bound $|\mathrm{d}B(\lambda)/\mathrm{d}\lambda| = \mathrm{O}(|B(\lambda)|/\lambda)$ implies that $|\mathrm{d}f_{uu}(\lambda)/\mathrm{d}\lambda| = \mathrm{O}(f_{uu}(\lambda)/\lambda)$ since $f_{uu} = \sigma_\varepsilon^2/2\pi|B(\lambda)|^2$. Similar implication holds for the spectral density matrix f_{xx}. Examples of scalar processes that satisfy Condition 4 are FARIMA model of Granger and Joyeux (1980) or Hosking (1981), and fractional Gaussian noise of Mandelbrot and van Ness (1968). These models satisfy $f(\lambda) \sim C\lambda^{-2d}$ as $\lambda \to 0+$ for some memory parameter $d \in [0, \frac{1}{2})$.

Condition 5 has been used by Robinson (1994) and Robinson and Hidalgo (1997). The condition restricts the collective memory of regressors and errors. For regressors with long memory parameter d_x and errors with long memory parameter d_u, Condition 5 imposes restriction $d_x + d_u < \frac{1}{2}$. This condition ensures that the standard least-squares estimation procedure of the slope coefficients is $\sqrt{T}$-consistent and leads to a Gaussian limit distribution (Robinson, 1994). As Hidalgo (2003) remarks, the first part of Condition 5 seems to be very mild and appears to be necessary and minimal for the central limit theorem for OLS estimates of slope coefficient to hold. In a related proposition of Giraitis and Surgailis (1990) an analogous condition is required for convergence of quadratic forms in linear processes. The validity of the CLT carries over to a functional CLT in the present paper.

The main result of this section can now be stated. Let B be a vector of p independent standard Brownian motion processes restricted to Λ where Λ is a subset of $[0,1]$ with closure in $(0,1)$.

Theorem 1. *Under Conditions* 1–5 *and under the null hypothesis*,

$$\sqrt{T}\begin{pmatrix} \hat{\beta}(\tau) - \beta \\ \hat{\delta}(\tau) \end{pmatrix} \Longrightarrow \frac{1}{\tau(1-\tau)} \begin{pmatrix} \Sigma^{-1}\Omega^{1/2}(\tau B(1) - \tau B(\tau)) \\ \Sigma^{-1}\Omega^{1/2}(B(\tau) - \tau B(1)) \end{pmatrix}$$

on Λ, *where* $\Omega = 2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda)\,\mathrm{d}\lambda$ *and* $\Sigma = \mathrm{E}(x_t x_t')$.

Theorem 1 implies in particular that

$$\sqrt{T}\hat{\delta}(\tau) \Longrightarrow \frac{1}{\tau(1-\tau)}\Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1)),$$

so that for each fixed $\tau \in \Lambda$,

$$\sqrt{T}\hat{\delta}(\tau) \xrightarrow{\mathrm{d}} \mathrm{N}(0, V(\tau)) \tag{9}$$

where

$$V(\tau) = \frac{1}{\tau(1-\tau)}\Sigma^{-1}\Omega\Sigma^{-1}.$$

It is interesting to note that when x_t or u_t are long memory processes, the limiting distribution remains to be a function of a Brownian motion rather than of a fractional Brownian motion that often arises in asymptotic results in long memory environment. A result that is crucial for validity of Theorem 1 is that $T^{-1/2}\sum_{j=1}^{T-1} I_{zu}(\lambda_j)$, which is asymptotically proportional to the partial sum $T^{-1/2}\sum_{t=1}^{[\tau T]} x_t u_t$, converges weakly to a Brownian motion. To achieve weak convergence of the partial sum $\sum_{t=1}^{[\tau T]} x_t$ for strongly dependent process x_t, normalization by $T^{-1/2-d}$ is required and the limiting process is a fractional Brownian motion. However, the case of the partial sum $\sum_{t=1}^{[\tau T]} x_t u_t$ is different. Intuitively, while the memory of the processes x_t and u_t is of a long range, their product $x_t u_t$ displays short memory behaviour. This phenomenon may be regarded as analogous to that of Robinson (1998) where the sample autocovariances of processes x_t and u_t are stochastically dampening each other in his estimator of Ω.

To assess the power of the test procedure, we examine limiting behaviour of the process $(\hat{\beta}(\tau)', \hat{\delta}(\tau)')'$ under alternatives. We restrict ourselves to the local alternatives

$$\beta_t = \beta + \frac{1}{\sqrt{T}} h\left(\frac{t}{T}\right) \quad \text{for some } \beta, \tag{10}$$

where h is a p-dimensional vector of bounded variation functions on $[0, 1]$. This class of alternatives comprises many types of structural change that may be of interest. For instance, a function $h(\tau) = \delta \mathbb{I}(\tau_0 \leqslant \tau)$ describes the alternative of an abrupt break of size δ at time τ_0. A step function h defines multiple structural breaks. A function h consisting of two constant segments connected by a smooth curve depicts smooth transition between two steady levels of a parameter. A general smooth function h captures continual change of the parameter.

For the limiting distribution under local alternatives the following result is obtained.

Theorem 2. *Under Conditions* 1–5 *and under the local alternative hypothesis* (10),

$$\sqrt{T}\begin{pmatrix}\hat{\beta}(\tau)-\beta \\ \hat{\delta}(\tau)\end{pmatrix} \Longrightarrow \frac{1}{\tau(1-\tau)}\begin{pmatrix}\Sigma^{-1}\Omega^{1/2}(\tau B(1)-\tau B(\tau)) \\ \Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1))\end{pmatrix} + \frac{1}{\tau(1-\tau)}\begin{pmatrix}\tau\int_\tau^1 h(u)\,\mathrm{d}u \\ (\int_0^\tau h(u)\,\mathrm{d}u-\tau\int_0^1 h(u)\,\mathrm{d}u)\end{pmatrix}$$

for $\tau \in \Lambda$.

By the continuous mapping theorem, an immediate consequence of Theorem 2 is the following corollary.

Corollary 1. *Let* ϕ *be a continuous functional on* $D(\Lambda)^p$. *Let* $Z_T = \phi(\sqrt{T}\hat{\delta}(\tau))$ *and*

$$Z_h = \phi\left(\frac{1}{\tau(1-\tau)}\Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1)) + \frac{1}{\tau(1-\tau)}\left(\int_0^\tau h(u)\,\mathrm{d}u - \tau\int_0^1 h(u)\,\mathrm{d}u\right)\right).$$

Under the conditions of Theorem 2,

$$Z_T \xrightarrow{\mathrm{d}} Z_h.$$

The corollary shows that a test based on Z_T has nontrivial local power against a broad range of alternatives. The limiting random variable Z_h is indexed by functions h specifying local alternatives. Under the null, when $h\equiv 0$, the test statistic Z_T converges in distribution to Z_0,

$$\phi\left(\sqrt{T}\hat{\delta}(\tau)\right) \xrightarrow{\mathrm{d}} \phi\left(\frac{1}{\tau(1-\tau)}\Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1))\right).$$

The asymptotic test at a significance level α is based on a critical region C_α constructed from the asymptotic null distribution, $P(Z_0 \in C_\alpha) = \alpha$. The asymptotic test rejects the null when $Z_T \in C_\alpha$.

The form of the limiting distributions in Theorems 1 and 2 explains the reason for the necessity of bounding the set Λ away from 0 and 1. The restriction on Λ guarantees that the convergence of the estimator $\hat{\delta}$, which is the basis of the test statistic, is uniform. Moreover, it can be shown that for $\Lambda = (0,1)$ many functionals, including the sup- and L_q-norms, diverge to infinity in probability.

The trimming restriction on Λ can be avoided by allowing the limiting distribution of the test statistic to be of a different form than a functional of the Brownian bridge. The results of Jaeschke (1979) and Eicker (1979) suggest that the supremum of $\hat{\delta}(\tau)$ taken over subsets of $(0,1)$ increasing towards $(0,1)$ at an appropriate speed and normalized by a suitable centring and rescaling sequences should converge to an extreme value distribution. However, relaxing the restriction on Λ in such a way comes at a cost. The rate of convergence of the test statistics to the extreme value distribution can be expected to be very slow. The asymptotic critical values would not be appropriate for tests in samples of moderate size and an elaborate bootstrap procedure would be required to improve on the performance of the asymptotic test. We do not pursue this possibility in the current paper.

The variance of the process $(B(\tau)-\tau B(1))/(\tau(1-\tau))$,

$$\operatorname{var}\frac{B(\tau)-\tau B(1)}{\tau(1-\tau)}=\frac{1}{\tau(1-\tau)},$$

varies across Λ. This means that under the null, the probability that the process $\|\hat{\delta}(\tau)\|$ crosses any vertical line above the real axis is smallest at $\tau=\frac{1}{2}$. This may lead us to inquire whether the power of the test based on supremum and other functionals can be improved by levelling the variance of the estimated process $\hat{\delta}$ across Λ. Given the restriction of Λ away from $(0,1)$, we may normalize the process $\hat{\delta}$ by multiplying it by $[\tau(1-\tau)]^{1/2}$. By Theorem 1, under the null,

$$[\tau(1-\tau)]^{1/2}\sqrt{T}\hat{\delta}(\tau)\Longrightarrow\frac{1}{[\tau(1-\tau)]^{1/2}}\Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1))$$

whose variance is equal to $\Sigma^{-1}\Omega\Sigma^{-1}$ across Λ. The rejection probabilities of the test based on the levelled process $\hat{\delta}$ in samples of moderate size is examined in a Monte Carlo experiment in Section 4.

Our test procedure is based on the behaviour of the OLS estimator of β coefficients. At the core of the limit behaviour of the test statistics lies the fact that $T^{-1/2}\sum_{j=1}^{T-1}w_{z(\tau)}(\lambda_j)\overline{w}_{\hat{u}}(\lambda_j)$ converges weakly to a Brownian motion process. Using this fact, the asymptotic behaviour of other tests based on the behaviour of OLS slope coefficient estimators can be obtained. For example, if $\hat{\beta}_{t_1}^{t_2}$ is the OLS estimator of β in the regression $y_t=\alpha+\beta' x_t+u_t$ for $t=t_1,\dots,t_2$, then under the local alternative (10)

$$\tau\sqrt{T}(\hat{\beta}_1^{[\tau T]}-\hat{\beta}_1^{T})\Longrightarrow\Sigma^{-1}\Omega^{1/2}(B(\tau)-\tau B(1))+\left(\int_0^\tau h(u)\,\mathrm{d}u-\tau\int_0^1 h(u)\,\mathrm{d}u\right)$$

in correspondence with the results of Ploberger et al. (1989). If $\hat{\Sigma}$ and $\hat{\Omega}$ are consistent estimates of Σ and Ω, then the Wald-statistic process based on partial sample slope estimators has limiting distribution

$$T\left(\hat{\beta}_1^{[\tau T]}-\hat{\beta}_{[\tau T]+1}^{T}\right)'\left(\frac{\Sigma^{-1}\Omega\Sigma^{-1}}{\tau(1-\tau)}\right)^{-1}\left(\hat{\beta}_1^{[\tau T]}-\hat{\beta}_{[\tau T]+1}^{T}\right)\Longrightarrow J(\tau)'J(\tau),$$

where

$$\begin{aligned}J(\tau)={}&\frac{1}{[\tau(1-\tau)]^{1/2}}(B(\tau)-\tau B(1))\\&+\frac{1}{[\tau(1-\tau)]^{1/2}}\Omega^{-1/2}\Sigma\left(\int_0^\tau h(u)\,\mathrm{d}u-\tau\int_0^1 h(u)\,\mathrm{d}u\right)\end{aligned}$$

as in Andrews (1993).

On the other hand, the limiting distribution of tests based on behaviour of the OLS residuals depends crucially on the weak convergence of $T^{-1/2}\sum_{t=1}^{T}\hat{u}_t$ to a limiting process. Under long memory, the asymptotic properties of this sum can be expected to be different than under short memory.

3. Bootstrap procedure

The limiting distribution of the process $\hat{\delta}$ in (9) depends on unknown parameters Ω and Σ. The process $\hat{\delta}$ can be normalized by consistent estimates $\hat{\Omega}$, $\hat{\Sigma}$ of these parameters. Such consistent estimates are for example

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} x_t x_t' \tag{11}$$

and

$$\hat{\Omega} = \frac{4\pi^2}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) I_{\hat{u}\hat{u}}(\lambda_j). \tag{12}$$

Consistency of $\hat{\Sigma}$ follows from ergodicity of x_t in the variance implied by Conditions 1 and 2. The estimator $\hat{\Omega}$ is based on results of Robinson (1998) and its consistency is asserted in the following theorem.

Theorem 3. *Under Conditions* 1–5 *and under the local alternative*,

$$\hat{\Omega} \xrightarrow{p} \Omega.$$

The normalized process $\tilde{\delta}(\tau) = \hat{\Omega}^{-1/2}\hat{\Sigma}\hat{\delta}(\tau)$ has a limiting distribution which is free of nuisance parameters,

$$\sqrt{T}\tilde{\delta}(\tau) \Longrightarrow \frac{B(\tau) - \tau B(1)}{\tau(1-\tau)}.$$

In special cases, distributions of functionals of Brownian motion are known analytically and quantiles of the distributions can be easily computed. Examples are supremum of a Brownian motion and supremum of a Brownian bridge. In other instances, critical values have been computed by simulation and tabulated, as in case of the supremum of a standardized tied-down Bessel process in Andrews (1993). However, in majority of cases, the critical values of the test statistic need to be simulated by the researcher.

One alternative to computing asymptotical critical values by simulation is to employ a bootstrap procedure. The core idea of bootstrap is to replace the unknown distribution of a random variable by the empirical distribution of a random sample drawn from that distribution. However, when the data are not independent and identically distributed, the basic bootstrap of Efron (1979) is not valid. In the time series context, an early adaptation of the basic bootstrap method rests on the assumption that the data are generated by a finite-order stationary ARMA process with independent identically distributed innovations (Efron and Tibshirani, 1993). In a direction towards nonparametric methods, Bühlmann (1997, 1998) approximates the linear infinite-dimensional process by a sieve of finite-dimensional autoregressive processes whose order is growing with the sample size. Diebold et al. (1998) propose a purely nonparametric bootstrap method. Their Cholesky factor bootstrap replaces estimates of parametric models with nonparametric estimation of the dynamics.

A different way of approximately preserving the temporal dependence structure of the data is to resample blocks of data. Carlstein (1986) and Künsch (1989) propose to resample from nonoverlapping and overlapping blocks of data, respectively, and to concatenate the blocks to generate a bootstrap sample. Politis and Romano (1992) introduce an idea of subsampling, regarding blocks of data—subseries—as new pseudo-samples.

A problem shared by nonparametric bootstrap methods is that they require an intervention by the researcher in choosing a dimension parameter of the procedure, be it lag length, bandwidth or block length. The performance of time-series bootstrap can be highly sensitive to the choice of the dimension parameter, particularly in samples of moderate size. Although automatic procedures for choosing the dimension have been devised for some methods, they can be computationally expensive.

Nonparametric bootstrap procedures can alternatively be carried out in the frequency domain where either frequency domain data or their squares, that is the discrete Fourier coefficients or periodograms, can be bootstrapped. This approach is motivated by the observation that converting a stochastic process from time domain to frequency domain reduces serial correlation of the process though it induces heteroskedasticity. Bootstrap method of Ramos (1984) for Fourier coefficients or Franke and Härdle (1992) and Dahlhaus and Janas (1996) for periodograms require a consistent estimate of the spectral density and therefore a choice of a bandwidth. Local periodogram bootstrap of Paparoditis and Politis (2000) avoids the need for estimating the spectrum but again demands a bandwidth choice.

Hidalgo (2003) proposes a method that eliminates the dimension choice. He suggests to bootstrap OLS residuals in frequency domain. His bootstrap procedure is easy to implement and computationally inexpensive. Moreover, it is one the first bootstrap procedures shown to be valid for long memory time series in a fairly general context, adding to a still thin body of the literature on long memory time series bootstrap.

