HIGHER ORDER APPROXIMATIONS FOR WALD STATISTICS IN TIME SERIES REGRESSIONS WITH INTEGRATED PROCESSES

BY

ZHIJIE XIAO and PETER C. B. PHILLIPS

COWLES FOUNDATION PAPER NO. 1094

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2004

http://cowles.econ.yale.edu/

ELSEVIER

Journal of Econometrics 108 (2002) 157–198

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Higher order approximations for Wald statistics in time series regressions with integrated processes

Zhijie Xiao[a], Peter C.B. Phillips[b,*]

[a]Department of Economics, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, IL 61820, USA

[b]Cowles Foundation for Research in Economics, Yale University, P.O. Box 208281, New Haven, CT 06520, USA; University of Auckland & University of York

Received 16 June 1998; received in revised form 6 September 2001; accepted 13 November 2001

Abstract

Asymptotic expansions are developed for Wald test statistics in time series regressions with integrated processes. These expansions provide an opportunity to reduce size distortion in testing by suitable bandwidth selection, and automated rules for doing so are calculated. A band spectral regression and the associated Wald test are also considered. Both the first order and second order properties of the estimator are studied. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: C14; C22

Keywords: Bandwidth selection; Higher order approximation; Moment expansion; Spectral regression

1. Introduction

This paper studies the Wald statistic in time series regression models of the form

$$y_t = \beta' x_t + u_t, \quad t = 1, \ldots, n, \tag{1.1}$$

where u_t is an $I(0)$ process with zero mean and continuous spectral density $f_{uu}(\lambda)$. The regressor x_t is a vector of integrated time series of order one ($I(1)$) or zero ($I(0)$). Both $I(0)$ and $I(1)$ regressors are studied here, but the main focus of the paper is on the case with $I(1)$ regressors. In this event, if x_t is $I(1)$, then both x_t and y_t are nonstationary and the linear combination $y_t - \beta' x_t$ is stationary, so that (1.1) is a cointegrating regression in the sense of Engle and Granger (1987). More will be

* Corresponding author. Tel.: +1-203-432-3695; fax: +1-203-432-5429.
E-mail address: peter.phillips@yale.edu (P.C.B. Phillips).

PII: S0304-4076(01)00134-8

said later about the alternate specifications of the model as we proceed to develop the higher order theory.

We consider the frequency domain version of (1.1), viz.

$$w_y(\lambda_s) = \beta' w_x(\lambda_s) + w_u(\lambda_s), \quad \lambda_s = 2\pi s/n, \ s = -n/2+1, \ldots, n/2, \tag{1.2}$$

where $w_y(\lambda_s)$, $w_x(\lambda_s)$, and $w_u(\lambda_s)$ are discrete Fourier transforms of y_t, x_t, and u_t, at the fundamental frequencies λ_s defined by $w_a(\lambda_s) = (1/\sqrt{2\pi n})\sum_{t=1}^{n} a_t e^{it\lambda_s}$. For convenience we simply assume that n is even. If the residual process in (1.1), u_t, is stationary with a continuous spectrum that is bounded away from the origin, the residuals of regression (1.2), $w_u(\lambda_s)$, are asymptotically uncorrelated but generally heteroskedastic. Regression in the frequency domain permits a nonparametric treatment of the regression errors and, utilizing a consistent estimate of $f_{uu}(\lambda)$, delivers a semi-parametric estimator of β that is asymptotically equivalent to GLS in the case of stationary u_t and x_t. Hence, it is unnecessary to parameterize the autocorrelation structure of u_t to achieve efficient estimation. In addition, these methods allow attention to be focused on the most relevant frequency in the regression, thereby providing a selective approach that accommodates more general formulations in which the parameter β may not be constant across frequency bands (Of course, in this case, the time domain counterpart of regression (1.2) will be more complicated—see Corbae et al., 2002.) When x_t is $I(1)$ and independent of u_t, OLS regression is asymptotically efficient in (1.1), as shown by Phillips and Park (1988), but efficiency gains still acrue from GLS in finite samples (see Section 2 for a discussion).

Regression analysis in the frequency domain was introduced by Hannan (1963a), following ideas of Whittle (1951), and extended to nonlinear models by Hannan (1971) and Robinson (1976). Robinson (1991) studied automatic frequency domain inference on semiparametric and nonparametric models where bandwidth selection for the spectral density estimate is determined from the data. Phillips (1991) showed how to apply this method to cointegrating regressions, developed a limit theory for frequency domain estimators in this case and established some optimality properties for such regressions. In addition, a recent paper by Corbae et al. (2002) provides asymptotic results on the behavior of discrete Fourier transforms of integrated processes and deterministic trends and explores spectral regression methods in the presence of deterministic trends.

The present paper studies the use of Wald statistics in linear time series regressions like (1.1). Finite sample problems of over-rejection in the use of Wald tests in such regressions have long been recognized by econometricians, an example being the well studied phenomenon of over-rejection of homogeneity conditions in empirical consumer demand analysis (see Barten, 1969; Byron, 1970; Lluch, 1971; Deaton, 1974). Monte Carlo results in such cases have shown that the Wald test is biased toward rejecting the null hypothesis (e.g. Laitinen, 1978; Meisner, 1979; Bera et al., 1981).

One of the mechanisms for improving asymptotic χ^2 approximations of Wald tests is the use of higher order expansions. The statistical theory of asymptotic expansions for Wald tests has been extensively studied in econometrics (see Sargan and Mikhail, 1971; Sargan, 1976; Phillips, 1977; Phillips, 1987; Phillips and Park, 1988; Rothenberg, 1984a,b; and Linton, 1995a, b). However, higher order expansions have not so far been developed and used in nonstationary time series environments, with an exception in Phillips (1987), largely because of the difficulty in developing valid higher order extensions of the underlying functional central limit theory on which the nonstationary regression asymptotics typically depend.

This paper seeks to implement a simple and usable approach to the development of a higher order theory for Wald tests in time series regressions with $I(1)$ regressors, avoiding the need for higher order extensions of the functional limit theory. Our approach is formal, rather than rigorous. In particular, we develop formal moment approximations, in terms of moments of approximating expansions, rather than rigorous asymptotic expansions of the moments of the Wald statistics. These formal expansions lead quite simply to formulae that can be used in practical applications. In efficient semiparametric time series regression a critical element in the construction of the estimator and associated tests is the choice of the bandwidth in the estimation of the spectrum. The idea behind the present development is to construct a bandwidth selection criterion by minimizing the second order effect on the expected value of the Wald statistic. A second order adjusted Wald statistic can then be constructed to correct the size distortion of the Wald test in finite samples. Results from this approach are given also for stationary time series regressions.

The paper is organized as follows. The model and test statistics are studied in the next section. Some preliminary results for spectral density estimation and spectral regression with integrated regressors are given in Section 3. The expansion for the Wald statistic and a modified Wald test in the univariate case is given in Section 4. The case of a narrow band spectral regression is studied in Section 5. Section 6 gives the results in the multivariate case. The results of a small Monte Carlo experiment are reported in Section 7. Section 8 concludes and some proofs are given in the Appendix.

2. Background and assumptions

Our subject is the Wald statistic in regression (1.1), where the regressor x_t is either an $I(1)$ or an $I(0)$ process. An intercept term can also be included in (1.1) simply by dropping the zero frequency in (1.2) in estimation. Our main assumptions are as follows, although, for some results, further specific assumptions may be needed and these are stated as they are required. When x_t is $I(1)$, we assume that $\Delta x_t = v_t$, and that x_t is initialized at $t=0$ by any $O_p(1)$ random variables. The quantities v_t and u_t are independent stationary and ergodic k-vector and scalar time series with zero mean, finite fourth moments and spectral densities $f_{vv}(\lambda)$ and $f_{uu}(\lambda)$, which are bounded away from zero. We assume that $\sum_h |h|^q |\gamma_u(h)| < \infty$, for some $q > 2$, where $\gamma_u(h) = cov(u_t, u_{t+h})$.[1] The vector x_t and partial sums of u_t both satisfy invariance principles with independent limit processes, so that, as $n \to \infty$, $n^{-1/2} x_{[nr]} \Rightarrow B_x(r) \equiv BM(2\pi f_{vv}(0))$, a vector Brownian motion of dimension k with covariance matrix $2\pi f_{vv}(0)$, and $n^{-1/2} \sum_{t=1}^{[nr]} u_t \Rightarrow B_u(r) \equiv BM(2\pi f_{uu}(0))$. In the case that x_t is an $I(0)$ regressor, we assume that x_t satisfies the same conditions as those given above for u_t.

[1] This summability condition not only implies that the spectral density is continuous and bounded, but also imply the uniform boundedness of $f_q(\lambda)$ which is defined below (see, e.g. Brillinger, 1980).

Table 1
Finite sample comparison between OLS and GLS

	n		OLS	GLS1	GLS2	GLS3
$\alpha = 0.3$	64	BIAS	−0.001469	−0.000921	−0.001120	−0.001292
		STD	0.049520	0.045074	0.046907	0.047975
	128	BIAS	−0.000117	−0.000087	−0.000018	−0.000012
		STD	0.025097	0.024189	0.024747	0.024798
	256	BIAS	0.000560	0.000556	0.000522	0.000563
		STD	0.013086	0.012652	0.012956	0.012995
$\alpha = 0.6$	64	BIAS	−0.002248	−0.001902	−0.001924	−0.001966
		STD	0.082229	0.065304	0.066750	0.068112
	128	BIAS	0.000185	0.000170	0.000189	0.000205
		STD	0.042423	0.036765	0.037645	0.038616
	256	BIAS	0.000959	0.000892	0.000949	0.000998
		STD	0.022698	0.020049	0.020938	0.021068
$\alpha = 0.8$	64	BIAS	−0.003498	−0.003421	−0.003606	−0.003802
		STD	0.148589	0.084445	0.090419	0.093211
	128	BIAS	0.000668	0.000603	0.000627	0.000678
		STD	0.079815	0.051375	0.056211	0.058081
	256	BIAS	0.001968	0.001343	0.001618	0.001731
		STD	0.044103	0.031097	0.032104	0.034001

The data were generated by (1.1) with $u_t = \alpha u_{t-1} + \varepsilon_t$, and $x_t = x_{t-1} + v_t$, where ε_t and v_t are iid N(0,1) and are independent ofeach other, $x_0 = u_0 = 0$. STD denotes standard deviation. GLS1, GLS2, GLS3 uses bandwidth $M = n^{1/3}, 2n^{1/3}, 3n^{1/3}$, respectively.