In this paper we propose to approximate the critical values of the testing procedure described in Section 2 by a bootstrap procedure based on the ideas of Hidalgo (2003). The procedure consists of the following steps.

Step 1: Compute OLS estimates $\hat{\beta}(\tau)$ and $\hat{\delta}(\tau)$ from (4) or (6) for $\tau \in \Lambda$. Compute $\hat{\tau} = \arg\max_{\tau\in\Lambda}\|\hat{\delta}(\tau)\|$, the OLS estimates $\hat{\beta} = \hat{\beta}(\hat{\tau})$ and $\hat{\delta} = \hat{\delta}(\hat{\tau})$ and the OLS residuals

$$\hat{u}_t = y_t - \hat{\beta}' x_t - \hat{\delta}' z_t(\hat{\tau}).$$

Step 2: Compute

$$w_{\hat{u}}(\lambda_j) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} \hat{u}_t e^{it\lambda_j} \quad \text{for} \quad j = 1, \ldots, T-1$$

and

$$\tilde{w}_{\hat{u}}(\lambda_j) = \frac{w_{\hat{u}}(\lambda_j) - (1/(T-1))\sum_{k=1}^{T-1} w_{\hat{u}}(\lambda_k)}{(1/(T-1))\sum_{j=1}^{T-1} |w_{\hat{u}}(\lambda_j) - (1/(T-1))\sum_{k=1}^{T-1} w_{\hat{u}}(\lambda_k)|^2}.$$

Step 3: Draw a random sample $\eta_1^*, \ldots, \eta_{[T/2]}^*$ from the distribution $P^*(\eta_j^* = \tilde{w}_{\hat{u}}(\lambda_k)) = 1/[T/2]$ for $k = 1, \ldots, [T/2]$ and generate a bootstrap sample

$$w_y^*(\lambda_j) = \hat{\beta}_0' w_x(\lambda_j) + |w_{\hat{u}}(\lambda_j)|\eta_j^*, \quad j = 1, \ldots, [T/2],$$

where $\hat{\beta}_0$ is the estimate of β from the null regression of $w_y(\lambda_j)$ on $w_x(\lambda_j)$ alone.

Step 4: Compute $(\hat{\beta}^*(\tau)', \hat{\delta}^*(\tau)')'$ as

$$\begin{pmatrix} \hat{\beta}^*(\tau) \\ \hat{\delta}^*(\tau) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} 2\,Re \begin{pmatrix} \sum_{j=1}^{[T/2]} I_{y^*x}(\lambda_j) \\ \sum_{j=1}^{[T/2]} I_{y^*z}(\lambda_j) \end{pmatrix},$$

where the right-hand side depends on τ through the definition $z_t = x_t \mathbb{1}(t \leqslant [\tau T])$ in (3).

Step 5: Compute the functional used for the original data, $Z_T^* = \phi(\sqrt{T}\hat{\delta}^*)$.

The distribution of the bootstrap test statistic Z_T^* can be used to approximate the asymptotic null distribution of Z_T, that is to construct a bootstrap test. To show the validity of the bootstrap procedure, we need to prove that the bootstrap process

$$\begin{pmatrix} \hat{\beta}^*(\tau) - \hat{\beta}_0 \\ \hat{\delta}^*(\tau) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \times 2\,Re \begin{pmatrix} \sum_{j=1}^{[T/2]} w_{x,j} |w_{\hat{u},j}| \eta_j^* \\ \sum_{j=1}^{[T/2]} w_{z(\tau),j} |w_{\hat{u},j}| \eta_j^* \end{pmatrix} \tag{13}$$

consistently estimates the null behaviour of the process $(\hat{\beta}(\tau)' - \beta', \hat{\delta}(\tau)')'$. It must be shown that under the null and under the local alternative the process $2\,Re\, T^{-1/2} \sum_{j=1}^{[T/2]} w_{z(\tau),j} |w_{\hat{u},j}| \eta_j^*$, conditionally on data, converges weakly in probability to the same process as $T^{-1/2} \sum_{j=1}^{T-1} I_{zu}(\lambda_j)$, that is to $(1/2\pi)\Omega^{1/2} B(\tau)$,

$$2\,Re \frac{1}{\sqrt{T}} \sum_{j=1}^{[T/2]} w_{z(\tau),j} |w_{\hat{u},j}| \eta_j^* \overset{p}{\Longrightarrow} \frac{1}{2\pi} \Omega^{1/2} B(\tau),$$

where "$\overset{p}{\Longrightarrow}$" stands for the weak convergence in probability as defined by Giné and Zinn (1990).

The consistency of the bootstrap is asserted in the following theorem.

Theorem 4. *Under Conditions* 1–5 *and under both the null and the local alternative hypotheses*,

$$\sqrt{T}\begin{pmatrix} \hat{\beta}^*(\tau) - \hat{\beta} \\ \hat{\delta}^*(\tau) \end{pmatrix} \Longrightarrow \frac{1}{\tau(1-\tau)} \begin{pmatrix} \Sigma^{-1}\Omega^{1/2}(\tau B(1) - \tau B(\tau)) \\ \Sigma^{-1}\Omega^{1/2}(B(\tau) - \tau B(1)) \end{pmatrix}$$

in probability.

A straightforward consequence of Theorem 4 and the continuous mapping theorem is the following corollary.

Corollary 2. *Let ϕ be a continuous functional on $D(\Lambda)^p$. Let*

$$Z_T^* = \phi(\sqrt{T}\hat{\delta}^*(\tau))$$

and let Z_0 be *Z_h of Corollary* 1 *with $h = 0$, i.e.*

$$Z_0 = \phi((\tau(1-\tau))^{-1}\Sigma^{-1}\Omega^{1/2}(B(\tau) - \tau B(1))).$$

Under the conditions of Theorem 4,

$$Z_T^* \xrightarrow{\text{d}} Z_0$$

in probability, *that is*

$$P(Z_T^* \leqslant x | \mathscr{F}_T \vee \mathscr{G}_T) \xrightarrow{\text{p}} P(Z_0 \leqslant x)$$

for each continuity point x of the right-hand side.

The bootstrap test is constructed using a critical region C_α^* based on the bootstrap distribution in such a way that $P(Z_T^* \in C_\alpha^*) = \alpha$, where α is a level of significance. The bootstrap test rejects when $Z_T \in C_\alpha^*$. Let $F_T^*(x) = P(Z_T^* \leqslant x | \mathscr{F}_T \vee \mathscr{G}_T)$ denote the distribution function of Z_T^* conditional on data and $F(x) = P(Z_0 \leqslant x)$ the null asymptotic distribution function. The bootstrap p-value for a one-tailed test is $p_T = 1 - F_T^*(Z_T)$. The bootstrap test rejects H_0 when Z_T is large, that is when p_T is small. By Corollaries 1 and 2, $Z_T \xrightarrow{\text{d}} Z_h$ and $F_T^* \Longrightarrow F$ in probability. The continuous mapping theorem implies that $p_T = 1 - F_T^*(Z_T) \xrightarrow{\text{d}} 1 - F(Z_h)$ in probability. The p-values based on the bootstrap distribution F_T^* are therefore asymptotically equivalent to the p-values based on the distribution F.

It should be noted that the proposed bootstrap is not the only possibility. The variables η_j^* in Step 3 are drawn from the empirical distribution of normalized discrete Fourier transform of the OLS residuals. Alternatively, external bootstrap can be carried out by drawing η_j^* from any complex distribution with zero mean, unit variance and $\mathrm{E}\eta_j^{*2} = 0$. A natural choice is a complex normal distribution. The proof of validity of the external bootstrap procedure remains identical to the current proof. Another valid modification is to multiply η_j^* in Step 3 by the value of $w_{\hat{u}}(\lambda_j)$ instead of its modulus. The proof of validity in this case goes through with only minor alterations as noted at the end of the proof of Proposition 7 in Section 6 below. A simulation study suggests that none of the methods above dominates the others in performance.

Hidalgo (2003) interchanges the resampling with the Fourier transformation, resampling first from the normalized time-domain residuals and then transforming the resampled data into the frequency domain. His simulation results seem to suggest that there is no substantial advantage in exchanging the order of the operations. In the simulation experiments in this paper we use the procedure given in Steps 1–5.

4. Monte Carlo

In order to assess the performance of the bootstrap procedure in finite samples, a small Monte Carlo study is conducted. Data are generated according to a simple linear model

$$y_t = \alpha + \beta_t x_t + u_t,$$

where scalar series $\{x_t\}$ and $\{u_t\}$ follow a FARIMA$(0, d, 0)$ process and where $\alpha = 0$. The long memory parameters d_x and d_u for the regressor x_t and errors u_t are either 0 (short memory) or 0.2 (stationary long memory). The series x_t and u_t are generated using the Davies–Harte (1987) algorithm. The set Λ of feasible break dates is taken to be the interval $[\varepsilon T, (1-\varepsilon)T]$ where $\varepsilon = 0.05$, so that approximately 5% of potential break dates are discarded from each side of the $1, \ldots, T$ range. The sample sizes considered are 32, 64, 128, 256. While a sample of length 32 may be too short to yield satisfactory results in the long memory case, the Monte Carlo simulation can still offer useful insights into the performance of the method for the short memory case. Two functionals are chosen on which to base the test procedure: a Kolmogorov–Smirnov- (or Bartlett-) type statistic, whose discrete version is

$$\mathrm{KS} = \sup_{[\varepsilon T] \leqslant j \leqslant [(1-\varepsilon)T]} \sqrt{T} \left| \hat{\delta}\left(\frac{j}{T}\right) \right|$$

and a Cramér–von Mises-type statistic based on L_2-distance, with a discrete version

$$\mathrm{CvM} = \sum_{j=[\varepsilon T]}^{[(1-\varepsilon)T]} \hat{\delta}^2\left(\frac{j}{T}\right).$$

The bootstrap test is based on the estimated process $\hat{\delta}$ obtained from (4) or (6). Since the limiting variance of the process $\hat{\delta}(\tau)$ varies with τ, we also consider a normalised version $[\tau(1-\tau)]^{1/2}\hat{\delta}(\tau)$, whose variance is level across Λ.

The asymptotic test is based on the process $\tilde{\delta}(\tau) = \hat{\Omega}^{-1/2}\hat{\Sigma}\hat{\delta}(\tau)$, where $\hat{\Sigma}$ and $\hat{\Omega}$ are computed as in (11) and (12), respectively. A levelled version $[\tau(1-\tau)]^{1/2}\tilde{\delta}(\tau)$ is also considered. The values of the Kolmogorov–Smirnov and Cramér–von Mises test statistics are compared with quantiles of their asymptotic distribution. These quantiles are estimated by approximating the limiting processes by their discrete versions over a grid of 10,000 points spaced equally across the interval $[0, 1]$ and by simulating the distribution of functionals of these processes by Monte Carlo. The number of Monte Carlo replications is 10^6.

The results in each of the tables are all obtained conditionally on a set of 5000 replications of a 256×2 matrix of independent identically distributed $N(0,1)$ elements. Within each replication, 1000 bootstrap samples are generated. The rejection probabilities are based on 5% nominal significance level.

For the examination of the level of the bootstrap and asymptotic tests, the results are given in Table 1. In this table and in Table 2, the heading "raw" denotes the size of the test based on the original process $\hat{\delta}(\tau)$ defined in (4) or (6) whereas the heading "norm" refers to the size of the test based on the levelled process $[\tau(1-\tau)]^{1/2}\hat{\delta}(\tau)$. The bootstrap test is nonconservative, with level approaching the nominal value from above as the sample size increases. Overall, neither KS nor CvM test statistic can be said to generate better test as far as level is concerned. The actual level tends to be closer to the nominal value when the memory of the error is of short range. Levelling the variance of the process $\hat{\delta}$ does not seem to bring substantial changes in the rejection probabilities under the null.

The asymptotic test performs poorly for the range of sample sizes under consideration. Again, neither of the Kolmogorov–Smirnov and Cramér–von

Table 1
Level of test at 5% nominal level

		Bootstrap test				Asymptotic test			
		KS		CvM		KS		CvM	
d_x	d_u	Raw	Norm	Raw	Norm	Raw	Norm	Raw	Norm
$T=32$									
0	0	9.9	9.9	9.4	9.3	46.7	41.5	52.3	34.6
0	0.2	12.3	12.2	11.9	10.5	48.8	43.1	54.6	36.1
0.2	0	9.9	10.4	10.2	9.4	49.7	44.6	56.9	41.0
0.2	0.2	12.2	12.6	12.3	11.0	50.5	45.5	58.8	42.9
$T=64$									
0	0	9.1	9.2	8.8	7.7	17.9	15.0	15.7	9.4
0	0.2	10.2	9.6	8.3	7.5	18.7	15.8	17.1	10.1
0.2	0	8.8	8.6	8.7	8.1	20.7	18.2	21.6	13.5
0.2	0.2	10.1	9.4	9.3	8.5	19.9	18.2	22.6	15.5
$T=128$									
0	0	6.5	6.3	6.7	6.5	7.6	4.6	6.4	4.7
0	0.2	6.9	6.4	6.7	6.5	8.2	5.0	6.8	4.5
0.2	0	6.4	6.5	6.9	6.7	9.5	6.2	8.5	6.1
0.2	0.2	7.4	7.3	7.2	7.1	8.7	6.4	9.6	7.4
$T=256$									
0	0	5.3	5.7	5.9	5.9	3.7	1.7	4.0	3.3
0	0.2	5.8	5.5	5.9	5.9	4.0	1.9	4.2	3.4
0.2	0	5.4	5.3	6.1	6.1	4.8	2.4	5.2	4.1
0.2	0.2	6.3	6.0	6.1	6.0	4.0	2.2	5.5	4.7

Table 2
Power against the alternative of one break at 5% nominal level

		Bootstrap test				Asymptotic test			
		KS		CvM		KS		CvM	
d_x	d_u	Raw	Norm	Raw	Norm	Raw	Norm	Raw	Norm
$T = 32$									
0	0	11.0	19.8	24.9	36.9	48.0	52.1	75.8	70.0
0	0.2	13.0	20.9	26.5	38.4	49.9	53.0	76.8	71.0
0.2	0	11.9	21.0	26.2	37.9	52.3	58.9	81.1	76.2
0.2	0.2	14.3	22.5	27.9	38.4	52.6	58.9	80.5	75.8
$T = 64$									
0	0	15.5	53.1	68.3	80.9	17.5	54.7	78.9	82.4
0	0.2	15.6	51.4	66.6	79.6	11.7	54.4	78.4	81.8
0.2	0	16.8	53.4	68.9	81.5	22.4	65.0	84.3	87.5
0.2	0.2	17.5	50.8	66.0	77.1	21.5	60.5	81.0	83.8
$T = 128$									
0	0	32.3	91.9	97.5	99.1	16.0	91.5	97.1	98.5
0	0.2	31.0	90.3	96.5	98.7	16.0	89.9	96.3	98.0
0.2	0	34.7	92.5	98.1	99.4	24.5	94.8	98.3	99.2
0.2	0.2	32.3	89.2	95.0	97.9	19.9	88.8	95.7	99.2
$T = 256$									
0	0	79.4	100.0	100.0	100.0	61.7	100.0	100.0	100.0
0	0.2	74.6	99.9	100.0	100.0	100.0	100.0	100.0	100.0
0.2	0	81.5	100.0	100.0	100.0	73.7	100.0	100.0	100.0
0.2	0.2	71.6	100.0	100.0	100.0	49.3	99.9	99.8	100.0

Mises tests dominates the other. Levelling the variance of the process $\tilde{\delta}$ actually seems to slightly damage the null rejection probabilities for a range of sample sizes.

In order to explore the power of the test, the alternative is set up as a break in the middle of the sample, $\tau_0 = 1/2$, with unit size of the jump, $\delta = 1$. In the experiment the alternative is fixed, that is the size of break does not change with the sample size. The outcome of the simulation of power is reported in Table 2. In terms of power, the CvM test appears to be strictly preferable to the KS test for both the bootstrap and the asymptotic test. This is in agreement with expectation of Ploberger and Krämer (1992) who suspected that L_2-norm CvM test might perform better than sup-norm KS test in case of the one-time structural break. The rejection probabilities of the asymptotic test are larger than those of the bootstrap test in a majority of parameter combinations. However, such a comparison is not informative since the actual critical values have not been corrected to yield 5% level of the tests. An important observation is that levelling the variance of the process $\hat{\delta}$ unambiguously and substantially improves the power of all forms of the test.

Overall, the outcome of the simulation exercise provides evidence that the bootstrap procedure proposed in the paper performs reasonably well already for samples of moderate size. The results of the exercise further seem to suggest that (a) the bootstrap test is preferable to the asymptotic test, at least for small to moderately sized samples, (b) Cramér–von Mises-type of test statistic is preferable to the Kolmogorov–Smirnov-type, at least for one-time change alternatives, and (c) levelling the variance of the test process $\hat{\delta}$ across Λ may be recommended at least for some forms of the alternative hypothesis.

5. Conclusion

The paper examines a test for parameter instability in a linear model where memory of both regressors and errors is allowed to be of a long range. The testing procedure is based on a process of OLS slope coefficient estimators. The choice of a continuous functional of this process for constructing the test statistic can reflect beliefs about the form of alternative and can improve the power of the test procedure.

A bootstrap procedure is proposed to approximate the distribution of the test statistic to the first order. The procedure is carried out in frequency domain and does not require choice of any tuning parameter such as block length in block bootstrap methods. A Monte Carlo study suggests that the bootstrap produces good results and is superior over the asymptotic test for moderate size samples.

There are several natural directions in which the current work can be extended. First, the condition that $\Omega < \infty$ could be relaxed to allow for greater degree of collective memory of regressors and errors. In this case, the OLS estimation procedure could be replaced by a GLS-type procedure. Second, partial structural change could be considered and gains in efficiency from allowing partial change evaluated. Third, a bootstrap procedure might be shown to approximate the distribution of the test statistics to an order higher than first.

Further, under the assumption that the alternative hypothesis holds and is of the one-time structural break form, the date of break could be estimated and, based on the distribution of the break date estimator, inference conducted.

6. Proofs

For notational simplicity, the process $\{x_t\}$ in Theorems 1–4 is taken to be scalar. Asymptotic results for vector processes can be obtained using Cramér–Wold device for stochastic processes as defined for example in Lemma A4 of Andrews (1993).

Validity of Theorems 1, 2 and 4 rests on the fact that under Conditions 1–5, as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{T-1} I_{zu}(\lambda_j) \Longrightarrow \frac{1}{2\pi} \Omega^{1/2} B(\tau), \tag{14}$$

$$\frac{1}{T} \sum_{j=1}^{T-1} w_{h_1 x,j} \overline{w}'_{h_2 x,j} \xrightarrow{\mathrm{p}} \Sigma \int_0^1 h_1(t) h_2(t)\, \mathrm{d}t \tag{15}$$

and

$$2\, Re \frac{1}{\sqrt{T}} \sum_{j=1}^{[T/2]} w_{z(\tau),j} |w_{\hat{u},j}| \eta_j^* \overset{\mathrm{p}}{\Longrightarrow} \frac{1}{2\pi} \Omega^{1/2} B(\tau) \tag{16}$$

over $[0, 1]$, where for any function h, $\{w_{hx,j}, j = 1, \ldots, T\}$ is the discrete Fourier transform of the sequence $\{h(t/T)x_t, t = 1, \ldots, T\}$ and where the random variables η_j^* are defined in Step 3 of the bootstrap procedure. In all three cases, the convergence is shown in two steps. First, convergence is proved for weighted innovation processes $\{\xi_t\}$ and $\{\varepsilon_t\}$. The result for the processes $\{x_t\}$, $\{u_t\}$ is then established by showing that the difference between the left-hand side of (14)–(16) and their weighted-innovation analogues converges to zero in probability uniformly over $[0, 1]$. Auxiliary results are given in Lemmas 1–5 and Propositions 1–7 establish convergence in (14)–(16). The validity of Theorems 1–4 is then argued employing Propositions 1–7.

Lemma 1. *Let h be a bounded variation function on* $[0, 1]$. *Let* $H(\lambda) = \sum_{t=1}^{T} h(t/T) \mathrm{e}^{\mathrm{i}t\lambda}$. *Then for some* $0 < C < \infty$ *independent of* T,

(a) $|H(\lambda)| \leqslant C/|\lambda|$ *for* $\lambda \in (0, \pi]$,
(b) $\int_0^{\lambda_j} |H(\lambda)|\, \mathrm{d}\lambda = \mathrm{O}(\log j)$ *uniformly over* $1 \leqslant j \leqslant [T/2]$.

Proof. (a) Letting $D_t(\lambda) = \sum_{k=1}^{t} \mathrm{e}^{\mathrm{i}k\lambda}$, noting that

$$|D_t(\lambda)| = \left| \mathrm{e}^{\mathrm{i}\lambda(1+t)/2} \frac{\sin \lambda t/2}{\sin \lambda/2} \right| \leqslant \frac{\pi}{|\lambda|}$$

for $0 < \lambda \leqslant \pi$, and using summation by parts, we have

$$\begin{aligned} |H(\lambda)| &\leqslant \sum_{t=1}^{T-1} |D_t(\lambda)| \left| h\left(\frac{t}{T}\right) - h\left(\frac{t+1}{T}\right) \right| + |D_T(\lambda)| |h(1)| \\ &\leqslant \frac{\pi}{|\lambda|} \left(\sum_{t=1}^{T-1} \left| h\left(\frac{t}{T}\right) - h\left(\frac{t+1}{T}\right) \right| + |h(1)| \right) \\ &\leqslant \frac{C}{|\lambda|} \end{aligned}$$

due to the boundedness of the total variation of the function h.

(b) $\int_0^{\lambda_j} |H(\lambda)|\, \mathrm{d}\lambda = \int_0^{1/T} |H(\lambda)|\, \mathrm{d}\lambda + \int_{1/T}^{\lambda_j} |H(\lambda)|\, \mathrm{d}\lambda \leqslant T \int_0^{1/T} \mathrm{d}\lambda + \int_{1/T}^{\lambda_j} \frac{C}{\lambda}\, \mathrm{d}\lambda = \mathrm{O}(\log j)$. $\square$

Lemma 2. *Let h be a bounded variation function on* $[0,1]$. *Let* $\{x_t\}$ *be a covariance stationary process satisfying Conditions* 1, 2 *and* 4. *Let* $H_T(\lambda)=\sum_{t=1}^{T}h(t/T)\mathrm{e}^{\mathrm{i}t\lambda}$ *and* $K_{h,T}(\lambda)=(1/2\pi T)|H_T(\lambda)|^2$. *Then*

$$\int_{-\pi}^{\pi}\left|\frac{A(\lambda)}{A_j}-1\right|^2 K_{h,T}(\lambda-\lambda_j)\,\mathrm{d}\lambda=\mathrm{O}\left(\frac{1}{j}\right)\quad\text{as } T\to\infty$$

uniformly over integers $1\leqslant j\leqslant[T/2]$.

Proof. The function A satisfies assumptions A1, A2′ of Robinson (1995b). Furthermore, the kernel H_T has the property

$$|H_T(\lambda)|\leqslant\frac{\pi}{|\lambda|},\quad 0<\lambda\leqslant\pi,\ T\geqslant 1$$

by Lemma 1. Therefore the lemma is valid by the arguments of Robinson (1995b) in the proof of his Lemma 3. □

Lemma 3. *Let* $\{x_{1t}\}$, $\{x_{2t}\}$ *be scalar covariance stationary processes satisfying Conditions* 1, 2 *and* 4. *Let* h_1, h_2 *be bounded variation functions on* $[0,1]$. *Denote by* A_k *the transfer functions of the processes* $\{x_{kt}\}$, $k=1,2$. *Let* $v_k(\lambda_j)=\sqrt{2\pi}w_{k,j}/(\sigma_\xi A_k(\lambda j))$ *where* $\{w_{k,j}, j=1,\ldots,T\}$ *is the discrete Fourier transform of the sequence* $\{h_k(t/T)x_{kt}, t=1,\ldots,T\}$. *Then*

(a) $\mathrm{E}\{v_k(\lambda_j)\bar{v}_l(\lambda_j)\}=T^{-1}\sum_{t=1}^{T}h_k(t/T)h_l(t/T)+\mathrm{O}(\log j/j)$ *and*
(b) $\mathrm{E}\{v_k(\lambda_j)v_l(\lambda_j)\}=\mathrm{O}(1)$

uniformly over integers $1\leqslant j\leqslant[T/2]$, *for* k, $l=1,2$.

Proof. (a) We have

$$\begin{aligned}&\mathrm{E}w_{k,j}\overline{w}_{l,j}-\frac{\sigma_\xi^2}{2\pi}\left(\frac{1}{T}\sum_{t=1}^{T}h_k\left(\frac{t}{T}\right)h_l\left(\frac{t}{T}\right)\right)A_k(\lambda_j)\overline{A_l}(\lambda_j)\\&\quad=\frac{\sigma_\xi^2}{2\pi}\int_{-\pi}^{\pi}(A_k(\lambda)\overline{A_l}(\lambda)-A_k(\lambda_j)\overline{A_l}(\lambda_j))K_{kl}(\lambda-\lambda_j)\,\mathrm{d}\lambda,\end{aligned}$$

where

$$K_{kl}(\lambda)=\frac{1}{2\pi T}\overline{H}_k(\lambda)H_l(\lambda)$$

and

$$H_k(\lambda)=\sum_{t=1}^{T}h_k\left(\frac{t}{T}\right)\mathrm{e}^{\mathrm{i}t\lambda},\quad k=1,2.$$

Condition 4 implies that we can choose $\eta>0$ such that for $\lambda\in(-\eta,0)\cup(0,\eta)$, for some d_k, $d_l\in[0,\frac{1}{2})$ and for some $0<C<\infty$, $|A_k(\lambda)\overline{A_l}(\lambda)|\leqslant C|\lambda|^{-(d_k+d_l)}$ and $|(\mathrm{d}/\mathrm{d}\lambda)A_k(\lambda)\overline{A_l}(\lambda)|\leqslant C|\lambda|^{-(d_k+d_l)-1}$. Furthermore by Lemma 1 the kernels K_{kl} and H_k display properties required in the proof of Theorem 2 of Robinson (1995a), namely $K_{kl}(\lambda)=\mathrm{O}(1/T\lambda^2)$ for $0<|\lambda|\leqslant\pi$ and $\int_{-D\lambda_j}^{D\lambda_j}|H_k(\lambda)|\,\mathrm{d}\lambda=\mathrm{O}(\log j)$, $k=1,2$.

The proof of part (a) therefore follows as in the first part of case (a) of Theorem 2 of Robinson (1995a). We obtain

$$\mathrm{E}w_{k,j}\overline{w}_{l,j} - \frac{\sigma_\xi^2}{2\pi}\left(\frac{1}{T}\sum_{t=1}^{T} h_k\left(\frac{t}{T}\right)h_l\left(\frac{t}{T}\right)\right)A_{k,j}\overline{A}_{l,j} = \mathrm{O}\left(\frac{\log j}{j}\lambda_j^{-(d_k+d_l)}\right)$$

from which it can be deduced that

$$\mathrm{E}v_{k,j}\overline{v}_{l,j} = \frac{1}{T}\sum_{t=1}^{T} h_k\left(\frac{t}{T}\right)h_l\left(\frac{t}{T}\right) + \mathrm{O}\left(\frac{\log j}{j}\right).$$

as required.

Part (b) follows from part (a) by the Schwarz inequality. □

Lemma 4. *Let g be a complex-valued function on* $[0,\pi]$ *which satisfies* (a) $|g|^2$ *is integrable on* $[0,\pi]$, (b) $g(\lambda) = \mathrm{O}(\lambda^{-d})$ *for* $\lambda \to 0+$ *for some* $d < \frac{1}{2}$ *and* (c) g *is bounded on any subinterval of* $(0,\pi]$. *Then for any* $\alpha \geqslant 0$, $\beta \geqslant 0$ *such that* $2d\alpha + \beta < 1$, as $T \to \infty$,

$$\frac{1}{T}\sum_{j=1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^{\beta}} \to \frac{1}{2\pi}\int_0^{\pi} \frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}}\,\mathrm{d}\lambda.$$

Proof. Fix $\varepsilon > 0$. For any small η,

$$\begin{aligned}&\left|\frac{1}{T}\sum_{j=1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^{\beta}} - \frac{1}{2\pi}\int_0^{\pi} \frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}}\,\mathrm{d}\lambda\right| \\ &\quad\leqslant \frac{1}{T}\sum_{j=1}^{[\eta T]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^{\beta}} + \frac{1}{2\pi}\int_0^{\eta} \frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}}\,\mathrm{d}\lambda \\ &\qquad + \left|\frac{1}{T}\sum_{j=[\eta T]+1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^{\beta}} - \frac{1}{2\pi}\int_{\eta}^{\pi} \frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}}\,\mathrm{d}\lambda\right|. \end{aligned} \tag{17}$$

By assumption, for small enough $\eta > 0$ and $0 < \lambda < \eta$,

$$\frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}} \leqslant C\lambda^{-2d\alpha-\beta} < C\lambda^{-1+\delta}$$

for some $\delta > 0$. Therefore

$$\frac{1}{T}\sum_{j=1}^{[\eta T]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^{\beta}} \leqslant \frac{C}{T}\sum_{j=1}^{[\eta T]} \left(\frac{j}{T}\right)^{-1+\delta} \leqslant C\eta^{\delta}.$$

Similarly

$$\frac{1}{2\pi}\int_0^{\eta} \frac{|g(\lambda)|^{2\alpha}}{\lambda^{\beta}}\,\mathrm{d}\lambda \leqslant C\eta^{\delta}.$$

The third term in (17) converges to zero as $T \to \infty$ by integrability of $|g|^2$. For small enough η and large enough T, the left-hand side of (17) is smaller than ε. $\square$

Lemma 5. *Let g be a function satisfying assumptions of the previous lemma. Then for any* $\alpha > 0, \beta \geqslant 0$, $\delta \geqslant 0$ *and* $\gamma \geqslant 1$, *as* $T \to \infty$,

(a) $(T^{-\alpha})\sum_{j=1}^{[T/2]} |g(\lambda_j)|^{2\alpha} \log^\delta j / j^\beta = \begin{cases} \mathrm{O}(\log^\delta T), & \alpha+\beta = 1, \\ \mathrm{o}(1), & \alpha+\beta > 1, \end{cases}$

(b) $(T^{-\alpha})\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]} g(\lambda_j)^\alpha \overline{g}(\lambda_k)^\alpha (\frac{\log^\delta j}{j^\beta}\frac{\log^\delta k}{k^\beta})^{1/2} \frac{1}{|j-k|_+^\gamma} = \mathrm{o}(1)$ *if* $\alpha+\beta > 1$, *where* $|j-k|_+ = \max\{1, |j-k|\}$.