Under these assumptions, a frequency domain efficient estimator of β is obtained from a regression with weighted averages of periodogram estimates at the fundamental frequencies, viz

$$\hat{\beta} = \left[\sum_s I_{xx}(\lambda_s)\hat{f}_{uu}^{-1}(\lambda_s)\right]^{-1}\left[\sum_s I_{xy}(\lambda_s)\hat{f}_{uu}^{-1}(\lambda_s)\right], \tag{2.1}$$

where $I_{xx}(\lambda_s) = w_x(\lambda_s)w_x(\lambda_s)^*$, $I_{xy}(\lambda_s) = w_x(\lambda_s)w_y(\lambda_s)^*$ (the affix * signifies the complex conjugate transpose), $\hat{f}_{uu}(\lambda_s)$ is a nonparametric estimator of $f_{uu}(\lambda_s)$ and the summation $\sum_s$ is taken over $[-\pi, \pi]$. When x_t is an $I(0)$ process, it is well known that the frequency domain GLS estimator (2.1) is more efficient than the OLS estimator and achieves asymptotically the Gauss–Markov efficiency bound under the given smoothness condition on the residual spectral density. In the case that x_t is an $I(1)$ process, OLS estimation of β is (first order) asymptotically equivalent to the GLS estimator (Phillips and Park, 1988). However, such equivalence holds only asymptotically and, in general, the GLS regression estimator (2.1) still has better small sample properties. Table 1 reports some Monte Carlo results on the comparison between OLS and (2.1) in a simple version of (1.1) with autocorrrelated errors. The results indicate that, although OLS is asymptotically efficient when the regressor is $I(1)$, (2.1) generally has superior finite sample performance.

We are interested in testing the linear hypothesis

$$H_0\colon R\beta = r, \tag{2.2}$$

where r is an $p \times 1$ vector of constant and R is an $p \times k$ matrix. The regression Wald statistic corresponding to $\hat{\beta}$ in (2.1) is given by

$$W = (R\hat{\beta} - r)'[R\hat{\Sigma}^{-1}R']^{-1}(R\hat{\beta} - r), \tag{2.3}$$

where

$$\hat{\Sigma} = \sum_s I_{xx}(\lambda_s)\hat{f}_{uu}^{-1}(\lambda_s).$$

All of the theory we develop carries over with minor changes to the case of analytic nonlinear restrictions on β in place of (2.2) and associated changes in the Wald statistic (2.3).

The following lemma is closely related to earlier work in Phillips (1991) and gives the limit theory of the estimator $\hat{\beta}$ and the corresponding Wald statistic (2.3) when x_t is an $I(1)$ process.

Lemma 1. *As* $n \to \infty$:

(1) $n(\hat{\beta} - \beta) \to [\int B_x B_x']^{-1}[\int B_x \, \mathrm{d}B_u] = MN(0, 2\pi f_{uu}(0)[\int B_x B_x']^{-1})$.

(2) *Under* H_0, $W \xrightarrow{\mathrm{d}} \chi_p^2$.

As the lemma shows, the asymptotic distribution of the regression estimator is mixed normal with matrix mixing variate $[\int B_x B_x']^{-1}$ that depends on the Brownian motion $B_x(r)$. This mixed normal distribution facilitates the construction of a regression Wald test about β by using an estimate of the conditional covariance matrix $\Sigma = \sum_s I_{xx}(\lambda_s) f_{uu}^{-1}(\lambda_s)$. As a result, the Wald statistic has an asymptotic χ^2 distribution under H_0, and the null hypothesis can be tested using conventional χ^2 limit theory distribution. However, the asymptotic χ^2 distribution by no means always provides a good approximation to the distribution of W, and Monte Carlo evidence indicates that Wald tests often over-reject the null hypothesis in finite samples. This paper proceeds to derive a higher order expansion for the expected Wald statistic. The expansion helps to explain the finite sample performance of the Wald test and to make compensating adjustments using second order effects.

3. Some preliminary expansions

We consider the 'leave-one-out' type of nonparametric estimator for the spectral density of the residual process u_t, viz

$$\begin{aligned}\hat{f}_{uu}(\lambda_s) &= \frac{1}{m} \sum_{\lambda_j \in B(\lambda_s), \lambda_j \neq \lambda_s} K(\lambda_j - \lambda_s) w_{\hat{u}}(\lambda_j) w_{\hat{u}}(\lambda_j)^* \\ &= \frac{1}{m} \sum_{\lambda_j \in B(\lambda_s), \lambda_j \neq \lambda_s} K(\lambda_j - \lambda_s) I_{\hat{u}\hat{u}}(\lambda_j) \\ &= \sum_{j \neq s} \omega_{sj} I_{\hat{u}\hat{u}}(\lambda_j),\end{aligned}$$

where

$$\omega_{sj} = \begin{cases} K(\lambda_j - \lambda_s)/m, & \lambda_j \in B(\lambda_s), \\ 0 & \text{otherwise,} \end{cases}$$

and $B(\lambda_s) = \{\omega\colon \lambda_s - \pi/2M < \omega \leqslant \lambda_s + \pi/2M\}$ is a frequency band of width π/M centered on $\lambda_s = 2\pi s/n$. Let $m = [n/2M]$, where $[\cdot]$ signifies integer part. Then, each band $B(\lambda_t)$ contains m fundamental frequencies λ_s. $K(\cdot)$ is a spectral window and satisfies conventional properties of being a real, even function with $1/m \sum_{\lambda_j \in B(\omega)} K(\lambda_j - \omega) = 1$. The corresponding lag window is $k(h/M) = 1/m \sum_{\lambda_s \in B(\omega)} K(\lambda_s - \omega) \mathrm{e}^{-\mathrm{i}h(\lambda_s - \omega)}$. Many candidate kernel functions are available and are discussed in standard texts of spectral analysis (e.g., Hannan, 1970; Brillinger, 1980; Priestley, 1981). The periodogram ordinates $I_{\hat{u}\hat{u}}(\lambda_s)$ are calculated using consistent estimates of the residuals. In our analysis, we use the residuals from an OLS regression on (1.2).

It is convenient to decompose the error term in the nonparametric spectral density estimator $\hat{f}_{uu}(\lambda_s)$ into a bias effect due to smoothing, B_s, a variance term arising from the periodogram, V_s, and an error coming from the preliminary estimation of $w_u(\lambda_s)$, P_s. Thus, we have

$$\hat{f}_{uu}(\lambda_s) = f_{uu}(\lambda_s) + B_s + V_s + P_s, \tag{3.1}$$

where $B_s = \sum_{j \neq s} \omega_{sj}[f_{uu}(\lambda_j) - f_{uu}(\lambda_s)]$, $V_s = \sum_{j \neq s} \omega_{sj}[I_{uu}(\lambda_j) - f_{uu}(\lambda_j)]$, $P_s = \sum_{j \neq s} \omega_{sj} [\hat{I}_{uu}(\lambda_j) - I_{uu}(\lambda_j)]$. Asymptotic results for the bias and variance in spectral density estimation are well known (e.g. Hannan, 1970) and are stated in the following lemma for convenience.

Lemma 2.

$$B_t \sim -M^{-q} k_q f_q(\lambda_t),$$

$$\sqrt{m} V_t \overset{\mathrm{d}}{\sim} N\left(0, \frac{1}{2} \int_{-\infty}^{\infty} k(x)^2 \, \mathrm{d}x f_{uu}^2(\lambda_t)\right),$$

where q is the characteristic exponent of the kernel function defined as $\lim_{x \to 0} \{1 - k(x)\}/|x|^q = k_q < \infty$, $f_q(\lambda) = 1/2\pi \sum_{h=-\infty}^{\infty} |h|^q \gamma_u(h) \mathrm{e}^{-\mathrm{i}h\lambda}$, $\sim$ *denotes asymptotic equivalence and* $\overset{\mathrm{d}}{\sim}$ *denotes asymptotically distributed.*

Most commonly used kernel functions have the property that $q = 2$ in which case the bias term is of order $\mathrm{O}(M^{-2})$. For example, this applies for the smoothed periodogram estimate corresponding to the Daniell kernel with $k(x) = \sin(\pi x/2)/(\pi x/2)$, and spectral window

$$K(\lambda) = \begin{cases} \pi^{-1} & \text{for } |\lambda| \leqslant \pi/2, \\ 0 & \text{otherwise.} \end{cases}$$

Since our interest is primarily in the analysis of second order effects, in what follows we will simply give results based on the smoothed periodogram estimate for which $a = \frac{1}{2} \int_{-\infty}^{\infty} k(x)^2 \, \mathrm{d}x = 1$, and $\sqrt{m} V_t \overset{\mathrm{d}}{\sim} N(0, f_{uu}^2(\lambda_t))$. The use of other kernels then involves only minor scale adjustments in the formulae.

In spectral regression for stationary time series, the error term coming from preliminary estimation of $w_u(\lambda_s)$ is of smaller order of magnitude and can be dropped. However, different results arise in regression models with $I(1)$ regressors. In particular, the periodogram average of an $I(1)$ process around the zero frequency has a larger order of magnitude than at high frequencies. As a result, in the spectral density estimator $\hat{f}_{uu}(\lambda_s)$, the errors from preliminary estimation are amplified around the zero frequency and thus enter the second order effect. The order of magnitude for P_t is summarized in the following lemma.

Lemma 3.

$$P_t \sim -\frac{1}{2m\pi}\int \mathrm{d}B_u B_x' \left[\int B_x B_x'\right]^{-1} \int B_x \,\mathrm{d}B_u, \quad for\ |\lambda_t| \leqslant \pi/(2M)$$

and

$$P_t = \mathrm{o_p}(m^{-1}) \quad for\ |\lambda_t| > \pi/(2M).$$

In view of the $\sqrt{m}$ rate of convergence of the spectral estimates and the M^{-q} order of magnitude of the bias, the expansions we deal with will typically be in terms of powers of $m^{-1/2}$ and M^{-q}. In these expansions, terms that are of order $\mathrm{O_p}(m^{-1})$ or $\mathrm{O_p}(M^{-2q})$ turn out to be of principal interest and will be referred to as "higher order" terms.