Proof. (a) We have

$$\frac{1}{T^\alpha}\sum_{j=1}^{[T/2]} |g(\lambda_j)|^{2\alpha} \frac{\log^\delta j}{j^\beta} \leqslant C \frac{\log^\delta T}{T^{\alpha+\beta}} \sum_{j=1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^\beta}.$$

First, for a small positive η as in the previous lemma:

$$\begin{aligned}\frac{\log^\delta T}{T^{\alpha+\beta}} \sum_{j=1}^{[\eta T]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^\beta} &\leqslant C \frac{\log^\delta T}{T^{\alpha+\beta}} \sum_{j=1}^{[\eta T]} \lambda_j^{-2d\alpha-\beta} = C T^{\alpha(2d-1)} \log^\beta T \sum_{j=1}^{[\eta T]} j^{-2d\alpha-\beta} \\ &= \begin{cases} \mathrm{O}(T^{\alpha(2d-1)}\log^\delta T), & 2d\alpha > 1-\beta, \\ \mathrm{O}(T^{1-\alpha-\beta}\log^{\delta+1} T), & 2d\alpha = 1-\beta, \\ \mathrm{O}(T^{1-\alpha-\beta}\log^\delta T), & 2d\alpha < 1-\beta. \end{cases}\end{aligned}$$

Second,

$$\frac{1}{T}\sum_{j=[\eta T]+1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^\beta} \to \frac{1}{2\pi}\int_{2\pi\eta}^{\pi} \frac{|g(\lambda)|^{2\alpha}}{\lambda^\beta}\,\mathrm{d}\lambda$$

by integrability of $|g|^2$ and $1/\lambda$ over any interval $[2\pi\varepsilon, \pi]$, $\varepsilon \in (0, \frac{1}{2})$. Therefore

$$\frac{\log^\delta T}{T^{\alpha+\beta}} \sum_{j=[\eta T]+1}^{[T/2]} \frac{|g(\lambda_j)|^{2\alpha}}{\lambda_j^\beta} = \mathrm{O}\left(\frac{\log^\delta T}{T^{\alpha+\beta-1}}\right) = \begin{cases} \mathrm{O}(\log^\delta T), & \alpha+\beta = 1, \\ \mathrm{o}(1), & \alpha+\beta > 1 \end{cases}$$

and part (a) is established.

(b) By the Schwarz inequality, the sum in question is bounded by

$$\begin{aligned}\frac{1}{T^\alpha}\sum_{j=1}^{[T/2]} |g(\lambda_j)|^{2\alpha} \frac{\log^\delta j}{j^\beta} \sum_{k=1}^{[T/2]} \frac{1}{|j-k|_+^\gamma} &\leqslant \frac{C}{T^\alpha}\sum_{j=1}^{[T/2]} |g(\lambda_j)|^{2\alpha} \frac{\log^\delta j}{j^\beta}\sum_{k=1}^{[T/2]} \frac{1}{k} \\ &= C\frac{\log^\delta T}{T^{\alpha+\beta-1}} \mathrm{O}(\log T) = \mathrm{o}(1)\end{aligned}$$

from part (a). $\square$

Proposition 1. *Let g be a complex-valued function on* $[0, \pi]$ *which satisfies* (a) $g(-\lambda) = \overline{g}(\lambda)$ *for all* $\lambda \in (0, \pi]$, (b) $|g|^2$ *is integrable on* $[0, \pi]$, (c) $g(\lambda) = \mathrm{O}(\lambda^{-d})$ *for* $\lambda \to 0+$ *for* *some* $d < \frac{1}{2}$ *and* (d) g *is bounded on any subinterval of* $(0, \pi]$. *Under Conditions* 1–3, *as* $T \to \infty$,

$$\frac{2\pi}{\sqrt{T}} \sum_{j=1}^{T-1} g(\lambda_j) w_{\zeta(\tau)}(\lambda_j) \overline{w}_\varepsilon(\lambda_j) \Longrightarrow \left(\sigma_\xi^2 \sigma_\varepsilon^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\lambda)|^2 \, \mathrm{d}\lambda \right)^{1/2} B(\tau) \tag{18}$$

on $[0, 1]$, *where* B *is a standard Brownian motion and where the sequence* $\{\zeta_t(\tau)\}$ *is defined as* $\{\zeta_t(\tau)\} = \{\xi_t \mathbb{I}(t \leqslant [\tau T]),\ t = 1, \ldots, T\}$.

Proof. The left-hand side of (18) can be written as

$$G_T(\tau) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} \xi_t \left(\sum_{s=1}^{T} \varepsilon_s c_{t-s} \right),$$

where

$$c_t = \frac{1}{T} \sum_{j=1}^{T-1} g(\lambda_j) \mathrm{e}^{\mathrm{i} t \lambda_j}.$$

Denoting $d_t = (1/\sqrt{T}) \sum_{s=1}^{T} \varepsilon_s c_{t-s}$, the process G_T can be written as

$$G_T(\tau) = \sum_{t=1}^{[\tau T]} \xi_t d_t.$$

The realizations of the process G_T belong to the space $D[0, 1]$ of real functions which are right continuous with left-hand limits. The sequence $\{\xi_t d_t, \mathscr{F}_{t-1} \vee \mathscr{G}_T,\ 1 \leqslant t \leqslant T\}$ is a martingale difference sequence. The first two moments of the process G_T are

$$\mathrm{E} G_T(\tau) = 0,$$
$$\mathrm{E}|G_T(\tau)|^2 = \sigma_\varepsilon^2 \sigma_\xi^2 \tfrac{[\tau T]}{T} \tfrac{1}{T} \textstyle\sum_{j=1}^{T-1} |g(\lambda_j)|^2 = \tfrac{[\tau T]}{T} \mathrm{E}|G_T(1)|^2.$$

The variance of the process G_T therefore increases asymptotically linearly in τ and the weak convergence of the process G_T in (18) holds if the following two conditions of Scott (1973) are satisfied:

(a) $\sum_{t=1}^{T} \mathrm{E}(|d_t \xi_t|^2 | \mathscr{F}_{t-1} \vee \mathscr{G}_T) \xrightarrow{\mathrm{p}} \sigma_\xi^2 \sigma_\varepsilon^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\lambda)|^2 \, \mathrm{d}\lambda$ as $T \to \infty$ and
(b) $\sum_{t=1}^{T} \mathrm{E}(|d_t \xi_t|^2 \mathbb{I}(|d_t \xi_t| > \delta) | \mathscr{F}_{t-1} \vee \mathscr{G}_T) \xrightarrow{\mathrm{p}} 0$ as $T \to \infty$, for any positive δ.

These two conditions have been checked by Hidalgo (2003) under similar assumptions on the weight function g and identical assumptions on the processes $\{\xi_t\}$, $\{\varepsilon_t\}$. After making appropriate adjustments for complex weight functions and replacing Lemma 1 there with our Lemma 4, the proof remains valid in our case. □

Proposition 2. *Under Conditions* 1–5,

$$\begin{pmatrix} \frac{1}{\sqrt{T}} \sum_{j=1}^{T-1} I_{xu}(\lambda_j) \\ \frac{1}{\sqrt{T}} \sum_{j=1}^{T-1} I_{zu}(\lambda_j) \end{pmatrix} \Longrightarrow \begin{pmatrix} \frac{1}{2\pi} \Omega^{1/2} B(1) \\ \frac{1}{2\pi} \Omega^{1/2} B(\tau) \end{pmatrix}$$

over $[0, 1]$ *as* $T \to \infty$, *where* $\Omega = 2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) \, \mathrm{d}\lambda$.

Proof. It suffices to show that $1/\sqrt{T}\sum_{j=1}^{T-1} I_{zu}(\lambda_j) \Longrightarrow \frac{1}{2\pi}\Omega^{1/2}B(\tau)$ over $[0,1]$. The function

$$g(\lambda) = \frac{1}{2\pi}A(\lambda)\overline{B}(\lambda)$$

satisfies the conditions of Proposition 1. The present proposition is then proved if

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} A(\lambda_j)\overline{B}(\lambda_j)\left(\frac{w_{z(\tau)}(\lambda_j)\overline{w}_u(\lambda_j)}{A(\lambda_j)\overline{B}(\lambda_j)} - w_{\zeta(\tau)}(\lambda_j)\overline{w}_\varepsilon(\lambda_j)\right) \Longrightarrow 0. \tag{19}$$

Denoting $f_j = f(\lambda_j)$ for any function f, the left-hand side of (19) can be written as

$$Y_1(\tau) + Y_2(\tau) + Y_3(\tau),$$

where

$$Y_1(\tau) = \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} A_j\overline{B}_j\left(\frac{w_{z(\tau),j}}{A_j} - w_{\zeta(\tau),j}\right)\left(\frac{\overline{w}_{u,j}}{\overline{B}_j} - \overline{w}_{\varepsilon,j}\right),$$

$$Y_2(\tau) = \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} A_j\overline{B}_j\left(\frac{w_{z(\tau),j}}{A_j} - w_{\zeta(\tau),j}\right)\overline{w}_{\varepsilon,j}$$

and

$$Y_3(\tau) = \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} A_j\overline{B}_j w_{\zeta(\tau),j}\left(\frac{\overline{w}_{u,j}}{\overline{B}_j} - \overline{w}_{\varepsilon,j}\right). \tag{20}$$

Processes Y_1, Y_2 and Y_3 are of the form

$$Y_i(\tau) = \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} g_j V_j(\tau)\overline{W}_j, \quad i = 1,2,3,$$

where D is a generic constant and $V_j(\tau)$ and $\overline{W}_j$ stand for the third and the fourth factor, respectively, of the summands of the processes Y_i. To prove that $Y_i \Longrightarrow 0$ for $\tau \in [0,1]$ it suffices to show that finite dimensional distributions of the process Y_i converge to zero in probability and that the process Y_i is tight. Take any $n \in N$, any numbers $\tau_1, \ldots, \tau_n$ from the interval $[0,1]$ and any finite complex constants $\alpha_1, \ldots, \alpha_n$. The first moment of $\sum_{l=1}^{n}\alpha_l Y_i(\tau_l)$ is zero for $i = 1,2,3$. The second moment is

$$\frac{1}{T}\sum_{j=1}^{T-1}\sum_{k=1}^{T-1} g_j\overline{g}_k \mathrm{E}s_{jk}\mathrm{E}W_j\overline{W}_k \leqslant \frac{4}{T}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]} |g_j\overline{g}_k\|\mathrm{E}s_{jk}\|\mathrm{E}W_j\overline{W}_k|,$$

where

$$s_{jk} = \sum_{l=1}^{n}\sum_{m=1}^{n}\alpha_l\overline{\alpha}_m V_j(\tau_l)\overline{V}_k(\tau_m).$$

For $i = 1,2$ the factor $V_j(\tau)$ is equal to $w_{z(\tau),j}/A_j - w_{\zeta(\tau),j}$. The total variation of functions $h_\tau(x) = \mathbb{I}(0 \leqslant x \leqslant \tau)$, $\tau \in (0,1]$ is equal to one, therefore by Lemma 3 part (a)

$$\sup_{\tau\in[0,1]} \mathrm{E}\left|\frac{w_{z(\tau),j}}{A_j} - w_{\zeta(\tau),j}\right|^2 \leqslant \frac{D\log j}{j} \tag{21}$$

as $T \to \infty$ uniformly over integers $1 \leqslant j \leqslant [T/2]$. Using the Schwarz inequality,

$$|\mathrm{E}s_{jk}| \leqslant D\left(\frac{\log j}{j}\frac{\log k}{k}\right)^{1/2} \sum_{l=1}^{n}\sum_{m=1}^{n} |\alpha_l\|\alpha_m| \leqslant D\left(\frac{\log j}{j}\frac{\log k}{k}\right)^{1/2}.$$

When $i = 3$ the factor $V_j(\tau)$ equals $w_{\zeta(\tau),j}$. For any τ, $\sigma \in [0,1]$,

$$\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),k} = \frac{1}{2\pi T}\sum_{t=1}^{[\tau T]}\sum_{s=1}^{[\sigma T]} \mathrm{E}\xi_t\xi_s \mathrm{e}^{\mathrm{i}t\lambda_j - \mathrm{i}s\lambda_k} = \frac{\sigma_\xi^2}{2\pi}\frac{1}{T}\sum_{t=1}^{[(\tau\wedge\sigma)T]} \mathrm{e}^{\mathrm{i}t(\lambda_j-\lambda_k)}.$$

For $j = k$, the last expression is equal to $(\sigma_\xi^2/2\pi)[(\tau\wedge\sigma)T]/T$, while for $j \neq k$,

$$\begin{aligned}\frac{1}{T}\left|\sum_{t=1}^{[(\tau\wedge\sigma)T]} \mathrm{e}^{\mathrm{i}t(\lambda_j-\lambda_k)}\right| &= \frac{1}{T}\left|\frac{\sin([(\tau\wedge\sigma)T]\frac{\lambda_j-\lambda_k}{2})}{\sin(\frac{\lambda_j-\lambda_k}{2})}\right| \leqslant \frac{1}{T}\frac{1}{|\sin(\frac{\lambda_j-\lambda_k}{2})|} \\ &\leqslant \frac{1}{T}\frac{\pi}{|\lambda_j-\lambda_k|} = \frac{1}{2|j-k|}.\end{aligned}$$

In sum,

$$\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),k} = \mathrm{O}\left(\frac{1}{|j-k|_+}\right) \tag{22}$$

uniformly over $(\tau,\sigma) \in [0,1]^2$ and $1 \leqslant j,k \leqslant [T/2]$, where $|j-k|_+ = \max\{1, |j-k|\}$. Therefore when $i = 3$,

$$|\mathrm{E}s_{jk}| \leqslant D\sum_{l=1}^{n}\sum_{m=1}^{n} |\alpha_l\|\alpha_m| \frac{1}{|j-k|_+} \leqslant \frac{D}{|j-k|_+}.$$

Turning to the factor W_j, for $i = 1,3$ it is equal to $w_{u,j}/B_j - w_{\varepsilon,j}$. By Lemma 3 part (a),

$$\mathrm{E}\left|\frac{w_{u,j}}{B_j} - w_{\varepsilon,j}\right|^2 \leqslant \frac{D\log j}{j} \tag{23}$$

as $T \to \infty$ and by the Schwarz inequality

$$|\mathrm{E}W_j\overline{W}_k| \leqslant D\left(\frac{\log j}{j}\frac{\log k}{k}\right)^{1/2}.$$

In case $i = 2$, $W_j = w_{\varepsilon,j}$ and

$$\mathrm{E}\overline{w}_{\varepsilon,j}w_{\varepsilon,k} = \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T} \mathrm{E}\varepsilon_t\varepsilon_s e^{-\mathrm{i}t\lambda_j + \mathrm{i}s\lambda_k} = \frac{\sigma_\varepsilon^2}{2\pi T}\sum_{t=1}^{T} \mathrm{e}^{-\mathrm{i}t(\lambda_j-\lambda_k)} = \frac{\sigma_\varepsilon^2}{2\pi}\mathbb{I}(j = k).$$

Collecting the bounds obtained for moments of the factors $V_j(\tau)$ and W_j and using Lemma 5, the following results are obtained:

$$\mathrm{E}\left|\sum_{l=1}^{n}\alpha_l Y_1(\tau_l)\right|^2 \leqslant \frac{D}{T}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_j g_k|\frac{\log j}{j}\frac{\log k}{k} = \mathrm{o}(1),$$

$$\mathrm{E}\left|\sum_{l=1}^{n}\alpha_l Y_2(\tau_l)\right|^2 \leqslant \frac{D}{T}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_j g_k|\left(\frac{\log j}{j}\frac{\log k}{k}\right)^{1/2}\mathbb{I}(j=k)$$
$$= \frac{D}{T}\sum_{j=1}^{[T/2]}|g_j|^2\frac{\log j}{j} = \mathrm{o}(1),$$

$$\mathrm{E}\left|\sum_{l=1}^{n}\alpha_l Y_3(\tau_l)\right|^2 \leqslant \frac{D}{T}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_j g_k|\left(\frac{\log j}{j}\frac{\log k}{k}\right)^{1/2}\frac{1}{|j-k|_+} = \mathrm{o}(1).$$

An application of the Cramér–Wold device together with the Markov inequality establishes convergence of finite dimensional distributions of processes Y_i, $i=1,2,3$, to zero in probability.

Tightness of the processes Y_i is implied by the moment condition of Billingsley (1999, Theorem 13.5, p. 142):

$$\mathrm{E}|Y_i(\rho)-Y_i(\sigma)|^2|Y_i(\tau)-Y_i(\rho)|^2 \leqslant (F(\tau)-F(\sigma))^{2\alpha}, \quad i=1,2,3, \tag{24}$$

where $\alpha > \frac{1}{2}$, $\sigma \leqslant \rho \leqslant \tau$ and F is a nondecreasing, continuous function on $[0,1]$. The fourth moment of the difference $Y_i(\tau)-Y_i(\sigma)$ is given by

$$\mathrm{E}|Y_i(\tau)-Y_i(\sigma)|^4$$
$$\leqslant \frac{16}{T^2}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}\sum_{l=1}^{[T/2]}\sum_{m=1}^{[T/2]}|g_j\bar{g}_k g_l \bar{g}_m|\,|\mathrm{E}V_j\overline{V}_k V_l \overline{V}_m|\,|\mathrm{E}W_j\overline{W}_k W_l\overline{W}_m|,$$

where $V_j = V_j(\tau)-V_j(\sigma)$. For $i=1,2$, $V_j = (w_{z(\tau),j}-w_{z(\sigma),j})/A_j - (w_{\zeta(\tau),j}-w_{\zeta(\sigma),j})$ and

$$\mathrm{cum}(V_j,\overline{V}_k,V_l,\overline{V}_m)$$
$$= \frac{\kappa_\xi}{(2\pi)^5}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\mu)}{\overline{A}_k}-1\right)\left(\frac{A(\zeta)}{A_l}-1\right)\left(\frac{A(-\lambda-\mu-\zeta)}{\overline{A}_m}-1\right)$$
$$\times H(\lambda+\lambda_j)H(\mu-\lambda_k)H(\zeta+\lambda_l)H(-\lambda-\mu-\zeta-\lambda_m)\,\mathrm{d}\lambda\,\mathrm{d}\mu\,\mathrm{d}\zeta,$$

where $\kappa_\xi = \mathrm{cum}(\xi_t,\xi_t,\xi_t,\xi_t)$, $H(\lambda)=\sum_{t=1}^{T}h(t/T)\mathrm{e}^{\mathrm{i}t\lambda}$ and $h(x)=\mathbb{I}(\sigma \leqslant x \leqslant \tau)$. Proceeding as in the proof of (4.8) in Robinson (1995b), we get

$$|\mathrm{cum}(V_j,\overline{V}_k,V_l,\overline{V}_m)| \leqslant DP_j^{1/2}P_k^{1/2}P_l^{1/2}P_m^{1/2},$$

where

$$P_j = \int_{-\pi}^{\pi} \left| \frac{A(\lambda)}{A_j} - 1 \right|^2 \frac{1}{2\pi T} |H(\lambda + \lambda_j)|^2 \, \mathrm{d}\lambda.$$

Denoting $K_{h,T}(\lambda) = (2\pi T)^{-1} |H(\lambda)|^2$, it can be seen that

$$\begin{aligned} P_j &= (\tau - \sigma) \int_{-\pi}^{\pi} \left| \frac{A(\lambda)}{A_j} - 1 \right|^2 K_{1,(\tau-\sigma)T}(\lambda + \lambda_j) \, \mathrm{d}\lambda \\ &= (\tau - \sigma) \mathrm{O}\left(\frac{1}{j} \right) \end{aligned}$$

uniformly over $(\tau, \sigma) \in [0,1]^2$ and $1 \leqslant j \leqslant [T/2]$ by Lemma 2.

Likewise

$$\begin{aligned} \mathrm{E}|V_j|^2 &= \frac{\sigma_\xi^2}{2\pi} \int_{-\pi}^{\pi} \left| \frac{A(\lambda)}{A_j} - 1 \right|^2 \frac{1}{2\pi T} |H(\lambda + \lambda_j)|^2 \, \mathrm{d}\lambda \\ &= DP_j = (\tau - \sigma) \mathrm{O}\left(\frac{1}{j} \right). \end{aligned}$$

By the Schwarz inequality,

$$\begin{aligned} &|\mathrm{E} V_j \overline{V}_k V_l \overline{V}_m| \\ &\quad \leqslant |\mathrm{cum}(V_j, \overline{V}_k, V_l, \overline{V}_m)| + 3(\mathrm{E}|V_j|^2 \mathrm{E}|V_k|^2 \mathrm{E}|V_l|^2 \mathrm{E}|V_m|^2)^{1/2} \\ &\quad \leqslant C(\tau - \sigma)^2 j^{-1/2} k^{-1/2} l^{-1/2} m^{-1/2}. \end{aligned} \tag{25}$$

For $i = 3$, $V_j = w_{\xi(\tau)-\xi(\sigma),j}$ and

$$\begin{aligned} \mathrm{cum}(V_j, \overline{V}_k, V_l, \overline{V}_m) = \frac{\kappa_\xi}{(2\pi)^5} \frac{1}{T^2} \int_{-\pi}^{\pi} \int \int & H(\lambda + \lambda_j) \\ & \times H(\mu - \lambda_k) H(\zeta + \lambda_j) H(-\lambda - \mu - \zeta - \lambda_k) \, \mathrm{d}\lambda \, \mathrm{d}\mu \, \mathrm{d}\zeta \end{aligned}$$

which by using periodicity of H and the Schwarz inequality can be shown to be $(\tau - \sigma)^2 \mathrm{O}(1)$ uniformly over $(\tau, \sigma) \in [0,1]^2$ and $1 \leqslant j, k, l, m \leqslant [T/2]$. Similarly,

$$\mathrm{E}|V_j|^2 = \frac{\sigma_\xi^2}{4\pi^2 T} \int_{-\pi}^{\pi} |H(\lambda + \lambda_j)|^2 \, \mathrm{d}\lambda \leqslant C(\tau - \sigma)$$

and so for $i = 3$, $|\mathrm{E} V_j \overline{V}_k V_l \overline{V}_m| = (\tau - \sigma)^2 \mathrm{O}(1)$.

Regarding the factor W_j, for $i = 1, 3$ we have $W_j = w_{u,j}/B_j - w_{\varepsilon,j}$ and reasoning as in case of V_j $(i = 1, 2)$ we obtain

$$|\mathrm{cum}(W_j, \overline{W}_k, W_l, \overline{W}_m)| \leqslant D P_{B,j}^{1/2} P_{B,k}^{1/2} P_{B,l}^{1/2} P_{B,m}^{1/2}$$

and

$$\mathrm{E}|W_j|^2 = DP_{B,j},$$

where

$$P_{B,j} = \int_{-\pi}^{\pi} \left| \frac{B(\lambda)}{B_j} - 1 \right|^2 \frac{1}{2\pi T} |H(\lambda + \lambda_j)|^2 \mathrm{d}\lambda$$

with $H(\lambda) = \sum_{t=1}^{T} h(t/T) \mathrm{e}^{\mathrm{i}t\lambda}$ and $h \equiv 1$. By Lemma 2, $P_{B,j} = \mathrm{O}(j^{-1})$, therefore

$$\mathrm{E} W_j \overline{W}_k W_l \overline{W}_m = \mathrm{O}(j^{-1/2} k^{-1/2} l^{-1/2} m^{-1/2})$$

uniformly over $1 \leqslant j, k, l, m \leqslant [T/2]$.

Finally, when $i = 2$, $W_j = w_{\varepsilon,j}$,

$$\mathrm{cum}(W_j, \overline{W}_k, W_l, \overline{W}_m) = \frac{\kappa_\xi}{4\pi^2} \frac{1}{T^2} \sum_{t=1}^{T} \mathrm{e}^{\mathrm{i}t(\lambda_j - \lambda_k + \lambda_l - \lambda_m)} = \mathrm{O}\left(\frac{1}{T}\right)$$

and

$$\mathrm{E} W_j \overline{W}_k = \frac{1}{2\pi} \mathbb{I}(j = k) = \mathrm{O}(1)$$

uniformly over $1 \leqslant j, k, l, m \leqslant [T/2]$.

Due to the bounds obtained above for moments of V_j and W_j, the following inequalities hold:

$$\begin{aligned} \mathrm{E}|Y_1(\tau) - Y_1(\sigma)|^4 &\leqslant \frac{D}{T^2} \sum_{j,k,l,m=1}^{[T/2]} |g_j g_k g_l g_m| (\tau - \sigma)^2 j^{-1} k^{-1} l^{-1} m^{-1} \\ &= D(\tau - \sigma)^2 \left(T^{-1/2} \sum_{j=1}^{[T/2]} \frac{|g_j|}{j} \right)^4 = D(\tau - \sigma)^2 \mathrm{o}(1) \end{aligned}$$

uniformly over $(\tau, \sigma) \in [0, 1]^2$ by Lemma 4,

$$\begin{aligned} \mathrm{E}|Y_2(\tau) - Y_2(\sigma)|^4 &\leqslant \frac{D}{T^2} \sum_{j,k,l,m=1}^{[T/2]} |g_j g_k g_l g_m| (\tau - \sigma)^2 j^{-1/2} k^{-1/2} l^{-1/2} m^{-1/2} \\ &= D(\tau - \sigma)^2 \left(T^{-1/2} \sum_{j=1}^{[T/2]} \frac{|g_j|}{j^{1/2}} \right)^4 = (\tau - \sigma)^2 \mathrm{O}(1) \end{aligned}$$

by the Schwarz inequality and Lemma 4. The same bound applies to $\mathrm{E}|Y_3(\tau) - Y_3(\sigma)|^4$. By the Schwarz inequality,

$$\mathrm{E}|Y_i(\rho) - Y_i(\sigma)|^2 |Y_i(\tau) - Y_i(\rho)|^2 \leqslant D((\rho - \sigma)^2 (\tau - \sigma)^2)^{1/2} \leqslant D(\tau - \sigma)^2$$

for $i = 1, 2, 3$ and the moment condition (24) is verified with $\alpha = 2$ and $F(\tau) = D\tau^2$. This proves the uniform convergence in (19) and the proposition is established. □

Proposition 3. *Let g be a function satisfying the assumptions of Proposition* 1. *Let h_1, h_2 be bounded variation functions on* $[0, 1]$. *Let $\{w_{h\xi,j},\ j = 1, \ldots, T\}$ be the discrete Fourier transform of the sequence $\{h(t/T)\xi_t, t = 1, \ldots, T\}$. Under Conditions* 1 *and* 2, *as* $T \to \infty$,

$$\frac{2\pi}{T}\sum_{j=1}^{T-1}|g(\lambda_j)|^2 w_{h_1\xi,j}\overline{w}_{h_2\xi,j} \xrightarrow{\mathrm{p}} \frac{\sigma_\xi^2}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t.$$

Proof. Denote $g_j = g(\lambda_j)$, $h_{kt} = h_k(t/T)$ and

$$Z = \frac{2\pi}{T}\sum_{j=1}^{T-1}|g_j|^2 w_{h_1\xi,j}\overline{w}_{h_2\xi,j}.$$

Then

$$\begin{aligned}
\mathrm{E}Z &= \frac{2\pi}{T}\sum_{j=1}^{T-1}|g_j|^2\frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{E}\xi_t\xi_s h_{1t}h_{2s}\mathrm{e}^{\mathrm{i}(t-s)\lambda_j}\\
&= \frac{\sigma_\xi^2}{T^2}\sum_{j=1}^{T-1}|g_j|^2\sum_{t=1}^{T}h_{1t}h_{2t} \to \frac{\sigma_\xi^2}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t
\end{aligned}$$

by Lemma 4. Further,

$$\begin{aligned}
\mathrm{E}|Z|^2 &= \frac{1}{T^4}\sum_{j,k=1}^{T-1}|g_jg_k|^2\sum_{t,s,r,v=1}^{T}E(\xi_t\xi_s\xi_r\xi_v)h_{1t}h_{2s}h_{1r}h_{2v}\mathrm{e}^{\mathrm{i}(t-s)\lambda_j}\mathrm{e}^{-\mathrm{i}(r-v)\lambda_k}\\
&= \frac{\kappa_\xi}{T^4}\sum_{j=1}^{T-1}\sum_{k=1}^{T-1}|g_jg_k|^2\sum_{t=1}^{T}h_{1t}^2h_{2t}^2 + \frac{\sigma_\xi^4}{T^4}\sum_{j=1}^{T-1}\sum_{k=1}^{T-1}|g_jg_k|^2\left(\sum_{t=1}^{T}h_{1t}h_{2t}\right)^2\\
&\quad + \frac{\sigma_\xi^4}{T^4}\sum_{j=1}^{T-1}\sum_{k=1}^{T-1}|g_jg_k|^2\sum_{t=1}^{T}h_{1t}^2\mathrm{e}^{\mathrm{i}t(\lambda_j-\lambda_k)}\sum_{t=1}^{T}h_{2t}^2\mathrm{e}^{-\mathrm{i}t(\lambda_j-\lambda_k)}\\
&\quad + \frac{\sigma_\xi^4}{T^4}\sum_{j=1}^{T-1}\sum_{k=1}^{T-1}|g_jg_k|^2\left|\sum_{t=1}^{T}h_{1t}h_{2t}\mathrm{e}^{\mathrm{i}t(\lambda_j+\lambda_k)}\right|^2.
\end{aligned}$$

The first term is $\mathrm{O}(1/T)$ by Lemma 4. Proceeding as in the computations leading to (22), it can be seen that

$$\sum_{t=1}^{T}h_{lt}^2\mathrm{e}^{\mathrm{i}t(\lambda_j-\lambda_k)} \leqslant \frac{CT}{|j-k|_+}, \quad l=1,2,$$

where $|j-k|_+ = \max\{1,|j-k|\}$. Therefore the third term is bounded in absolute value by

$$\frac{D}{T^2}\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_jg_k|^2\frac{1}{|j-k|_+^2}$$

which is o(1) by Lemma 5.

Similarly, the fourth term is o(1). The second term is dominant and converges to

$$\left(\frac{\sigma_\xi^2}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t\right)^2$$

by Lemma 4. In sum,

$$\mathrm{E}Z \to \frac{\sigma_\xi^2}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t$$

and $\mathrm{E}|Z|^2 \to |\mathrm{E}Z|^2$. An application of the Markov inequality completes the proof of convergence of Z in probability to $(\sigma_\xi^2/2\pi)\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t$. □

Proposition 4. *Let h_1, h_2 be bounded variation functions on* $[0,1]$. *Let* $\{w_{h\xi,j}, j = 1,\dots,T\}$ *be the discrete Fourier transform of the sequence* $\{h(t/T)\xi_t,\dots,t=1,\dots,T\}$. *Under Conditions* 1–5, *as* $T\to\infty$,

$$\frac{1}{T}\sum_{j=1}^{T-1} w_{h_1x,j}\overline{w}_{h_2x,j} \xrightarrow{\mathrm{p}} \frac{1}{2\pi}\Sigma\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t$$

where $\Sigma = \mathrm{E}x_t^2$.