An expansion for the inverse of the spectral density estimator, $\hat{f}_{uu}(\lambda_t)^{-1}$, can be obtained based on (3.1) and Lemmas 2 and 3. The results are given in Lemma 4 and readers are referred to Xiao and Phillips (1998) for details of the derivation.

Lemma 4.

$$\begin{aligned}\hat{f}_{uu}(\lambda_t)^{-1} = {} & f_{uu}(\lambda_t)^{-1} - \bar{f}_{uu}(\lambda_t)^{-2}V_t + \bar{f}_{uu}(\lambda_t)^{-3}V_t^2 - f_{uu}(\lambda_t)^{-2}B_t \\ & + f_{uu}(\lambda_t)^{-3}B_t^2 - \check{f}_{uu}(\lambda_t)^{-2}P_t + R_1(\lambda_t),\end{aligned} \tag{3.2}$$

where

$$\bar{f}_{uu}(\lambda_s) = m^{-1}\sum_{j\neq s} K(\lambda_j - \lambda_s) f_{uu}(\lambda_j), \quad \check{f}_{uu}(\lambda_s) = m^{-1}\sum_{j\neq s} K(\lambda_j - \lambda_s) I_{uu}(\lambda_j),$$

and

$$\begin{aligned}R_1(\lambda_t) = {} & -\bar{f}_{uu}(\lambda_t)^{-1} f_{uu}(\lambda_t)^{-3}B_t^3 - \check{f}_{uu}(\lambda_t)^{-1}\bar{f}_{uu}(\lambda_t)^{-3}V_t^3 \\ & + \hat{f}_{uu}(\lambda_t)^{-1}\check{f}_{uu}(\lambda_t)^{-2}P_t^2.\end{aligned} \tag{3.3}$$

By assumption, $f_{uu}(\lambda)$ is bounded away from zero and so, from Lemmas 2–4, we deduce that $R_1(\lambda_t) = \mathrm{o_p}(m^{-1} + M^{-2q})$.

The following lemma gives a useful limiting result for periodogram averages of $I(1)$ processes. Its derivation is based on a related result of Phillips (1991) and it will be used in the analysis of higher order asymptotics throughout later sections.

Lemma 5. *If the function* $g(\lambda)$ *has an absolutely and uniformly convergent Fourier representation* $g(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \varphi_k e^{ik\lambda}$, *then*

$$\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) g(\lambda_s) \xrightarrow{d} \frac{1}{2\pi} g(0) \int_0^1 B_x B_x'.$$

In particular, if $g(\lambda)$ *is a power function of* $f_{uu}(\lambda)$ *and its derivatives to order* q, *the above result holds.*

Remark 1. As in Phillips (1991), contributions from the zero-frequency ordinates dominate the limit behavior of these periodogram averages of $I(1)$ series and the limits are characterized in terms of quadratic functionals of the Brownian motion weak limit of $n^{-1/2} x_{[n\cdot]}$.

4. The Wald expansion in the univariate case

We start the development with the scalar case ($k = 1$) to simplify the derivations and illustrate the main results. The multivariate case will be treated later in Section 6. As mentioned in the introduction, here and elsewhere in the paper we proceed with formal moment approximations, rather than rigorous moment expansions.

In the univariate case, the linear hypothesis $R\beta = r$ reduces to H_0: $\beta = \beta_0$, and the corresponding regression Wald statistic is simply $W = (\hat{\beta} - \beta_0)^* \hat{\Sigma} (\hat{\beta} - \beta_0)$. Under the null hypothesis that $\beta = \beta_0$,

$$\hat{\beta} - \beta_0 = \left[\sum_s I_{xx}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1} \right]^{-1} \left[\sum_s I_{xu}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1} \right],$$

and

$$W = Z^* \hat{H}^{-1} Z, \tag{4.1}$$

where

$$\hat{H} = \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1}, \tag{4.2}$$

$$H = \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}, \tag{4.3}$$

$$Z = \frac{1}{n} \sum_s I_{xu}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1}. \tag{4.4}$$

By an application of Lemma 4, we decompose $D = \hat{H} - H$ and Z as follows:

$$\begin{aligned} D = &-\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \bar{f}(\lambda_s)^{-2} V_s + \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \bar{f}(\lambda_s)^{-3} V_s^2 \\ &- \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s + \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 \\ &- \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} P_s + R_3 \\ = &-D_{V1} + D_{V2} - D_{B1} + D_{B2} - D_{P1} + R_3 \end{aligned} \tag{4.5}$$

and

$$
\begin{aligned}
Z &= \frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} - \frac{1}{n}\sum_s I_{xu}(\lambda_s)\bar{f}(\lambda_s)^{-2}V_s \\
&\quad + \frac{1}{n}\sum_s I_{xu}(\lambda_s)\bar{f}_{uu}(\lambda_s)^{-3}V_s^2 - \frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2}B_s \\
&\quad + \frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-3}B_s^2 - \frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2}P_s + R_4 \\
&= Z_0 - Z_{V1} + Z_{V2} - Z_{B1} + Z_{B2} - Z_{P1} + R_4,
\end{aligned} \tag{4.6}
$$

where $R_3 = 1/n^2 \sum_s I_{xx}(\lambda_s) R_1(\lambda_s)$, $R_4 = 1/n \sum_s I_{xu}(\lambda_s) R_1(\lambda_s)$, and $R_1(\lambda_t) = o_p(m^{-1} + M^{-2q})$ is defined in Lemma 4.

The orders of magnitude of the terms in expansions (4.5) and (4.6) are analyzed in Lemma A1 in the appendix. Expanding $\hat{H}^{-1}$ around H^{-1} to the third term in a geometric expansion, substituting the truncated forms of D, Z, and $\hat{H}^{-1}$ into (4.1), and collecting terms up to $O_p(m^{-1} + M^{-q})$, we obtain a formal moment expansion for $E[W]$, which we summarize for convenience in the following formal result.

Theorem 1.

$$
\begin{aligned}
E[W] &= Q_0 + m^{-1}Q_1 + M^{-q}Q_2 + o(m^{-1} + M^{-q}) \\
&= \bar{W} + o(m^{-1} + M^{-q}),
\end{aligned}
$$

where Q_0, Q_1, and Q_2 are $O(1)$ *quantities defined as*

$$
\begin{aligned}
Q_0 &= E[H^{-1}Z_0^* Z_0], \\
Q_1 &= Q_{11} + Q_{12}, \\
Q_{11} &= m\{E[H^{-1}Z_{V1}^* Z_{V1}] + 2E[H^{-1}Z_0^* Z_{V2}] \\
&\quad + E[H^{-3}D_{V1}^2 Z_0^* Z_0] - E[H^{-2}D_{V2}Z_0^* Z_0]\}, \\
Q_{12} &= m\{E[H^{-1}D_{P1}Z_0^* Z_0] - 2E[H^{-1}Z_0^* Z_{P1}]\}, \\
Q_2 &= M^q\{E[H^{-2}D_{B1}Z_0^* Z_0] - 2E[H^{-1}Z_0^* Z_{B1}]\},
\end{aligned}
$$

and $\bar{W} = Q_0 + m^{-1}Q_1 + M^{-q}Q_2$ is the truncated expectation of the Wald statistic.

Remark 2. In this expansion, the leading term, Q_0, is just the expected value of the infeasible version of W in which the true spectral densities appear.

Remark 3. The second order effects (up to $O(m^{-1} + M^{-q})$) include: Q_{11}, a variance effect; Q_{12}, second order effect coming from the preliminary estimation; and Q_2, a bias effect. Here, the bias term Q_2 dominates the squared bias terms in order of magnitude and thus enters the second order effect.

Remark 4. Among the second order effects, the first term, $m^{-1}Q_1$, is a variance term and is always positive. The effect coming from preliminary estimation, $m^{-1}Q_{12}$, is asymptotically positive. The third term, $M^{-q}Q_2$, is a bias term that depends on the curvature of the spectral density function $f_{uu}(\omega)$. In fact, if we perform the expansion further to the order M^{-2q}, it can be shown that the $\mathrm{O}(M^{-2q})$ term is also positive. Thus, it is apparent from the expansion that, in most cases, the Wald test is likely to over-reject the null hypothesis.

When Q_2 is negative, we get two terms in the second order effect with different signs: $m^{-1}Q_1$ and $M^{-q}Q_2$. Whether or not the Wald test will over-reject H_0 depends on which of these terms dominates. In this case, a (second order) optimal bandwidth can be selected by minimizing the absolute value of the second order effects, that is, M can be chosen by equating $m^{-1}Q_1$ to $M^{-q}|Q_2|$, giving the optimal bandwidth

$$M = \left(\frac{|Q_2|}{2Q_1}\right)^{1/(q+1)} n^{1/(q+1)}.$$

Choosing the bandwidth by this formula enhances second order efficiency.

When Q_2 is positive, both terms in the second order effect of $\mathrm{E}[W]$ are positive. As a result, at least to the second order, the Wald statistic (2.2) tends to over-reject the null hypothesis. In this case, a bandwidth selection criterion can be defined by minimizing the expected value of the second order effect in the Wald statistic, i.e. M can be chosen to minimize $m^{-1}Q_1 + M^{-q}Q_2$, yielding

$$M = \left(\frac{qQ_2}{2Q_1}\right)^{1/(q+1)} n^{1/(q+1)}.$$

Choosing the bandwidth in such a way, the truncated expected Wald statistic is

$$Q_0 + n^{-q/(q+1)}[q^{1/(q+1)} + q^{-q/(q+1)}](2Q_1)^{q/(q+1)}Q_2^{1/(q+1)}.$$

A second order adjusted Wald statistic can also be constructed based on the above expansion. If $\hat{Q}_1$ and $\hat{Q}_2$ are consistent estimates of Q_1 and Q_2, we define the following second order modified Wald statistic:

$$W_M = W - n^{-q/(q+1)}\left[q^{1/(q+1)} + q^{-q/(q+1)}\right](2\hat{Q}_1)^{q/(q+1)}\hat{Q}_2^{1/(q+1)}.$$

This modified Wald statistic is asymptotically χ^2 and has an $O_{\mathrm{p}}(n^{-q/(q+1)})$ second order effect removed from its expected value.