Proof. The function $g(\lambda) = A(\lambda)/\sqrt{2\pi}$ satisfies the conditions of Proposition 1. It is sufficient to prove that

$$\frac{1}{T}\sum_{j=1}^{T-1}|A_j|^2\left(\frac{w_{h_1x,j}\overline{w}_{h_2x,j}}{|A_j|^2} - w_{h_1\xi,j}\overline{w}_{h_2\xi,j}\right) \xrightarrow{\mathrm{p}} 0, \tag{26}$$

where $A_j = A(\lambda_j)$. The left-hand side is equal to

$$\begin{aligned}
&\frac{1}{T}\sum_{j=1}^{T-1}|A_j|^2\left(\frac{w_{h_1x,j}}{A_j} - w_{h_1\xi,j}\right)\left(\frac{\overline{w}_{h_2x,j}}{\overline{A}_j} - \overline{w}_{h_2\xi,j}\right)\\
&\quad + \frac{1}{T}\sum_{j=1}^{T-1}|A_j|^2\left(\frac{w_{h_1x,j}}{A_j} - w_{h_1\xi,j}\right)\overline{w}_{h_2\xi,j}\\
&\quad + \frac{1}{T}\sum_{j=1}^{T-1}|A_j|^2 w_{h_1\xi,j}\left(\frac{\overline{w}_{h_2x,j}}{\overline{A}_j} - \overline{w}_{h_2\xi,j}\right).
\end{aligned}$$

By the Schwarz inequality, the expectation of the modulus of the first term is bounded by

$$\begin{aligned}
&\frac{2}{T}\sum_{j=1}^{[T/2]}|A_j|^2\left(\mathrm{E}\left|\frac{w_{h_1x,j}}{A_j} - w_{h_1\xi,j}\right|^2\right)^{1/2}\left(\mathrm{E}\left|\frac{w_{h_2x,j}}{A_j} - w_{h_2\xi,j}\right|^2\right)^{1/2}\\
&\quad \leqslant \frac{D}{T}\sum_{j=1}^{[T/2]}|A_j|^2\frac{\log j}{j} = \mathrm{o}(1)
\end{aligned}$$

by Lemma 3 part (a) and Lemma 5. A bound for the expectation of the absolute value of the second term is

$$\frac{2}{T}\sum_{j=1}^{[T/2]}|A_j|^2\left(\mathrm{E}\left|\frac{w_{h_1x,j}}{A_j}-w_{h_1\xi,j}\right|^2\right)^{1/2}(\mathrm{E}|w_{h_2\xi,j}|^2)^{1/2}$$
$$\leqslant\frac{D}{T}\sum_{j=1}^{[T/2]}|A_j|^2\left(\frac{\log j}{j}\right)^{1/2}=\mathrm{o}(1)$$

by Lemma 3 part (a) and Lemma 5. The third term can be bounded in the same way as the second term. Therefore (26) holds and by Proposition 3

$$\frac{1}{T}\sum_{j=1}^{T-1}w_{h_1x,j}\overline{w}_{h_2x,j}\xrightarrow{\mathrm{p}}\frac{\sigma_\xi^2}{2\pi}\int_0^\pi\frac{1}{2\pi}|A(\lambda)|^2\,\mathrm{d}\lambda\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t$$
$$=\frac{1}{2\pi}\Sigma\int_0^1 h_1(t)h_2(t)\,\mathrm{d}t.\qquad\square$$

Proposition 5. *Under Conditions* 1–5, *with a function g satisfying the conditions of Proposition* 1,

$$\frac{1}{T}\sum_{j=1}^{T-1}|g(\lambda_j)|^2 w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}|w_{\varepsilon,j}|^2\Longrightarrow(\tau\wedge\sigma)\frac{\sigma_\xi^2\sigma_\varepsilon^2}{4\pi^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda$$

uniformly over $(\tau,\sigma)\in[0,1]^2$.

Proof. Denote $g(\lambda_j)=g_j$. First moment of the expression on the left-hand side is

$$\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}\mathrm{E}|w_{\varepsilon,j}|^2=\frac{[(\tau\wedge\sigma)T]}{T}\frac{\sigma_\xi^2\sigma_\varepsilon^2}{4\pi^2}\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2$$
$$\to(\tau\wedge\sigma)\frac{\sigma_\xi^2\sigma_\varepsilon^2}{4\pi^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda$$

by Lemma 4 because

$$\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}=\frac{1}{2\pi T}\sum_{t=1}^{[\tau T]}\sum_{s=1}^{[\sigma T]}\mathrm{E}\xi_t\xi_s\mathrm{e}^{\mathrm{i}(t-s)\lambda_j}=\frac{[(\tau\wedge\sigma)T]}{T}\frac{\sigma_\xi^2}{2\pi}.\tag{27}$$

Second moment of that expression is

$$\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}\bar{w}_{\zeta(\tau),k}w_{\zeta(\sigma),k}\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2$$
$$=\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{cum}(w_{\zeta(\tau),j},\bar{w}_{\zeta(\sigma),j},\bar{w}_{\zeta(\tau),k},w_{\zeta(\sigma),k})\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2$$

$$+\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}\mathrm{E}\bar{w}_{\zeta(\tau),k}w_{\zeta(\sigma),k}\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2$$
$$+\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\tau),k}\mathrm{E}\bar{w}_{\zeta(\sigma),j}w_{\zeta(\sigma),k}\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2$$
$$+\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}w_{\zeta(\tau),j}w_{\zeta(\sigma),k}\mathrm{E}\bar{w}_{\zeta(\tau),k}\bar{w}_{\zeta(\sigma),j}\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2. \tag{28}$$

Now

$$\mathrm{cum}(w_{\zeta(\tau),j},\bar{w}_{\zeta(\sigma),j},\bar{w}_{\zeta(\tau),k},w_{\zeta(\sigma),k})$$
$$=\frac{1}{4\pi^2T^2}\sum_{t=1}^{[(\tau\wedge\sigma)T]}\mathrm{cum}(\xi_t,\xi_t,\xi_t,\xi_t)=\frac{\kappa_\xi}{4\pi^2}\frac{1}{T}\frac{[(\tau\wedge\sigma)T]}{T}=\mathrm{O}\left(\frac{1}{T}\right)$$

uniformly over $(\tau,\sigma)\in[0,1]^2$. The fourth moments of ε_t are finite, therefore the first term of (28) is bounded by $(D/T^3)\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_jg_k|^2$ which is $\mathrm{O}(1/T)$ by Lemma 4. Further, from (22),

$$|\mathrm{E}w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),k}|\leqslant\frac{C}{|j-k|_+}$$

uniformly over $(\tau,\sigma)\in[0,1]^2$ and $1\leqslant j,k\leqslant[T/2]$, where $|j-k|_+=\max\{1,|j-k|\}$, and the third term of (28) is bounded by $(D/T^2)\sum_{j=1}^{[T/2]}\sum_{k=1}^{[T/2]}|g_jg_k|^2 1/|j-k|_+^2$ which is o(1) by Lemma 5. Similarly, the fourth term is o(1). Therefore we are left with the dominant second term:

$$\mathrm{E}\left|\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}|w_{\varepsilon,j}|^2\right|^2$$
$$=\left(\frac{[(\tau\wedge\sigma)T]}{T}\right)^2\frac{\sigma_\xi^4}{4\pi^2}\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2+\mathrm{o}(1).$$

Since

$$\mathrm{cum}(w_{\varepsilon,j},\bar{w}_{\varepsilon,j},w_{\varepsilon,k},\bar{w}_{\varepsilon,k})=\frac{1}{4\pi^2T^2}\sum_{t=1}^{T}\mathrm{cum}(\varepsilon_t,\varepsilon_t,\varepsilon_t,\varepsilon_t)=\frac{\kappa}{4\pi^2T}$$

and $\mathrm{E}w_{\varepsilon,j}\bar{w}_{\varepsilon,k}=(\sigma_\varepsilon^2/2\pi)\mathbb{I}(j=k)$, we have

$$\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2\mathrm{E}|w_{\varepsilon,j}|^2|w_{\varepsilon,k}|^2$$
$$=\frac{\kappa}{4\pi^2T}\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2+\frac{\sigma_\varepsilon^4}{4\pi^2}\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_jg_k|^2$$

$$+\frac{\sigma_\varepsilon^4}{4\pi^2}\frac{1}{T^2}|g_{T/2}|^4\mathbb{I}(T \text{ even}) + \frac{\sigma_\varepsilon^4}{4\pi^2}\frac{1}{T^2}\sum_{j=1}^{T-1}|g_j|^4$$

$$= \frac{\sigma_\varepsilon^4}{4\pi^2}\frac{1}{T^2}\sum_{j=1}^{T-1}\sum_{j=1}^{T-1}|g_j g_k|^2 + \mathrm{o}(1)$$

by Lemmas 4 and 5. That means that

$$\mathrm{E}\left|\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2 w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}|w_{\varepsilon,j}|^2\right|^2 \to \left((\tau \wedge \sigma)\frac{\sigma_\varepsilon^2\sigma_\xi^2}{4\pi^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda\right)^2.$$

The second moment of the process $(1/T)\sum_{j=1}^{T-1}|g_j|^2 w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}|w_{\varepsilon,j}|^2$ therefore converges to the square of the limit of its first moment. By the Markov inequality,

$$\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2 w_{\zeta(\tau),j}\bar{w}_{\zeta(\sigma),j}|w_{\varepsilon,j}|^2 \xrightarrow{\mathrm{p}} (\tau \wedge \sigma)\frac{\sigma_\varepsilon^2\sigma_\xi^2}{4\pi^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda$$

for each $(\tau, \sigma) \in [0, 1]^2$. Since the limiting function is continuous and increasing in τ and σ, the convergence is uniform. □

Proposition 6. *Let* $h_1, \ldots, h_4$ *be bounded variation functions on* $[0, 1]$. Let $\{x_t\}$ *be a covariance stationary process satisfying Conditions* 1, 2 *and* 4. *Let* $\{w_{h_r x,j}, j = 1, \ldots, T\}$ *be the discrete Fourier transform of the scalar sequence* $\{h_r(t/T)x_t, t = 1, \ldots, T\}$, $r = 1, \ldots, 4$. *Let* $I_{h_r x, h_s x, j} = w_{h_r x,j}\overline{w}_{h_s x,j}$. *Then*

$$\frac{1}{T}\sum_{j=1}^{T-1} I_{h_1 x, h_2 x, j} I_{h_3 x, h_4 x, j} = \mathrm{o_p}(T).$$

Proof. We have

$$\frac{1}{T}\sum_{j=1}^{T-1} I_{h_1 x, h_2 x, j} I_{h_3 x, h_4 x, j} = \frac{1}{T}\sum_{j=1}^{T-1} f_{xx,j}^2 (a_j + b_j + c_j + d_j), \tag{29}$$

where

$$a_j = \left(\frac{I_{h_1 x, h_2 x, j}}{f_{xx,j}} - 2\pi\frac{I_{h_1\xi, h_2\xi, j}}{\sigma_\xi^2}\right)\left(\frac{I_{h_3 x, h_4 x, j}}{f_{xx,j}} - 2\pi\frac{I_{h_3\xi, h_4\xi, j}}{\sigma_\xi^2}\right),$$

$$b_j = \left(\frac{I_{h_1 x, h_2 x, j}}{f_{xx,j}} - 2\pi\frac{I_{h_1\xi, h_2\xi, j}}{\sigma_\xi^2}\right)2\pi\frac{I_{h_3\xi, h_4\xi, j}}{\sigma_\xi^2},$$

$$c_j = 2\pi\frac{I_{h_1\xi, h_2\xi, j}}{\sigma_\xi^2}\left(\frac{I_{h_3 x, h_4 x, j}}{f_{xx,j}} - 2\pi\frac{I_{h_3\xi, h_4\xi, j}}{\sigma_\xi^2}\right)$$

and

$$d_j = \frac{4\pi^2}{\sigma_\xi^4} I_{h_1\xi,h_2\xi,j} I_{h_3\xi,h_4\xi,j}.$$

The second moment of the first factor of a_j is

$$\mathrm{E}|u_{1j}\bar{u}_{2j} - v_{1j}\bar{v}_{2j}|^2 = a_{1j} + a_{2j},$$

where

$$u_{r,j} = \frac{\sqrt{2\pi}}{\sigma_\xi}\frac{w_{h_r x,j}}{A_j}, \quad v_{r,j} = \frac{\sqrt{2\pi}}{\sigma_\xi} w_{h_r\xi,j},$$

$$\begin{aligned} a_{1j} = {} & \mathrm{cum}(u_{1j}, \bar{u}_{2j}, \bar{u}_{1j}, u_{2j}) - \mathrm{cum}(u_{1j}, \bar{u}_{2j}, \bar{v}_{1j}, v_{2j}) \\ & - \mathrm{cum}(v_{1j}, \bar{v}_{2j}, \bar{u}_{1j}, u_{2j}) + \mathrm{cum}(v_{1j}, \bar{v}_{2j}, \bar{v}_{1j}, v_{2j}) \end{aligned}$$

and, denoting $h_{rs} = \frac{1}{T}\sum_{t=1}^{T} h_r(t/T)h_s(t/T)$ for $r,s = 1,2$,

$$\begin{aligned} a_{2j} = {} & (\mathrm{E}u_{1j}\bar{u}_{2j} - h_{12})(\mathrm{E}\bar{u}_{1j}u_{2j} - h_{12}) + (\mathrm{E}u_{1j}\bar{u}_{2j} - h_{12}) + (\mathrm{E}\bar{u}_{1j}u_{2j} - h_{12}) + h_{12}^2 \\ & + (\mathrm{E}u_{1j}\bar{u}_{1j} - h_{11})(\mathrm{E}\bar{u}_{2j}u_{2j} - h_{22}) + (\mathrm{E}u_{1j}\bar{u}_{1j} - h_{11}) + (\mathrm{E}\bar{u}_{2j}u_{2j} - h_{22}) \\ & + h_{11}h_{22} + \mathrm{E}u_{1j}u_{2j}\mathrm{E}\bar{u}_{2j}\bar{u}_{1j} - (\mathrm{E}u_{1j}\bar{u}_{2j} - h_{12})(\mathrm{E}\bar{v}_{1j}v_{2j} - h_{12}) \\ & - (\mathrm{E}u_{1j}\bar{u}_{2j} - h_{12}) - (\mathrm{E}\bar{v}_{1j}v_{2j} - h_{12}) - h_{12}^2 - (\mathrm{E}u_{1j}\bar{v}_{1j} - h_{11})(\mathrm{E}\bar{u}_{2j}v_{2j} - h_{22}) \\ & - (\mathrm{E}u_{1j}\bar{v}_{1j} - h_{11}) - (\mathrm{E}\bar{u}_{2j}v_{2j} - h_{22}) - h_{11}h_{22} - \mathrm{E}u_{1j}v_{2j}\mathrm{E}\bar{u}_{2j}\bar{v}_{1j} \\ & - (\mathrm{E}v_{1j}\bar{v}_{2j} - h_{12})(\mathrm{E}\bar{u}_{1j}u_{2j} - h_{12}) - (\mathrm{E}v_{1j}\bar{v}_{2j} - h_{12}) - (\mathrm{E}\bar{u}_{1j}u_{2j} - h_{12}) - h_{12}^2 \\ & - (\mathrm{E}v_{1j}\bar{u}_{1j} - h_{11})(\mathrm{E}\bar{v}_{2j}u_{2j} - h_{22}) - (\mathrm{E}v_{1j}\bar{u}_{1j} - h_{11}) - (\mathrm{E}\bar{v}_{2j}u_{2j} - h_{22}) \\ & - h_{11}h_{22} - \mathrm{E}v_{1j}u_{2j}\mathrm{E}\bar{v}_{2j}\bar{u}_{1j} + (\mathrm{E}v_{1j}\bar{v}_{2j} - h_{12})(\mathrm{E}\bar{v}_{1j}v_{2j} - h_{12}) \\ & + (\mathrm{E}v_{1j}\bar{v}_{2j} - h_{12}) + (\mathrm{E}\bar{v}_{1j}v_{2j} - h_{12}) + h_{12}^2 + (\mathrm{E}v_{1j}\bar{v}_{1j} - h_{11})(\mathrm{E}\bar{v}_{2j}v_{2j} - h_{22}) \\ & + (\mathrm{E}v_{1j}\bar{v}_{1j} - h_{11}) + (\mathrm{E}\bar{v}_{2j}v_{2j} - h_{22}) + h_{11}h_{22} + \mathrm{E}v_{1j}v_{2j}\mathrm{E}\bar{v}_{2j}\bar{v}_{1j}. \end{aligned}$$

The term a_{1j} is equal to

$$\begin{aligned} & \frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^3}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)A(\mu)}{|A(\lambda_j)|^2} - 1\right)\left(\frac{A(\zeta)A(-\lambda-\mu-\zeta)}{|A(\lambda_j)|^2} - 1\right) \\ & \quad \times H_1(\lambda+\lambda_j)H_2(\mu-\lambda_j)H_1(\zeta+\lambda_j)H_2(-\lambda-\mu-\zeta-\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\mu\,\mathrm{d}\zeta, \end{aligned} \tag{30}$$

where $H_r(\lambda) = \sum_{t=1}^{T} h_r(t/T)\mathrm{e}^{\mathrm{i}t\lambda}$, $r = 1,2$. Proceeding as in the proof of (4.8) in Robinson (1995b), expression (30) can be written as a sum of components of three types. The first component is

$$\frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^3}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\mu)}{\overline{A}_j}-1\right)$$
$$\times\left(\frac{A(\zeta)}{A_j}-1\right)\left(\frac{A(-\lambda-\mu-\zeta)}{\overline{A}_j}-1\right)$$
$$\times H_1(\lambda+\lambda_j)H_2(\mu-\lambda_j)H_1(\zeta+\lambda_j)H_2(-\lambda-\mu-\zeta-\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\mu\,\mathrm{d}\zeta$$

where $A_j = A(\lambda_j)$. Using the Schwarz inequality, periodicity of the integrand and the fact that $\int_{-\pi}^{\pi}|H_r(\lambda)|^2\,\mathrm{d}\lambda = \mathrm{O}(T)$, this component can be shown to be bounded in absolute value by

$$CP_{1,j}P_{2,j},$$

where

$$P_{r,j} = \int_{-\pi}^{\pi}\left|\frac{A(\lambda)}{A_j}-1\right|^2 K_r(\lambda-\lambda_j)\,\mathrm{d}\lambda$$

and $K_r(\lambda) = (1/2\pi T)|H_r(\lambda)|^2$.

A typical representative of the second type of component of (30) is

$$\frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^3}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\mu)}{\overline{A}_j}-1\right)\left(\frac{A(\zeta)}{A_j}-1\right)$$
$$\times H_1(\lambda+\lambda_j)H_2(\mu-\lambda_j)H_1(\zeta+\lambda_j)H_2(-\lambda-\mu-\zeta-\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\mu\,\mathrm{d}\zeta$$

whose absolute value can be similarly shown to be bounded by

$$CP_{1,j}P_{2,j}^{1/2}.$$

The last type of component is exemplified by

$$\frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^3}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\zeta)}{A_j}-1\right)$$
$$\times H_1(\lambda+\lambda_j)H_2(\mu-\lambda_j)H_1(\zeta+\lambda_j)H_2(-\lambda-\mu-\zeta-\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\mu\,\mathrm{d}\zeta$$
$$=\frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^3}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\zeta)}{A_j}-1\right)$$
$$\times H_1(\lambda+\lambda_j)H_2(-\lambda-\zeta-\theta-\lambda_j)H_1(\zeta+\lambda_j)H_2(\theta-\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\zeta$$
$$=\frac{\kappa}{\sigma_\xi^4}\frac{1}{(2\pi)^2}\frac{1}{T^2}\int_{-\pi}^{\pi}\int\left(\frac{A(\lambda)}{A_j}-1\right)\left(\frac{A(\zeta)}{A_j}-1\right)$$
$$\times H_1(\lambda+\lambda_j)H_1(\zeta+\lambda_j)H_2^{(2)}(-\lambda-\zeta-2\lambda_j)\,\mathrm{d}\lambda\,\mathrm{d}\zeta \tag{31}$$

since

$$\int_{-\pi}^{\pi}H_r(u+\lambda)H_r(v-\lambda)\,\mathrm{d}\lambda = 2\pi H_r^{(2)}(u+v),$$

where $H_r^{(2)}(\lambda) = \sum_{t=1}^{T} h_r^2(t/T)\mathrm{e}^{\mathrm{i}\lambda t}$. Since $\int_{-\pi}^{\pi} |H_r^{(2)}(\lambda)|\mathrm{d}\lambda = \mathrm{O}(T)$, the modulus of (31) is bounded by

$$CT^{-1/2}P_{1,j}.$$

By Lemma 2 the term a_{1j} is $\mathrm{O}(j^{-2} + j^{-3/2} + j^{-1}T^{-1/2})$. Applying Lemma 3 gives $a_{2j} = \mathrm{O}(1)$. Therefore the first factor of a_j is $\mathrm{O}(1)$. Likewise, the second factor of a_j, and therefore a_j itself, is $\mathrm{O}(1)$.

Denoting $h_{rt} = h_r(t/T)$, the second moment of $I_{h_1\xi,h_2\xi,j}$ is

$$\begin{aligned}
\mathrm{E}|I_{h_1\varepsilon,h_2\varepsilon,j}|^2 &= \frac{1}{4\pi^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{v=1}^{T} h_{1t}h_{1s}h_{2r}h_{2v}E\varepsilon_t\varepsilon_s\varepsilon_r\varepsilon_v\mathrm{e}^{\mathrm{i}(t-s+r-v)\lambda_j} \\
&= \frac{1}{4\pi^2T^2}\left(\mathrm{E}\varepsilon_t^4\sum_{t=1}^{T} h_{1t}^2h_{2t}^2 + \sigma_\varepsilon^4\sum_{t=1}^{T}\sum_{\substack{s=1\\ s\neq t}}^{T} h_{1t}^2h_{2s}^2\right. \\
&\quad \left. +\sigma_\varepsilon^4\sum_{t=1}^{T}\sum_{\substack{s=1\\ s\neq t}}^{T} h_{1t}h_{1s}h_{2t}h_{2s}\mathrm{e}^{\mathrm{i}(t-s)2\lambda_j} + \sigma_\varepsilon^4\sum_{t=1}^{T}\sum_{\substack{s=1\\ s\neq t}}^{T} h_{1t}h_{1s}h_{2s}h_{2t}\right) \\
&= \mathrm{O}(1)
\end{aligned}$$

because the fourth moments of ε_t are finite. In the same way, the factor $I_{h_3\xi,h_4\xi,j}$ is $\mathrm{O}(1)$. Using the Schwarz inequality, the sum $a_j + b_j + c_j + d_j$ in (29) is $\mathrm{O}(1)$ uniformly over integers $1 \leqslant j \leqslant [T/2]$. The proof of the proposition is then completed by applying Lemma 5 part (a) with $g(\lambda) = A(\lambda)$. □

Proposition 7. *Under Conditions* 1–5, *as* $T \to \infty$,

$$2\,Re\begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{j=1}^{[T/2]} w_{x,j}|w_{\hat{u},j}|\eta_j^* \\ \frac{1}{\sqrt{T}}\sum_{j=1}^{[T/2]} w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^* \end{pmatrix} \overset{\mathrm{p}}{\Longrightarrow} \begin{pmatrix} \frac{1}{2\pi}\Omega^{1/2}B(1) \\ \frac{1}{2\pi}\Omega^{1/2}B(\tau) \end{pmatrix}$$

over $\tau \in [0,1]$.

Proof. Define $\eta_{T-j}^* = \overline{\eta_j^*}$ for $j = [T/2]+1,\ldots,T-1$. Then

$$2\,Re\frac{1}{\sqrt{T}}\sum_{j=1}^{[T/2]} w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^* = \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^* + r(\tau),$$

where $r(\tau) = T^{-1/2}w_{z(\tau),T/2}|w_{\hat{u},T/2}|\eta_{T/2}^*\mathbb{I}(T\text{ even}) = \mathrm{O_p}(T^{-1/2})$ uniformly over $\tau \in [0,1]$. It is therefore sufficient to show that $T^{-1/2}\sum_{j=1}^{T-1} w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^* \overset{\mathrm{p}}{\Longrightarrow} (1/2\pi)\Omega^{1/2}B(\tau)$ over $\tau \in [0,1]$. We need to prove that

(a)
$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} g_j w_{\zeta(\tau),j}|w_{\varepsilon,j}|\eta_j^* \overset{\mathrm{p}}{\Longrightarrow} \frac{1}{2\pi}\Omega^{1/2}B(\tau), \tag{32}$$

(b)
$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1}(g_j w_{\zeta(\tau),j}|w_{\varepsilon,j}|\eta_j^* - w_{z(\tau),j}|w_{u,j}|\eta_j^*) \overset{\mathrm{p}}{\Longrightarrow} 0 \tag{33}$$

and

(c)
$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1}(w_{z(\tau),j}|w_{u,j}|\eta_j^* - w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^*) \overset{\mathrm{p}}{\Longrightarrow} 0 \tag{34}$$

over $\tau \in [0,1]$ for any $\varepsilon > 0$ where $g(\lambda) = A(\lambda)\overline{B}(\lambda)$ and $g_j = g(\lambda_j)$.

To prove the convergence in part (a), we need to show that finite dimensional distributions of the process $Y_T = T^{-1/2}\sum_{j=1}^{T-1} g_j w_{\zeta(\tau),j}|w_{\varepsilon,j}|\eta_j^*$ converge in probability to the finite dimensional distributions of a centered Gaussian process with covariance function $K(\tau,\sigma) = (\tau \wedge \sigma)(1/4\pi^2)\Omega$ and that the process Y_T is tight. First, $\mathrm{E}^* Y_T(\tau) = 0$ and

$$\mathrm{var}^*\, Y_T(\tau) = \frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2|w_{\zeta(\tau),j}|^2|w_{\varepsilon,j}|^2,$$

where E^* and var^* denotes mean and variance, respectively, taken conditionally on data. By Proposition 5, the last expression converges in probability to

$$\tau\frac{1}{2\pi}\frac{\sigma_\xi^2\sigma_\varepsilon^2}{4\pi^2}\int_{-\pi}^{\pi}|g(\lambda)|^2\,\mathrm{d}\lambda = \frac{\tau}{4\pi^2}\Omega.$$

Second, we need to show that the Lindeberg condition is satisfied,

$$\sum_{j=1}^{T-1}\mathrm{E}^*|T^{-1/2}A_jB_jw_{\zeta(\tau),j}|w_{\varepsilon,j}|\eta_j^*|^2\mathbb{I}(|T^{-1/2}A_jB_jw_{\zeta(\tau),j}|w_{\varepsilon,j}|\eta_j^*|^2 > \varepsilon) \overset{\mathrm{p}}{\to} 0 \tag{35}$$

for each $\varepsilon > 0$.

We examine $\sup_\tau(1/T)|A_jB_j|^2|I_{\zeta\zeta,j}I_{\varepsilon\varepsilon,j}|$. From An et al. (1983), we have

$$\sup_{j=1,\dots,[T/2]}\left(\frac{2\pi}{\sigma_\varepsilon^2}\frac{1}{\log T}|w_{\varepsilon,j}|^2\right) \leqslant 1 \quad \text{a.s.}$$

and

$$\sup_{j=1,\dots,[T/2]}\left(\frac{2\pi}{\sigma_\xi^2}\frac{1}{\log T}|w_{\xi,j}|^2\right) \leqslant 1 \quad \text{a.s.}$$

Therefore

$$\begin{aligned}\sup_{j=1,\dots,[T/2]}\frac{1}{T}|A_jB_j|^2|I_{\zeta\zeta,j}I_{\varepsilon\varepsilon,j}| &\leqslant D\sup_j\frac{1}{T}|A_jB_j|^2\log^2 T \quad \text{a.s.}\\ &\leqslant DT^{2d-1}\log^2 T \quad \text{a.s.,}\end{aligned}$$

where $d = d_x + d_u < \frac{1}{2}$. As η_j^*, given the data, are independent identically distributed variables, the sum in (35) is bounded by

$$\mathrm{E}^*|\eta_j^*|^2 \mathbb{1}(|\eta_j^*|^2 > \varepsilon T^{1-2d} \log^{-2} T) \frac{2}{T} \sum_{j=1}^{[T/2]} |A_j B_j|^2 I_{\zeta\zeta,j} I_{\varepsilon\varepsilon,j}.$$

The first factor converges to zero as $T \to \infty$ since η_j^* has finite moments and $1 - 2d > 0$. The second factor is $\mathrm{O_p}(1)$ by Proposition 5 with $g_j = A_j \overline{B}_j$. Therefore the left-hand side of (35) is $\mathrm{o_p}(1)$ and by the Lindeberg–Feller central limit theorem the pointwise convergence

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{T-1} g_j w_{\zeta(\tau),j} |w_{\varepsilon,j}| \eta_j^* \xrightarrow{\mathrm{d}} N\Big(0, \frac{\tau}{4\pi^2} \Omega\Big)$$

in probability is proved.

Further,

$$\mathrm{cov}^*(Y_T(\tau), Y_T(\sigma)) = \frac{1}{T} \sum_{j=1}^{T-1} |g_j|^2 w_{\zeta(\tau),j} \bar{w}_{\zeta(\sigma),j} |w_{\varepsilon,j}|^2$$

which converges in probability to $(\tau \wedge \sigma)(1/4\pi^2)\Omega$ by Proposition 5. The proof of convergence of the finite dimensional distributions in part (a) is completed by using the Cramér–Wold device.

We now prove tightness of the process $Y_T(\tau)$. By Theorem 13.5 of Billingsley (1999) it is sufficient to check the moment condition

$$\mathrm{E}^*|Y_T(\rho) - Y_T(\sigma)|^2 |Y_T(\tau) - Y_T(\rho)|^2 \leqslant (1 + \mathrm{o_p}(1))(F(\tau) - F(\sigma))^\alpha, \tag{36}$$

where $\alpha > \frac{1}{2}$, $\sigma \leqslant \rho \leqslant \tau$, F is a nondecreasing continuous function on Λ and $\mathrm{o_p}(1)$ is uniform over $(\tau, \sigma) \in \Lambda^2$. Denoting $w_j = w_{\zeta(\tau),j} - w_{\zeta(\sigma),j}$, we have

$$\begin{aligned}
\mathrm{E}^*|Y_T(\tau) - Y_T(\sigma)|^4 &= \frac{1}{T^2} \sum_{j=1}^{T-1} |g_j|^4 |w_j|^4 |w_{\varepsilon,j}|^4 \mathrm{E}^* |\eta_j^*|^4 \\
&\quad + \frac{2}{T^2} \sum_{j=1}^{T-1} \sum_{\substack{k=1 \\ k \neq j}}^{T-1} |g_j g_k|^2 |w_j|^2 |w_k|^2 |w_{\varepsilon,j}|^2 |w_{\varepsilon,k}|^2 E^* |\eta_j^*|^2 |\eta_k^*|^2 \\
&\quad + \frac{1}{T^2} \sum_{j=1}^{T-1} \sum_{\substack{k=1 \\ k \neq j}}^{T-1} g_j^2 \bar{g}_k^2 w_j^2 \bar{w}_k^2 |w_{\varepsilon,j}|^2 |w_{\varepsilon,k}|^2 \mathrm{E}^* \eta_j^{*2} \bar{\eta}_k^{*2} \\
&\leqslant \frac{C}{T^2} \sum_{j=1}^{T-1} \sum_{k=1}^{T-1} |g_j g_k|^2 |w_j|^2 |w_k|^2 |w_{\varepsilon,j}|^2 |w_{\varepsilon,k}|^2 \\
&= C \left(\frac{1}{T} \sum_{j=1}^{T-1} |g_j|^2 |w_j|^2 |w_{\varepsilon,j}|^2 \right)^2.
\end{aligned}$$

By Proposition 5,

$$C\left(\frac{1}{T}\sum_{j=1}^{T-1}|g_j|^2|w_j|^2|w_{\varepsilon,j}|^2\right)^2 \Rightarrow C(\tau-\sigma)^2\frac{1}{4\pi^2}\Omega$$

over $(\tau,\sigma)\in[0,1]^2$. It follows that by the Schwarz inequality the left-hand side of (36) is bounded by $D^2(\tau-\sigma)^2(1+\mathrm{o_p}(1))$ since $(\tau-\rho)(\rho-\sigma)\leqslant(\tau-\sigma)^2$. The moment condition (36) is thus verified with $F(\tau)=D\tau$ and $\alpha=2$. This establishes tightness in probability of the process Y_T and completes the proof of the uniform convergence in part (a).