The asymptotics of the second order terms can be analyzed using the results of Lemma 5 (see Lemma A.2 in the appendix for this analysis), leading us to the following approximation for the truncated expected Wald statistic. Again, we state these formal approximation results in summary form as a theorem.

Theorem 2.

$$\bar{W} \sim 1 + \frac{4}{m} + \frac{k_q}{M^q} f_{uu}(0)^{-1} f_q(0).$$

Since the effect at the zero frequency dominates, the second order effect in $\bar{W}$ is asymptotically determined by the spectral density of the residual process at the origin. As a result, the size of the test based on W will be largely affected by the curvature of the spectral density function $f_{uu}(\cdot)$ at the origin. When $f_q(0)$ is positive, a simple formula for bandwidth selection can be defined that is based on the unknown spectrum and its smoothness component, viz.

$$M = \left(\frac{q k_q f_{uu}(0)^{-1} f_q(0)}{8} \right)^{1/(q+1)} n^{1/(q+1)}. \tag{4.7}$$

If we plug consistent estimates $\hat{f}_{uu}(0)$ and $\hat{f}_q(0)$ into the formula, then a practical formula for bandwidth selection is obtained. A simple second order modified Wald test can also be constructed using this formula as follows:

$$W_{\mathrm{SO}} = W - n^{-q/(q+1)}[q^{1/(q+1)} + q^{-q/(q+1)}](8)^{q/(q+1)}[k_q \hat{f}_q(0)/\hat{f}_{uu}(0)]^{1/(q+1)}. \tag{4.8}$$

In the extreme case that $f_q(0) = 0$, the leading bias term $M^{-q} k_q f_{uu}(0)^{-1} f_q(0)$ is zero, indicating that the optimal m will be of larger order of magnitude. In this case, in principle, further expansion to higher order is needed. In the special case where u_t are i.i.d., we choose $m = n/2$ and thus the OLS estimator is optimal.

Remark 5. If x_t is an $I(0)$ process, the higher order asymptotics are different. It can be shown in this case that

$$\mathrm{E}[W] \sim 1 + \frac{2}{m} + \frac{k_q}{M^q} \Omega^{-1} \Gamma,$$

where

$$\Omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{xx}(\omega) f_{uu}(\omega)^{-1} \, \mathrm{d}\omega,$$

$$\Gamma = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{xx}(\omega) f_{uu}(\omega)^{-2} f_q(\omega) \, \mathrm{d}\omega,$$

and the corresponding bandwidth formula based on minimization of the second order term in the expected Wald statistic is

$$M = \left(\frac{q k_q \Omega^{-1} \Gamma}{4} \right)^{1/(q+1)} n^{1/(q+1)}.$$

Such a formula depends on integrated forms of the spectral density functions, rather than their values at a single point. As a result, we cannot simply substitute consistent nonparametric density estimates for a single point into the formula to obtain a bandwidth choice rule. Instead, we have to estimate the spectral density at points over the interval $[-\pi, \pi]$ and approximate Ω, Γ, by Riemann sum discrete approximations. Another way to find a bandwidth in this case is to prespecify a parametric model for the error process u_t, estimate the parameters and utilize these estimates in a plug-in procedure to obtain an estimator of the optimal bandwidth. The bandwidth rule is then optimal for this parametric model, and has the correct order of magnitude even when the parametric model is misspecified.

5. Band spectral regression

One of the advantages of frequency domain regression is that it provides a selective procedure which allows us to focus on the most relevant frequencies. In economic applications, it is possible that a model applies more accurately at some frequencies than others, a famous example being the permanent income hypothesis model where permanent consumption is modeled as a linear function of permanent income. In this model, the marginal propensity to consume out of permanent income can well be very different from that out of transitory income. Since it is reasonable to associate permanent income with zero- and low-frequency components in the data, regressions based on high-frequency components can be expected to deliver different estimates of the marginal propensity to consume than those from low frequency components.

The idea of band spectral regression has a long history. Hannan (1963b) proposed averaging over a nondegenerate subset of frequency in regression and Engle (1974) applied it to test the permanent income hypothesis, naming the procedure band spectral regression. Later, Corbae et al. (1994) extended it to cases with stochastic trends, and Corbae et al. (2002) studied the effects of deterministic detrending in the time and frequency domains.

In what follows, we consider for model (1.2) a simple band spectral regression based on averaging over a frequency band B which contains a subset of fundamental frequencies and is symmetric about the origin. As we have seen in previous sections, due to the dominating effects of the zero frequency components in the nonstationary regressor case, estimators based on a frequency band that includes the zero frequency will have similar asymptotic properties to those of $\hat{\beta}$ in the previous sections. Thus, the following discussion considers a band B that concentrates on some non-zero frequencies and we focus on the nonstationary case. In particular, for some frequency $\omega > 0$, we set $B = B_\omega \cup B_{-\omega}$ with

$$B_\omega = \left[\omega - \frac{\pi}{2L}, \omega + \frac{\pi}{2L}\right],$$

where $L/n + 1/L \to 0$ as $n \to \infty$, and define

$$\ell = \#\{\lambda_s \in B_\omega\}.$$

The band B then has 2ℓ fundamental frequencies with $\ell \to \infty$ as $n \to \infty$, while the band itself shrinks to the symmetric frequencies ω and $-\omega$. We will assume that $\ell/n^{1/4} \to 0$. The band B spectral regression estimator is

$$\hat{\beta}_B = \left[\sum_{\lambda_s \in B} I_{xx}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\sum_{\lambda_s \in B} I_{xy}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}\right],$$

so that $\sqrt{2\ell}(\hat{\beta}_B - \beta) = \hat{H}_B^{-1}\hat{\Psi}_B$, where

$$\hat{H}_B = \frac{1}{2\ell}\sum_{\lambda_s \in B} I_{xx}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1} \quad \text{and} \quad \hat{\Psi}_B = \frac{1}{\sqrt{2\ell}}\sum_{\lambda_s \in B} I_{xu}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}.$$

Under some regularity conditions, notably a linear process condition on u_t (see Assumption A. In the appendix), we have the following first order result for $\hat{\beta}_B$.

Lemma 6.

$$\sqrt{2\ell}(\hat{\beta}_B - \beta) \xrightarrow{\mathrm{d}} MN(0, h_x(\omega)^{-1} f_{uu}(\omega)), \tag{5.1}$$

where

$$h_x(\omega) = f_{xx}(\omega) + \frac{1}{2\pi}\frac{B_x(1)^2}{|1-\mathrm{e}^{\mathrm{i}\omega}|^2} = \frac{f_{vv}(\omega)}{|1-\mathrm{e}^{\mathrm{i}\omega}|^2} + \frac{1}{2\pi}\frac{B_x(1)^2}{|1-\mathrm{e}^{\mathrm{i}\omega}|^2}.$$

Here the mixing variate in (5.1) involves $B_x(1)^2$ and appears via $h_x(\omega)$. The Wald statistic for H_0 ($\beta = \beta_0$) in this band spectral regression can be written as $W_B = \hat{\Psi}_B^* \hat{H}_B^{-1} \hat{\Psi}_B$. The expansion for W_B follows steps similar to those for W in Section 4. However, since the band B concentrates on the non-zero frequencies ω and $-\omega$, the periodogram averages have different orders of magnitude and need to be standardized differently. Lemma A.3 in the appendix gives orders of magnitude for some of the higher order components and is useful in the formal moment expansion for the band spectral regression Wald statistic that follows.

Theorem 3.

$$\begin{aligned} \mathrm{E}[W_B] &= \Phi_0 + m^{-1}\Phi_1 + M^{-q}\Phi_2 + \mathrm{o}(m^{-1} + M^{-q}) \\ &= \bar{W}_B + \mathrm{o}(m^{-1} + M^{-q}), \end{aligned}$$

where Φ_0, Φ_1, *and* Φ_2 *are* $\mathrm{O}(1)$ *quantities defined as*

$$\Phi_0 = \mathrm{E}[H_B^{-1}\Psi_0^*\Psi_0],$$

$$\begin{aligned} \Phi_1 = m\{&\mathrm{E}[H_B^{-1}\Psi_{V1}^*\Psi_{V1}] + 2\mathrm{E}[H_B^{-1}\Psi_0^*\Psi_{V2}] \\ &+ \mathrm{E}[H_B^{-3}\Delta_{V1}^2\Psi_0^*\Psi_0] - \mathrm{E}[H_B^{-2}\Delta_{V2}\Psi_0^*\Psi_0]\}, \end{aligned}$$

$$\Phi_2 = M^q\{\mathrm{E}[H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_0] - 2\mathrm{E}[H_B^{-1}\Psi_0^*\Psi_{B1}]\},$$

$H_B = (1/2\ell)\sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}$, $\Psi_0 = (1/\sqrt{2\ell})\sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1}$, *and* Ψ_{V1}, Ψ_{V2}, Ψ_{B1}, Δ_{B1}, Δ_{V1}, *and* Δ_{V2} *are quantities defined in Lemma A.3 in the appendix, and* $\bar{W}_B = \Phi_0 + m^{-1}\Phi_1 + M^{-q}\Phi_2$ *is the truncated expected Wald statistic.*

Remark 6. This expansion is similar to that in Section 4. However, given the narrow band spectral regression estimator that we consider in this section, the quantities in the formula have different definitions and different asymptotics. As discussed earlier, for testing (2.2) using the regression Wald test, a bandwidth selection criterion can be defined by minimizing the expected value of the second order effect in the Wald Statistic, giving the following bandwidth choice:

$$M = \left(\frac{q\Phi_2}{2\Phi_1}\right)^{1/(q+1)} n^{1/(q+1)}.$$

Some asymptotic results for the quantities in Theorem 3 are given in Lemma A.4 in the appendix. These results deliver the following explicit approximation for the truncated expected Wald statistic $\bar{W}_B$,

$$\bar{W}_B \sim 1 + \frac{3}{m} + \frac{k_q}{M^q} f_{uu}(\omega)^{-1} f_q(\omega).$$

Remark 7. Since we are considering a narrow band spectral regression, the higher order terms converge to quantities evaluated at the corresponding frequency ω. Consequently, the second order effect in $\bar{W}_B$ is asymptotically determined by the spectral density of the residual process at frequency ω. In particular, similar to the results in Section 4, the size of the test based on W_B will be largely affected by the curvature of the spectral density function $f_{uu}(\cdot)$ at ω. Thus, a formula for bandwidth selection and a simple second order modified Wald test can be obtained in a similar way as (4.7) and (4.8) based on consistent estimates of $f_{uu}(\omega)$ and $f_q(\omega)$.

6. The multivariate case

This section considers the multivariate case and the Wald test for a general linear hypothesis with p restrictions. Thus, the null hypothesis has the general format H_0: $R\beta = r$, where r is a $p \times 1$ vector and R is a $p \times k$ matrix. The quantities x_t and β are now $k \times 1$ vectors, $I_{xx}(\lambda)$ and $f_{xx}(\lambda)$ are $k \times k$ matrices, and I_{xy}, I_{xu} are $k \times 1$ vectors. The corresponding regression Wald statistic is given by (2.3). Under the null hypothesis that $R\beta = r$, we have

$$W = (\hat{\beta} - \beta)' R' [R\hat{\Sigma}^{-1} R']^{-1} R(\hat{\beta} - \beta).$$

For convenience, we continue to use $\hat{H}, H$, and Z, which are defined by formulae (4.2) –(4.4) in the scalar case, for their matrix counterparts. Then, the Wald statistic can be written as

$$W = Z^* \hat{H}^{-1} R' [R\hat{H}^{-1} R']^{-1} R\hat{H}^{-1} Z.$$

The expansion for W follows lines similar to those of Section 4, and the higher order asymptotics are similar to those in Theorems 1 and 2. However, the expansion is now complicated by the multivariate nature of the model. Detailed expansions are given in the appendix and we only state the major steps of the expansions here. Denoting $R\hat{H}^{-1}R'$ by $\hat{J}_0$, and $RH^{-1}R'$ by J_0, we expand $\hat{J}_0^{-1}$ around J_0 to the third order and get the following expansion for W,

$$Z^* \hat{H}^{-1} R' [J_0^{-1} - J_0^{-1}(\hat{J}_0 - J_0) J_0^{-1} + J_0^{-1}(\hat{J}_0 - J_0) J_0^{-1} (\hat{J}_0 - J_0) J_0^{-1} - R_5]^{-1} \\ R\hat{H}^{-1} Z, \tag{6.1}$$

where $R_5 = \hat{J}_0^{-1}(\hat{J}_0 - J_0)J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1}$. Notice that $\hat{J}_0 - J_0 = R(\hat{H}^{-1} - H^{-1})R'$, $\hat{H}^{-1}$ can be expanded around H as $\hat{H}^{-1} = H^{-1} - H^{-1}DH^{-1} + H^{-1}DH^{-1}DH^{-1} - R_6$, where $R_6 = \hat{H}^{-1}DH^{-1}DH^{-1}DH^{-1}$, and D and Z follow similar expansions as (4.6), (4.7), substituting all these preliminary expansions into (6.1) and collecting terms up to $\mathrm{O_p}(m^{-1} + M^{-q})$, we get the following formal moment approximation for the Wald statistic.

Theorem 4.

$$\begin{aligned} \mathrm{E}[W] &= S_0 + m^{-1}S_1 + M^{-q}S_2 + \mathrm{o_p}(m^{-1} + M^{-q}) \\ &= \bar{W}_G + \mathrm{o_p}(m^{-1} + M^{-q}), \end{aligned}$$

where S_0, S_1 *and* S_2 *are* $\mathrm{O}(1)$ *quantities defined as*

$$S_0 = \mathrm{E}[Z_0^* H^{-1}R'(RH^{-1}R')^{-1}RH^{-1}Z_0],$$

$$\begin{aligned} S_1 = m\{ &\mathrm{E}[Z_{V1}^* H^{-1}R'J_0^{-1}RH^{-1}Z_{V1}] + \mathrm{E}[Z_0^* H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0] \\ &- \mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0] \\ &+ \mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0] \\ &- 2\mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0] \\ &+ 2\mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}Z_0] \\ &+ \mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V2}H^{-1}R'J_0^{-1}RH^{-1}Z_0] \\ &+ 2\mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}Z_{V2}] - 2\mathrm{E}[Z_{P1}^* H^{-1}R'J_0^{-1}RH^{-1}Z_0] \\ &+ 2\mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{P1}H^{-1}Z_0] - 2E[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{V2}H^{-1}Z_0] \\ &- \mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{P1}H^{-1}R'J_0^{-1}RH^{-1}Z_0]\}, \end{aligned}$$

$$\begin{aligned} S_2 = M^q\{ &2\mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{B1}H^{-1}Z_0] - 2E[Z_{B1}^* H^{-1}R'J_0^{-1}RH^{-1}Z_0] \\ &- \mathrm{E}[Z_0^* H^{-1}R'J_0^{-1}RH^{-1}D_{B1}H^{-1}R'J_0^{-1}RH^{-1}Z_0]\}, \end{aligned}$$

and $\bar{W}_G = S_0 + m^{-1}S_1 + M^{-q}S_2$ *is the truncated expected Wald statistic.*

Remark 8. The orders of magnitude of the second order terms are the same as those in the univariate case. However, these second order effects do depend on the linear hypothesis. The number of restrictions, p, in the null hypothesis appears in both the first and the second order asymptotics. Lemma A.5 in the appendix gives asymptotic results for the higher order terms in the expansion. As a result, the truncated expected Wald statistic $\bar{W}_G$ has the following approximation, similar to that given in Theorem 2:

$$\bar{W}_G \sim p + \frac{4p}{m} + \frac{pk_q}{M^q} f_{uu}^{-1}(0)f_q(0).$$

Remark 9. Just as in previous sections, an optimal bandwidth rule can be determined by minimizing the second order term, giving formula (4.8), and a second order modified Wald test can then be constructed as in (4.9).

Remark 10. When x_t is a stationary process, the zero frequency no longer has a dominating effect and the value of spectral density elsewhere appears in the asymptotics. As a result, certain terms cannot be cancelled out and the higher order asymptotics of the Wald test have a more complicated representation than in the univariate case. We have been able to show that the Wald statistic in this case has the following moment approximation:

$$\mathrm{E}[W] \sim p + \frac{2p}{m} + \frac{k_q}{M^q} tr\{(R\Omega^{-1}R')^{-1}R\Omega^{-1}\Gamma\Omega^{-1}R'\},$$

where Ω and Γ are $k \times k$ matrices of integrals of spectral densities given by the same formulae as those in Remark 5.

7. Monte Carlo results

We conducted a small Monte Carlo experiment to evaluate the second order theory for Wald tests in cointegrating regressions of the type considered here. In particular, we examined the size properties of Wald statistics using different bandwidth choices and the second order adjusted Wald test, and looked at the effect of the second order bandwidth selection criterion on test size.

The data were generated according to the model

$$y_t = \beta x_t + u_t,$$
$$x_t = x_{t-1} + v_t,$$
$$u_t = \alpha u_{t-1} + \varepsilon_t,$$

where v_t and ε_t are iid $\mathrm{N}(0,\sigma_v^2)$ and $\mathrm{N}(0,\sigma_\varepsilon^2)$ variates and are independent of each other. Three values of autoregressive coefficient, $\alpha = 0.3, 0.6$, and 0.8 were considered. The null hypothesis H_0: $\beta = 1$ was tested. The corresponding Wald test is asymptotically χ_1^2 and the nominal 5% level critical value for the test based on the asymptotic distribution is 3.84.

Three sample sizes ($n = 64$, 128, 256) were used in the experiment, the number of replications was 10,000, and the Daniell window was used in the density estimation. The size properties of four Wald statistics are examined in the Monte Carlo analysis. They are as follows:

W_0: Wald statistic (2.3) using the optimal bandwidth formula (4.8).
W_1: Wald statistic (2.3) using a fixed bandwidth choice $M = 8$.
W_2: Wald statistic (2.3) using a fixed bandwidth choice $M = 10$.
W_{SO}: second order adjusted Wald statistic(4.9) using the optimal bandwidth.

Table 2
Empirical size (nominal size = 5%)

	n	W_0	W_1	W_2	W_{SO}
$\alpha = 0.3$	64	0.1466	0.1587	0.1598	0.1056
$\alpha = 0.3$	128	0.1102	0.1213	0.1215	0.0825
$\alpha = 0.3$	256	0.0806	0.0849	0.0855	0.0646
$\alpha = 0.6$	64	0.1787	0.2005	0.2017	0.1346
$\alpha = 0.6$	128	0.1315	0.1445	0.1458	0.1088
$\alpha = 0.6$	256	0.0903	0.1041	0.1047	0.0785
$\alpha = 0.8$	64	0.2068	0.2348	0.2360	0.1634
$\alpha = 0.8$	128	0.1529	0.1720	0.1728	0.1226
$\alpha = 0.8$	256	0.1033	0.1232	0.1235	0.0902

Table 3
Empirical critical values

	n	W_0	W_1	W_2	W_{SO}
$\alpha = 0.3$	64	7.4006	8.0124	8.0596	6.4001
$\alpha = 0.3$	128	5.8121	6.2244	6.2279	5.1852
$\alpha = 0.3$	256	4.7062	4.9326	4.9488	4.3536
$\alpha = 0.6$	64	9.5463	10.347	10.349	8.4762
$\alpha = 0.6$	128	6.8556	7.3700	7.4274	6.2219
$\alpha = 0.6$	256	5.2824	5.5971	5.5990	4.8546
$\alpha = 0.8$	64	11.811	12.475	12.548	10.835
$\alpha = 0.8$	128	7.3028	8.2632	8.1636	6.9366
$\alpha = 0.8$	256	6.0802	6.6325	6.6416	5.7716

Different combinations of σ_v^2 and σ_ε^2 were tried. Although in cases where x_t is endogenous the size distortion is known to be larger for small ratios of $\sigma_v^2/\sigma_\varepsilon^2$, with exogenous regressors, the effect of $\sigma_v^2/\sigma_\varepsilon^2$ on the empirical size of the Wald statistic in our model is very small and thus we report results only for $\sigma_v^2 = \sigma_\varepsilon^2 = 1$ here. The empirical size of the tests from Monte Carlo rejection rates based on the critical value for a nominal asymptotic size of 5% are reported in Table 2. The empirical critical values of the tests from quantiles in the simulations under the null hypothesis are given in Table 3.