For the convergence in part (b), we have

$$\begin{aligned}
&\mathrm{E}^*\left|\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1}(A_j\overline{B}_j w_{\zeta(\tau),j}|w_{\varepsilon,j}| - w_{z(\tau),j}|w_{u,j}|)\eta_j^*\right| \\
&\leqslant \frac{D}{\sqrt{T}}\sum_{j=1}^{T-1}|A_jB_j|\left|\frac{w_{z(\tau),j}}{A_j}-w_{\zeta(\tau),j}\right|\left|\frac{w_{u,j}}{B_j}-w_{\varepsilon,j}\right| \\
&\quad+\frac{D}{\sqrt{T}}\sum_{j=1}^{T-1}|A_jB_j|\left|\frac{w_{z(\tau),j}}{A_j}-w_{\zeta(\tau),j}\right||w_{\varepsilon,j}| \\
&\quad+\frac{D}{\sqrt{T}}\sum_{j=1}^{T-1}|A_jB_j||w_{\zeta(\tau),j}|\left|\frac{w_{u,j}}{B_j}-w_{\varepsilon,j}\right|
\end{aligned}$$

and proceeding as in the proof of Proposition 2 above it can be shown that the last expression is $\mathrm{o_p}(1)$ uniformly over $\tau\in[0,1]$.

To verify the convergence in part (c), we write the difference between errors and residuals under the local alternative as

$$u_t-\hat{u}_t=(\hat{\alpha}-\alpha)+(\hat{\beta}-\beta)x_t+\hat{\delta}\hat{z}_t-\frac{1}{\sqrt{T}}x_th_t,$$

where $\hat{z}_t=z_t(\hat{\tau})=x_t\mathbb{I}(t\leqslant[\hat{\tau}T])$ and $h_t=h(t/T)$. Therefore

$$w_{u,j}-w_{\hat{u},j}=(\hat{\beta}-\beta)w_{x,j}+\hat{\delta}w_{z(\hat{\tau}),j}-\frac{1}{\sqrt{T}}w_{hx,j},$$

$j=1,\dots,T-1$, where $w_{hx,j}$ is the discrete Fourier transform of the sequence $\{h_tx_t,1\leqslant t\leqslant T\}$. Since $||w_{u,j}|-|w_{\hat{u},j}||^2\leqslant|w_{u,j}-w_{\hat{u},j}|^2$,

$$\begin{aligned}
&\mathrm{E}^*\left|\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1}w_{z(\tau),j}(|w_{u,j}|-|w_{\hat{u},j}|)\eta_j^*\right|^2 \\
&=\frac{1}{T}\sum_{j=1}^{T-1}|w_{z(\tau),j}|^2\||w_{u,j}|-|w_{\hat{u},j}|\|^2\mathrm{E}^*|\eta_j^*|^2 \\
&\leqslant\frac{1}{T}\sum_{j=1}^{T-1}|w_{z(\tau),j}|^2|w_{u,j}-w_{\hat{u},j}|^2
\end{aligned}$$

$$\leqslant 3(\hat{\beta}-\beta)^2 \frac{1}{T}\sum_{j=1}^{T-1} |w_{z(\tau),j}|^2 |w_{x,j}|^2$$
$$+ 3\hat{\delta}^2 \frac{1}{T}\sum_{j=1}^{T-1} |w_{z(\tau),j}|^2 |w_{z(\hat{\tau}),j}|^2$$
$$+ \frac{3}{T^2}\sum_{j=1}^{T-1} |w_{z(\tau),j}|^2 |w_{hx,j}|^2. \tag{37}$$

By Theorem 2, $\hat{\beta}-\beta = \mathrm{O_p}(T^{-1/2})$ and $\hat{\delta} = \mathrm{O_p}(T^{-1/2})$. Also, by Proposition 6 with functions $h_1(x) = h_2(x) = \mathbb{I}(0 \leqslant x \leqslant \tau)$ and $h_3(x) = h_4(x) = \mathbb{I}(0 \leqslant x \leqslant \hat{\tau})$,

$$\frac{1}{T}\sum_{j=1}^{T-1} |w_{z(\tau),j}|^2 |w_{z(\hat{\tau}),j}|^2 = \mathrm{o_p}(T)$$

uniformly over $\tau \in [0,1]$ and similarly for the other sums. Therefore the right-hand side of the last displayed inequality is $\mathrm{o_p}(1)$ uniformly over $[0,1]$. The uniform convergence in (34) is established by using the Markov inequality. □

Replacing $|w_{\varepsilon,j}|$, $|w_{u,j}|$ and $|w_{\hat{u},j}|$ in (32)–(34) by $w_{\varepsilon,j}$, $w_{u,j}$ and $w_{\hat{u},j}$, and drawing η_j^* from any complex distribution with mean zero, unit variance, finite fourth moment and with $\mathrm{E}\eta_j^{*2} = 0$, it can be seen that the proof remains valid with only small modifications. In particular, expressions for $\mathrm{var}^* Y_T(\tau)$ and $\mathrm{cov}^*(Y_T(\tau), Y_T(\sigma))$ do not change, inequalities in part (a) for suprema in the Lindeberg condition and for $\mathrm{E}^*|Y_T(\tau) - Y_T(\sigma)|^4$, in part (b) for the conditional first moment and in part (c) for the conditional second moment continue to hold with minor changes in intermediate steps where required. This observation shows that there are several valid modifications of the basic bootstrap procedure described in Section 3.

Proof of Theorems 1 and 2. Under the local alternative,

$$\sqrt{T}\begin{pmatrix} \hat{\beta}(\tau)-\beta \\ \hat{\delta}(\tau) \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \frac{1}{T}\sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} I_{xu}(\lambda_j) \\ \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} I_{zu}(\lambda_j) \end{pmatrix}$$
$$+ \begin{pmatrix} \frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \frac{1}{T}\sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} w_{x,j}\overline{w}_{hx,j} \\ \frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} w_{z,j}\overline{w}_{hx,j} \end{pmatrix}. \tag{38}$$

By Proposition 4 with $h_1(x) = 1$ and $h_2(x) = \mathbb{I}(x \leqslant \tau)$,

$$\frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \xrightarrow{\mathrm{p}} \frac{\tau}{2\pi}\Sigma.$$

Similarly, $(1/T)\sum_{j=1}^{T-1} I_{xx}(\lambda_j) \xrightarrow{p} (1/2\pi)\Sigma$ and $1/T\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \xrightarrow{p} (\tau/2\pi)\Sigma$, and therefore

$$\begin{pmatrix} \frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \frac{1}{T}\sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix} \Longrightarrow \begin{pmatrix} 1 & \tau \\ \tau & \tau \end{pmatrix} \otimes \frac{1}{2\pi}\Sigma$$

over $[0, 1]$ as $T \to \infty$. Since matrix inverse is a continuous function for $\tau \in \Lambda$,

$$\begin{pmatrix} \frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \frac{1}{T}\sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \Longrightarrow \begin{pmatrix} 1 & \tau \\ \tau & \tau \end{pmatrix}^{-1} \otimes 2\pi\Sigma^{-1}$$

over Λ. Under the null, that is when $h \equiv 0$, the second term on the right-hand side of (38) vanishes. By Proposition 2,

$$\sqrt{T}\begin{pmatrix} \hat{\beta}(\tau) - \beta \\ \hat{\delta}(\tau) \end{pmatrix} \Longrightarrow \frac{1}{\tau(1-\tau)} \begin{pmatrix} \Sigma^{-1}\Omega^{1/2}(\tau B(1) - \tau B(\tau)) \\ \Sigma^{-1}\Omega^{1/2}(B(\tau) - \tau B(1)) \end{pmatrix}$$

and Theorem 1 is proved.

Under the alternative, $h \neq 0$ and by Proposition 4

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T-1} w_{z,j}\overline{w}_{hx,j} \Longrightarrow \frac{1}{2\pi}\Sigma \int_0^\tau h(t)\,\mathrm{d}t$$

over $[0, 1]$. Therefore the second term in (38) converges to

$$\frac{1}{\tau(1-\tau)} \begin{pmatrix} \tau\int_\tau^1 h(u)\,\mathrm{d}u \\ (\int_0^\tau h(u)\,\mathrm{d}u - \tau\int_0^1 h(u)\,\mathrm{d}u) \end{pmatrix}$$

and Theorem 2 is established. $\square$

Proof of Theorem 3. By Theorem 1 of Robinson (1998):

$$\frac{4\pi^2}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) I_{uu}(\lambda_j) \xrightarrow{p} \Omega. \tag{39}$$

Proceeding as in part (c) of the proof of Proposition 7, write

$$w_{u,j} - w_{\hat{u},j} = (\hat{\beta} - \beta)w_{x,j} + \hat{\delta}w_{z(\hat{\tau}),j} - \frac{1}{\sqrt{T}}w_{hx,j}.$$

Therefore

$$I_{\hat{u}\hat{u},j} - I_{uu,j} = |w_{u,j} - w_{\hat{u},j}|^2 - 2\,Re(w_{\hat{u},j} - w_{u,j})\overline{w}_{u,j}$$

and

$$\left|\frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j)(I_{\hat{u}\hat{u},j} - I_{uu,j})\right| \leqslant \frac{1}{T}\sum_{j=1}^{T-1} |w_{x,j}|^2 |w_{u,j} - w_{\hat{u},j}|^2 + \frac{1}{T}\sum_{j=1}^{T-1} |w_{x,j}|^2 |w_{u,j} - w_{\hat{u},j}\|w_{u,j}|.$$

The first term is $o_p(1)$ as shown for (37) in Proposition 7, part (c). By the Schwarz inequality, the second term is bounded by

$$\left(\frac{1}{T}\sum_{j=1}^{T-1} |w_{x,j}|^2 |w_{u,j} - w_{\hat{u},j}|^2\right)^{1/2} \left(\frac{1}{T}\sum_{j=1}^{T-1} |w_{x,j}|^2 |w_{u,j}|^2\right)^{1/2}$$

whose second factor is $O_p(1)$ because of (39). Therefore indeed $\hat{\Omega} \xrightarrow{p} \Omega$. □

Proof of Theorem 4. Write

$$\sqrt{T}\begin{pmatrix} \hat{\beta}^*(\tau) - \hat{\beta} \\ \hat{\delta}^*(\tau) \end{pmatrix} = \begin{pmatrix} \frac{1}{T}\sum_{j=1}^{T-1} I_{xx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{xz}(\lambda_j) \\ \frac{1}{T}\sum_{j=1}^{T-1} I_{zx}(\lambda_j) & \frac{1}{T}\sum_{j=1}^{T-1} I_{zz}(\lambda_j) \end{pmatrix}^{-1} \times 2\, Re \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{j=1}^{[T/2]} w_{x,j}|w_{\hat{u},j}|\eta_j^* \\ \frac{1}{\sqrt{T}}\sum_{j=1}^{[T/2]} w_{z(\tau),j}|w_{\hat{u},j}|\eta_j^* \end{pmatrix}.$$

Applying Propositions 4 and 7, it can be seen that Theorem 4 holds. □

Acknowledgements

This paper is a part of my Ph.D. Dissertation at London School of Economics. I gratefully acknowledge research support from the Denis Sargan Memorial Fund and from the Economic and Social Research Council through grant R000238212. The first version of this paper was presented at the International Conference on Modelling Structural Breaks, Long Memory and Stock Market Volatility held in London at Cass Business School, 6–7 December 2002. I would like to thank Javier Hidalgo, Liudas Giraitis, three referees, participants of the conference at Cass Business School, 2003 North American Summer Meeting in Evanston, 2003 European Meeting of the Econometric Society in Stockholm, 2003 European Winter Meeting of the Econometric Society in Madrid and 2004 North American Winter Meeting of the Econometric Society in San Diego, and participants in seminars at LSE and Queen Mary London for helpful suggestions.

References

An, H.-Z., Zhao-Guo, C., Hannan, E.J., 1983. The maximum of periodogram. Journal of Multivariate Analysis 13, 383–400.

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

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Billingsley, P., 1999. Convergence of Probability Measures, second ed. Wiley, New York.

Brown, R.L., Durbin, J., Evans, J.M., 1975. Techniques for testing the constancy of regression relationships over time. Journal of the Royal Statistical Society Series B 37, 149–192.

Bühlmann, P., 1997. Sieve bootstrap for time series. Bernoulli 3, 123–148.

Bühlmann, P., 1998. Sieve bootstrap for smoothing in nonstationary time series. The Annals of Statistics 26, 48–83.

Carlstein, E., 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. The Annals of Statistics 14, 1171–1179.

Chu, C.-S., Hornik, K., Kuan, C.-M., 1994. MOSUM tests for parameter constancy. Biometrika 82, 603–617.

Dahlhaus, R., Janas, D., 1996. A frequency domain bootstrap for ratio statistics in time series analysis. The Annals of Statistics 24, 1934–1963.

Davis, R.B., Harte, D.S., 1987. Tests for Hurst effect. Biometrika 74, 95–101.

Diebold, F.X., Ohanian, L.E., Berkowitz, J., 1998. Dynamic equilibrium economies: A framework for comparing models and data. Review of Economic Studies 65, 433–451.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. The Annals of Statistics 7, 1–26.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Eicker, F., 1979. The asymptotic distribution of the suprema of the standardized empirical processes. The Annals of Statistics 7, 116–138.

Franke, J., Härdle, W., 1992. On bootstrapping kernel spectral estimates. The Annals of Statistics 20, 121–145.

Giné, E., Zinn, J., 1990. Bootstrapping general empirical measures. The Annals of Probability 18, 851–869.

Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15–29.

Giraitis, L., Surgailis, D., 1990. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. Probability Theory and Related Fields 86, 87–104.

Hidalgo, J., 2003. An alternative bootstrap to moving blocks for time series regression models. Journal of Econometrics 117, 369–399.

Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165–176.

Jaeschke, D., 1979. The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. The Annals of Statistics 7, 108–115.

Keenan, D.M., 1987. Limiting behavior of functionals of higher-order sample cumulant spectra. The Annals of Statistics 15, 134–151.

Kuan, C.-M., Hornik, K., 1995. The generalized fluctuation test: a unifying view. Econometric Reviews 14, 135–161.

Künsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations. The Annals of Statistics 17, 1217–1241.

Mandelbrot, B.B., van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. SIAM Review 10, 422–437.

Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223–230.

Paparoditis, E., Politis, D.N., 2000. The local bootstrap for kernel estimators under general dependence conditions. Annals of the Institute of Statistical Mathematics 52, 139–159.

Ploberger, W., Krämer, W., 1990. The local power of the CUSUM of squares tests. Econometric Theory 6, 335–347.

Ploberger, W., Krämer, W., 1992. The CUSUM test with OLS residuals. Econometrica 60, 271–285.

Ploberger, W., Krämer, W., Kontrus, K., 1989. A new test for structural stability in the linear regression model. Journal of Econometrics 40, 307–318.

Politis, D.N., Romano, J.P., 1992. A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation. The Annals of Statistics 20, 1985–2007.

Quandt, R.E., 1960. Tests of the hypothesis that a linear regression system obeys two separate regimes. Journal of the American Statistical Association 55, 324–330.

Ramos, E., 1984. A bootstrap for time series. Qualifying paper, Department of Statistics, Harvard University.

Robinson, P.M., 1994. Rates of convergence and optimal spectral bandwidth for long range dependence. Probability Theory and Related Fields 99, 443–473.

Robinson, P.M., 1995a. Log-periodogram regression of time series with long range dependence. The Annals of Statistics 23, 1048–1072.

Robinson, P.M., 1995b. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 23, 1630–1661.

Robinson, P.M., 1998. Inference-without-smoothing in the presence of nonparametric autocorrelation. Econometrica 66, 1163–1182.

Robinson, P.M., Hidalgo, F.J., 1997. Time series regression with long-range dependence. The Annals of Statistics 25, 77–104.

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.

Sen, P.K., 1980. Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 52, 203–218.

Taniguchi, M., 1982. On estimation of the integrals of the fourth order cumulant spectral density. Biometrika 69, 117–122.