In these experiments, use of the optimal bandwidth selection criterion improves the finite sample size performance of the Wald test. Although size distortion still exists for all of these tests, the empirical size of the Wald statistic, W_0, which uses the optimal bandwidth, is smaller than that of the Wald statistics (W_1 and W_2) that use arbitrary bandwidth parameters. The empirical critical values of W_0 are also closer to the nominal asymptotic critical values than those of W_1 and W_2. Size distortion is further reduced in the second order adjusted Wald statistic. As the sample size increases from 64 to 256, the empirical size decreases and the empirical critical values move closer to the nominal values. However, there is noticeably more size distortion in all cases as the autocorrelation in the residual process u_t increases.

8. Conclusion

This paper develops higher order expansions for Wald test statistics in efficient, frequency domain semi-parametric regression models with integrated regressors. These expansions address some of the problems presented by first order asymptotic theory. In particular, they provide an opportunity to reduce size distortion in statistical testing by suitable bandwidth choices and second order adjustments. Since the effect at the zero frequency dominates the asymptotics for periodogram averages of $I(1)$ processes, the second order effect in the expected value of the Wald statistic is asymptotically determined by the spectral density of the residual process at the origin. As a result, the size properties of Wald tests are largely affected by the curvature of the spectral density function at the origin.

A bandwidth selection criterion that is based on minimizing the second order effect on the expected value of the Wald statistic is proposed to improve statistical testing and second order modified Wald tests are proposed that use consistent estimates of the second order terms. Expansions for the Wald test are also developed for a narrow band spectral regression estimator. When the frequency band under consideration concentrates to some non-zero frequency, the estimator of the cointegrating vector converges to its true value at a different rate and, the second order effects in the expected value of the Wald statistics are functions of the spectral density at the corresponding frequency. Interestingly, some of the effects of the nonstationarity that dominates low frequency limit behavior also carry over to high frequency asymptotics in band spectral regression models.

Acknowledgements

We thank Peter Robinson, Roger Koenker and two referees for helpful comments on earlier versions of this paper. Phillips thanks the NSF for research support under Grant Nos. SBR94-22922 and SBR97-30295 and Xiao thanks the Hong Kong Research Grants Council for research support under Grant No. CUHK4078/98 and the Cowles Foundation for hospitality during a visit in May–June 1998.

Appendix A

This appendix provides derivations of results given in the body of the paper. Second order terms are often only sketched and rigorous derivations are not always given. Further specific assumptions may also be needed in some cases for rigorous results for second order terms to apply. For notational convenience, we denote $A \sim B$ if $A = B(1 + o_p(1))$.

A.1. Proof of Lemma 1

Notice that, by a geometric expansion on $\hat{f}_{uu}(\lambda_s)^{-1}$,

$$\begin{aligned}
&n^{-1}\sum_s I_{xu}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}\\
&\quad= n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1} + \sum_{k=2}^{p}(-1)^{k-1}n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-k}\\
&\qquad(\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^{k-1}\\
&\qquad+(-1)^p n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-p}\hat{f}_{uu}(\lambda_s)^{-1}(\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^p.
\end{aligned}$$

We show that the leading term, $n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}$, converges weakly to $[2\pi f_{uu}(0)]^{-1}\int B_x\,\mathrm{d}B_u$ and other terms are asymptotically zeros. Using the Fourier series representation of $f_{uu}(\lambda_s)^{-1}$ (see, e.g., Naimark, 1960),

$$f_{uu}(\lambda_s)^{-1} = \frac{1}{2\pi}\sum_{g=-\infty}^{\infty} D_g \mathrm{e}^{\mathrm{i}g\lambda_s},$$

we have

$$\begin{aligned}
&n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\\
&\quad= \frac{1}{n}\sum_s\left[\frac{1}{2\pi}\sum_r C_{xu}(r)\mathrm{e}^{-\mathrm{i}\lambda_s r}\right]\left[\frac{1}{2\pi}\sum_{g=-\infty}^{\infty} D_g \mathrm{e}^{\mathrm{i}g\lambda_s}\right]\\
&\quad= \frac{1}{n}\left(\frac{1}{2\pi}\right)^2\sum_{g=-\infty}^{\infty} D_g\sum_r C_{xu}(r)\sum_s \mathrm{e}^{\mathrm{i}\lambda_s(g-r)}.
\end{aligned}$$

Since $\sum_s \mathrm{e}^{\mathrm{i}\lambda_s(g-r)} = 0$ unless $g - r = l * n$, for $l = 0, \pm1, \pm2, \ldots$, we have

$$\begin{aligned}
&n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\\
&\quad= \left(\frac{1}{2\pi}\right)^2\sum_r C_{xu}(r)\sum_{|l|=0}^{\infty} D_{r+l*n}\\
&\quad= \left(\frac{1}{2\pi}\right)^2\sum_{|r|\leqslant M} C_{xu}(r)\sum_{|l|=0}^{\infty} D_{r+l*n} + \left(\frac{1}{2\pi}\right)^2\sum_{|r|>M} C_{xu}(r)\sum_{|l|=0}^{\infty} D_{r+l*n}.
\end{aligned}$$

Following a similar argument to that of Phillips (1991) and by summability of D_r, we have

$$\left(\frac{1}{2\pi}\right)^2 \sum_{|r|\leqslant M} C_{xu}(r) \sum_{|l|=0}^{\infty} D_{r+l*n} \Rightarrow \frac{1}{2\pi} \int B_x \, \mathrm{d}B_u \left(\frac{1}{2\pi} \sum_{r=-\infty}^{\infty} D_r\right)$$

$$=[2\pi f_{uu}(0)]^{-1} \int B_x \, \mathrm{d}B_u,$$

$$\left(\frac{1}{2\pi}\right)^2 \sum_{|r|>M} C_{xu}(r) \sum_{|l|=0}^{\infty} D_{r+l*n} \to 0.$$

Thus

$$n^{-1}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} \Rightarrow [2\pi f_{uu}(0)]^{-1} \int B_x \, \mathrm{d}B_u. \tag{9.1}$$

Notice that since $f_{uu}(\lambda_s)$ is bounded away from zero and $\hat{f}_{uu}(\lambda_s)$ is a uniformly consistent estimator,

$$\left| n^{-1}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-p} \hat{f}_{uu}(\lambda_s)^{-1} (\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^p \right|$$

$$\leqslant c_1 n^{-1} \sum_s |I_{xu}(\lambda_s)| \, |f_{uu}(\lambda_s)^{-p-1} (\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^p|$$

for some c_1, and

$$|\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s)|^p \leqslant c_* \{|B_s|^p + |V_s|^p + |P_s|^p\}$$

for some constant c_*. Under some regularity conditions, we can obtain uniform rates of convergence for B_s, V_s, and P_s. For B_s and V_s, by standard results in spectral density estimation (see, e.g. Brillinger, 1980; Theorem 7.7.4), we have

$$\sup_t |B_t| = \mathrm{O}(M^{-q})$$

$$\sup_t |V_t| = \mathrm{O}_p(m^{-1/2+\varepsilon}) \quad \text{for any } \varepsilon > 0.$$

For P_s, under the assumption that v_t, u_t, and x_0 all have uniformly bounded moments, it can be verified that $C_{xu}(h)$ and $n^{-1}C_{xx}(h)$ are uniformly $\mathrm{O_p}(1)$ quantities. Thus, the uniform rate of the term P_t is obtained since

$$P_t = -2(\hat{\beta}_{\mathrm{OLS}} - \beta)\mathrm{Re}(\hat{f}_{xu}(\lambda_t)) + (\hat{\beta}_{\mathrm{OLS}} - \beta)' \hat{f}_{xx}(\lambda_t)(\hat{\beta}_{\mathrm{OLS}} - \beta),$$

$$n(\hat{\beta}_{\mathrm{OLS}} - \beta) = \mathrm{O_p}(1),$$

where $\hat{\beta}_{\mathrm{OLS}}$ is the OLS estimator, and

$$\hat{f}_{xu}(\lambda_t) = \frac{1}{2\pi} \sum_h k\left(\frac{h}{M}\right) C_{xu}(h) \mathrm{e}^{\mathrm{i}\lambda_t h},$$

$$\hat{f}_{xx}(\lambda_t) = \frac{1}{2\pi} \sum_h k\left(\frac{h}{M}\right) C_{xx}(h) \mathrm{e}^{\mathrm{i}\lambda_t h}.$$

We have

$$\sup|\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s)|^p = \mathrm{O_p}(M^{-pq} + m^{-p/2+\varepsilon}).$$

Giving the bandwidth choice and for $p > 2$,

$$\left| n^{-1}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-p} \hat{f}_{uu}(\lambda_s)^{-1}(\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^p \right| = \mathrm{o_p}(1).$$

Now we show that terms like $n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}B_s$, $n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s$ and $n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}P_s$, etc., are $\mathrm{o_p}(1)$. By definition, $V_s = \sum_j \omega_{sj}[I_{uu}(\lambda_j) - f_{uu}(\lambda_j)] = m^{-1}\sum_{\lambda_j \in B(\lambda_s)} K(\lambda_j - \lambda_s)[I_{uu}(\lambda_j) - f_{uu}(\lambda_j)]$, and using the Fourier series representation of $I_{xu}(\lambda_s)$, $I_{uu}(\lambda_j)$, $f_{uu}(\lambda_j)$, and $f_{uu}(\lambda_s)^{-2}$, denoting the sample autocorrelation of u_t and the autocorrelation function of u_t as $C_{uu}(h)$ and $\gamma_{uu}(k)$ respectively, we have

$$\begin{aligned}
&n^{-1}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s \\
&\quad = n^{-1}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} \sum_j \omega_{sj}[I_{uu}(\lambda_j) - f_{uu}(\lambda_j)] \\
&\quad = n^{-1}\sum_s \left[\frac{1}{2\pi}\sum_r C_{xu}(r)\mathrm{e}^{-\mathrm{i}\lambda_s r}\right]\left[\frac{1}{2\pi}\sum_{g=-\infty}^{\infty} B_g \mathrm{e}^{\mathrm{i}g\lambda_s}\right] \\
&\qquad \times \sum_j \omega_{sj}\left[\frac{1}{2\pi}\sum_h C_{uu}(h)\mathrm{e}^{-\mathrm{i}\lambda_j h} - \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_{uu}(k)\mathrm{e}^{-\mathrm{i}\lambda_j k}\right] \\
&\quad = n^{-1}\sum_s \left[\frac{1}{2\pi}\sum_r C_{xu}(r)\mathrm{e}^{-\mathrm{i}\lambda_s r}\right]\left[\frac{1}{2\pi}\sum_{g=-\infty}^{\infty} B_g \mathrm{e}^{\mathrm{i}g\lambda_s}\right] \\
&\qquad \times \frac{1}{2\pi}\sum_j \omega_{sj}\sum_h [C_{uu}(h) - \gamma_{uu}(h)]\mathrm{e}^{-\mathrm{i}\lambda_j h} \\
&\qquad - n^{-1}\sum_s \left[\frac{1}{2\pi}\sum_r C_{xu}(r)\mathrm{e}^{-\mathrm{i}\lambda_s r}\right]\left[\frac{1}{2\pi}\sum_{g=-\infty}^{\infty} B_g \mathrm{e}^{\mathrm{i}g\lambda_s}\right] \\
&\qquad \times \sum_j \omega_{sj}\left[\frac{1}{2\pi}\sum_{|k|\geqslant n}\gamma_{uu}(k)\mathrm{e}^{-\mathrm{i}\lambda_j k}\right] \\
&\quad = n^{-1}\left(\frac{1}{2\pi}\right)^3 \sum_r \sum_h \sum_{g=-\infty}^{\infty} B_g[C_{uu}(h) - \gamma_{uu}(h)]C_{xu}(r)\sum_s\sum_j \omega_{sj}\mathrm{e}^{\mathrm{i}(g-r)\lambda_s}\mathrm{e}^{-\mathrm{i}\lambda_j h} \\
&\qquad - n^{-1}\left(\frac{1}{2\pi}\right)^3 \sum_r \sum_{|k|\geqslant n} \sum_{g=-\infty}^{\infty} B_g \gamma_{uu}(k) C_{xu}(r)\sum_s\sum_j \omega_{sj}\mathrm{e}^{\mathrm{i}(g-r)\lambda_s}\mathrm{e}^{-\mathrm{i}\lambda_j k}.
\end{aligned}$$

where $(1/2\pi)\sum_{g=-\infty}^{\infty} B_g \mathrm{e}^{\mathrm{i}g\lambda_s}$ is the Fourier series representation of $f_{uu}(\lambda_s)^{-2}$. Notice that

$$\lambda_j \in B(\lambda_s) = \left\{\lambda_j\colon \lambda_s - \frac{\pi}{2M} < \lambda_j \leqslant \lambda_s + \frac{\pi}{2M}\right\}$$
$$= \left\{\lambda_s - \frac{\pi}{2M} + 2\pi t/n\colon t = 1, 2, \ldots, m = n/2M\right\}$$

if we use the Daniell's spectral window,

$$\sum_s \sum_j \omega_{sj} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}\lambda_j h}$$
$$= m^{-1} \sum_{s=1}^{n} \sum_{\lambda_j \in B(\lambda_s)} K(\lambda_j - \lambda_s) \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}\lambda_j h}$$
$$= (\pi m)^{-1} \sum_{s=1}^{n} \sum_{\lambda_j \in B(\lambda_s)} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}\lambda_j h}$$
$$= (\pi m)^{-1} \sum_{s=1}^{n} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \sum_{t=1}^{m} \mathrm{e}^{-\mathrm{i}(\lambda_s - \pi/2M + 2\pi t/n)h}$$
$$= (\pi m)^{-1} \sum_{s=1}^{n} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}(\lambda_s - \pi/2M)h} \sum_{t=1}^{m} \mathrm{e}^{-2\mathrm{i}\pi th/n}$$
$$= \mathrm{e}^{\mathrm{i}h\pi/2M} \left[\frac{1}{\pi m} \sum_{t=1}^{m} \mathrm{e}^{-2\mathrm{i}\pi th/n}\right] \left[\sum_{s=1}^{n} \mathrm{e}^{\mathrm{i}(g-r-h)\lambda_s}\right].$$

Since $\sum_s \mathrm{e}^{\mathrm{i}(g-r-h)\lambda_s} = 0$ unless $g - r - h = l * n$ (when $\sum_s \mathrm{e}^{\mathrm{i}(g-r-h)\lambda_s} = n$) for $l = 0, \pm 1, \pm 2, \ldots$, we have

$$n^{-1} \left(\frac{1}{2\pi}\right)^3 \sum_r \sum_h \sum_{g=-\infty}^{\infty} B_g [C_{uu}(h) - \gamma_{uu}(h)] C_{xu}(r) \sum_s \sum_j \omega_{sj} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}\lambda_j h}$$
$$= \left(\frac{1}{2\pi}\right)^3 \sum_h [C_{uu}(h) - \gamma_{uu}(h)] \mathrm{e}^{\mathrm{i}h\pi/2M} \left[\frac{1}{\pi m} \sum_{t=1}^{m} \mathrm{e}^{-2\mathrm{i}\pi th/n}\right]$$
$$\left[\sum_r \sum_{|l|=0}^{\infty} B_{r+h+l*n} C_{xu}(r)\right]$$
$$= \mathrm{o}_{\mathrm{p}}(1),$$

since

$$\frac{1}{\pi m} \sum_{t=1}^{m} \mathrm{e}^{-2\mathrm{i}\pi th/n}, \sum_r \sum_{|l|=0}^{\infty} B_{r+h+l*n} C_{xu}(r) = \mathrm{O}_{\mathrm{p}}(1) \quad \text{for all } h.$$

Similarly, for the second term,

$$n^{-1}\left(\frac{1}{2\pi}\right)^3 \sum_r \sum_{|k|\geqslant n} \sum_{g=-\infty}^{\infty} B_g \gamma_{uu}(k) C_{xu}(r) \sum_s \sum_j \omega_{sj} \mathrm{e}^{\mathrm{i}(g-r)\lambda_s} \mathrm{e}^{-\mathrm{i}\lambda_j k}$$

$$= \left(\frac{1}{2\pi}\right)^3 \sum_{|k|>n} \gamma_{uu}(k) \mathrm{e}^{\mathrm{i}k\pi/2M} \left[\frac{1}{\pi m}\sum_{t=1}^{m} \mathrm{e}^{-2\mathrm{i}\pi tk/n}\right] \left[\sum_r \sum_{|l|=0}^{\infty} B_{r+k+l*n} C_{xu}(r)\right]$$

$$= \mathrm{o_p}(1).$$

The proofs for other terms are similar to that of (9.1). Thus we have

$$n^{-1}\sum_s I_{xu}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1} \Rightarrow [2\pi f_{uu}(0)]^{-1} \int B_x \, \mathrm{d}B_u.$$

In a related fashion and as in Phillips (1991), we can show

$$n^{-2}\sum_s I_{xx}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1} \xrightarrow{\mathrm{d}} [2\pi f_{uu}(0)]^{-1} \int B_x B_x'.$$

A.2. Proof of Lemma 2

The results for B_s and V_s come from standard analysis of the bias and variance of kernel spectral density estimates.

A.3. Proof of Lemma 3

Denote the preliminary OLS estimator for β by $\hat{\beta}$, and notice that

$$P_t = -2(\hat{\beta}-\beta)\,\mathrm{Re}\left(\sum_{s\neq t}\omega_{ts} I_{xu}(\lambda_s)\right) + (\hat{\beta}-\beta)' \sum_{s\neq t} \omega_{ts} I_{xx}(\lambda_s)(\hat{\beta}-\beta)$$

$$= -2(\hat{\beta}-\beta)\,\mathrm{Re}(\hat{f}_{xu}(\lambda_t)) + (\hat{\beta}-\beta)'\hat{f}_{xx}(\lambda_t)(\hat{\beta}-\beta)$$

and

$$n(\hat{\beta}-\beta) \Rightarrow \left[\int B_x B_x'\right]^{-1} \int B_x \, \mathrm{d}B_u.$$

Following Phillips (1991), it can be verified that

$$\sum_s \omega_{ts} I_{xu}(\lambda_s) = \hat{f}_{xu}(\lambda_t)$$

$$= \frac{1}{2\pi}\sum_h k\left(\frac{h}{M}\right) C_{xu}(h)\, \mathrm{e}^{\mathrm{i}\lambda_t h}$$

$$\sim \frac{1}{2\pi}\sum_h k\left(\frac{h}{M}\right) \mathrm{e}^{\mathrm{i}\lambda_t h} \int B_x \, \mathrm{d}B_u.$$

Notice that for the Daniell kernel

$$K(\lambda) = \frac{1}{2\pi M}\sum_h k\left(\frac{h}{M}\right) \mathrm{e}^{\mathrm{i}\lambda_t (h/M)} = \begin{cases} \pi^{-1} & \text{for } |\lambda| \leqslant \pi/2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\sum_s \omega_{ts} I_{xu}(\lambda_s) \sim M \cdot K(M\lambda_t) \int B_x \, \mathrm{d}B_u$$

$$= \begin{cases} \mathrm{O_p}(M) & \text{if } |\lambda_t| \leqslant \pi/(2M), \\ 0 & \text{otherwise.} \end{cases} .$$

Similarly

$$\sum_s \omega_{ts} I_{xx}(\lambda_s) = \hat{f}_{xx}(\lambda_t)$$

$$= \frac{n}{2\pi} \sum_h k\left(\frac{h}{M}\right) \left[\frac{C_{xx}(h)}{n}\right] \mathrm{e}^{\mathrm{i}\lambda_t h}$$

$$\sim \frac{n}{2\pi} \sum_h k\left(\frac{h}{M}\right) \mathrm{e}^{\mathrm{i}\lambda_t h} \int B_x B_x'$$

$$= MnK(M\lambda_t) \int B_x B_x'$$

$$= \begin{cases} \mathrm{O_p}(nM) & \text{if } |\lambda_t| \leqslant \pi/(2M), \\ 0 & \text{otherwise.} \end{cases} .$$

Thus, for $|\lambda_t| > \pi/(2M)$,

$$P_t = -2(\hat{\beta} - \beta) \sum_{s \neq t} \omega_{ts} I_{xu}(\lambda_s) + (\hat{\beta} - \beta)' \sum_{s \neq t} \omega_{ts} I_{xx}(\lambda_s)(\hat{\beta} - \beta)$$

$$= \mathrm{o_p}(m^{-1}),$$

and for $|\lambda_t| \leqslant \pi/(2M)$,

$$P_t = -2(\hat{\beta} - \beta) \sum_{s \neq t} \omega_{ts} I_{xu}(\lambda_s) + (\hat{\beta} - \beta)' \sum_{s \neq t} \omega_{ts} I_{xx}(\lambda_s)(\hat{\beta} - \beta)$$

$$\sim -\frac{1}{2m} K(M\lambda_t) \int \mathrm{d}B_u B_x' \left[\int B_x B_x'\right]^{-1} \int B_x \, \mathrm{d}B_u. \qquad \square$$

A.4. Proof of Lemma 5

Since $g(\lambda)$ has uniformly and absolutely convergent Fourier representation, we can write the sum as

$$n^{-2} \sum_j I_{xx}(\lambda_j) g(\lambda_j) = n^{-2} \sum_j I_{xx}(\lambda_j) \left[\frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g \mathrm{e}^{\mathrm{i}g\lambda_j}\right]$$

$$= n^{-2} \sum_j \left[\frac{1}{2\pi} \sum_r C_{xx}(r) \mathrm{e}^{-\mathrm{i}\lambda_j r}\right] \left[\frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g \mathrm{e}^{\mathrm{i}g\lambda_j}\right]$$

$$= \frac{1}{4\pi^2} n^{-2} \sum_g D_g \sum_r C_{xx}(r) \sum_j \mathrm{e}^{\mathrm{i}(g-r)\lambda_j}.$$

Note that the summations may be interchanged because the infinite series in h and g converge absolutely and the sum over j is finite. Notice that $\sum \mathrm{e}^{\mathrm{i}\lambda_j(g-r)} = 0$ unless $g - r = n * l$ for some l. Thus,

$$n^{-2}\sum_j I_{xx}(\lambda_j)g(\lambda_j) = \frac{1}{4\pi^2}\, n^{-2}\sum_r C_{xx}(r)\sum_{|l|=0}^{\infty} D_{r+n*l}$$

$$\to_{\mathrm{d}} \frac{1}{2\pi}\, g(0)\int B_x B_x'.$$

A.5. Lemma A.1

$$Z_0 = \mathrm{O_p}(1), \quad Z_{V1} = \mathrm{O_p}(m^{-1/2}), \quad Z_{V2} = \mathrm{O_p}(m^{-1}),$$

$$Z_{B1} = \mathrm{O_p}(M^{-q}), \quad Z_{B2} = \mathrm{O_p}(M^{-2q}), \quad Z_{P1} = \mathrm{O_p}(m^{-1}),$$

and

$$D_{V1} = \mathrm{O_p}(m^{-1/2}), \quad D_{V2} = \mathrm{O_p}(m^{-1}), \quad D_{B1} = \mathrm{O_p}(M^{-q}), \quad D_{B2} = \mathrm{O_p}(M^{-2q}),$$

$$D_{P1} = \mathrm{O_p}(m^{-1}).$$

Proof. First, we show that $D_{V1} = \mathrm{O_p}(m^{-1/2})$ by verifying that $\mathrm{E}[D_{V1}]^2 = \mathrm{O_p}(m^{-1})$. Notice that

$$D_{V1} = \frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2}\sum_{s\neq t}\omega_{ts}[I_{uu}(\lambda_s) - f_{uu}(\lambda_s)]$$

$$= \frac{1}{n}\sum_s \varphi_s[I_{uu}(\lambda_s) - f_{uu}(\lambda_s)],$$

where

$$\varphi_s = \sum_t \omega_{ts}\left(\frac{1}{n} I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2}\right)$$

$$\sim \frac{1}{n} f_{uu}(\lambda_s)^{-2}\sum_t \omega_{ts} I_{xx}(\lambda_t)$$

$$\sim \frac{1}{n} f_{uu}(\lambda_s)^{-2} nM \cdot K(M\lambda_s)\int B_x^2$$

$$\sim \begin{cases} M f_{uu}(\lambda_s)^{-2} K(0)\int B_x^2 & \text{if } |\lambda_s| \leqslant \pi/(2M), \\ 0 & \text{otherwise.} \end{cases}$$

Under the regularity conditions given in Section 2, $\mathrm{E}[I_{uu}(\lambda_s) - f_{uu}(\lambda_s)] = \mathrm{O}(n^{-1})$, uniformly in λ_s, (see, e.g., Brillinger, 1980, Theorem 5.22), and it can be shown that

$\mathrm{E}D_{V1} = \mathrm{O}(n^{-1})$. For the second moment, notice that $Var(I_{uu}(\lambda_t)) \sim f_{uu}(\lambda_t)^2$, and

$$\mathrm{E}[D_{V1}]^2 \sim \frac{1}{n^2} \sum_{|\lambda_s| \leqslant \pi/(2M)} M^2 f_{uu}(\lambda_s)^{-4} K(0)^2 \mathrm{E}\left(\int B_x^2\right)^2 f_{uu}(\lambda_s)^2 + \frac{1}{n^2} \sum_{s \neq t} \varphi_s \varphi_t \, Cov[I_{uu}(\lambda_s), I_{uu}(\lambda_t)].$$

Since λ_s and λ_t are fundamental frequencies, $Cov[I_{uu}(\lambda_s), I_{uu}(\lambda_t)]$ is of order $\mathrm{O}(n^{-1})$, uniformly in $s \neq t$. Thus,

$$\mathrm{E}[D_{V1}]^2 \sim \frac{1}{(2m)^2} K(0)^2 \sum_{|\lambda_s| \leqslant \pi/(2M)} \mathrm{E}\left(\int B_x^2\right)^2 f_{uu}(\lambda_s)^{-2} \sim \frac{1}{4m} K(0)^2 \mathrm{E}\left(\int B_x^2\right)^2 f_{uu}(0)^{-2}.$$

As a result, $D_{V1} = \mathrm{O_p}(m^{-1/2})$. Next,

$$\begin{aligned}
\mathrm{E}D_{V2} &= \mathrm{E}\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} V_t^2 \\
&\sim \mathrm{E}\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} \frac{1}{m} f_{uu}(\lambda_t)^2 \\
&\sim \frac{1}{m} \mathrm{E}\left[\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1}\right] \\
&\sim \frac{1}{2\pi m} f_{uu}(0)^{-1} \mathrm{E} \int B_x^2 \\
&= \mathrm{O_p}(m^{-1}).
\end{aligned}$$

For D_{B1}, by an application of Lemma 5,

$$\begin{aligned}
D_{B1} &= \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} B_t \\
&\sim \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} M^{-q} k_q f_q(\lambda_t) \\
&\sim M^{-q} k_q \frac{1}{2\pi} f_{uu}(0)^{-2} f_q(0) \int B_x^2 \\
&= \mathrm{O_p}(M^{-q}),
\end{aligned}$$

and similarly for D_{B2} and

$$\begin{aligned}
D_{P1} &\sim \frac{1}{n^2} \sum_{|\lambda_t| \leqslant \pi/(2M)} I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} \left[\frac{1}{2m\pi} \left(\int \mathrm{d}B_u B_x\right)^2 \left(\int B_x^2\right)^{-1}\right] \\
&= \mathrm{O_p}(m^{-1}).
\end{aligned}$$

For Z_0, Z_{V1}, Z_{P1}, etc., we need to check their second moments. For example,

$$\begin{aligned}
\mathrm{E}|Z_0|^2 &= \mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)I_{uu}(\lambda_t)f_{uu}(\lambda_t)^{-2} \\
&\quad + \mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t)w_u(\lambda_s)^* w_u(\lambda_t)f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1} \\
&= A + B \\
&\sim \mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)f_{uu}(\lambda_t)^{-1} \sim \mathrm{E}\,\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x^2.
\end{aligned}$$

Under some regularity conditions and if the serial correlation in u_t is weak, the second term B is seen to be of smaller order of magnitude than the leading term. For example, if we assume that u_t is the linear process $u_t = C(L)\varepsilon_t$, (with $C(L) = \sum_{j=0}^{\infty} c_j L^j$, and where the ε_t are iid—see Assumption A, below) then, as discussed in the proof of Lemma 6 (see (9.5) below),

$$w_u(\lambda_s) = C(\mathrm{e}^{\mathrm{i}\lambda_s})w_\varepsilon(\lambda_s) + r(\lambda_s),$$

where the $w_\varepsilon(\lambda_s)$ are uncorrelated and the remainder term $r(\lambda_s)$ is small. Under the summability condition $\sum_{j=1}^{\infty} j^\rho|c_j| < \infty$, for $\rho \geqslant 1$, an explicit expression for $r(\lambda_s)$ is available from the spectral BN decomposition (see (9.5) below). It is then easily seen that the order of magnitude of $r(\lambda_s)$ is $\mathrm{O_p}(n^{-1/2})$ uniformly in λ_s (c.f. Priestley, 1981, Theorem 6.2.1) and that $\mathrm{E}|w_\varepsilon(\lambda_s)r(\lambda_t)| = \mathrm{O}(1/\sqrt{n})$, $\mathrm{E}|r(\lambda_t)|^2 = \mathrm{O}(1/n)$, and $\mathrm{E}|r(\lambda_s)|^2|r^*(\lambda_t)|^2 = \mathrm{O}(1/n^2)$, uniformly in λ_s, λ_t. Then, writing

$$\begin{aligned}
B &= \mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t)C(\mathrm{e}^{\mathrm{i}\lambda_s})w_\varepsilon(\lambda_s)r^*(\lambda_t)f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1} \\
&\quad + \mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t)r(\lambda_s)w_\varepsilon(\lambda_t)^* C(\mathrm{e}^{\mathrm{i}\lambda_t})^* f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1} \\
&\quad + \mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t)r(\lambda_s)r^*(\lambda_t)f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1} \\
&= B_1 + B_2 + B_3,
\end{aligned}$$

we have

$$\begin{aligned}
&\mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t)C(\mathrm{e}^{\mathrm{i}\lambda_s})w_\varepsilon(\lambda_s)r^*(\lambda_t)f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1} \\
&\quad \leqslant \max_{s,t}\{|C(\mathrm{e}^{\mathrm{i}\lambda_s})|\mathrm{E}\{|w_\varepsilon(\lambda_s)||r^*(\lambda_t)|\}\} \\
&\qquad \times \frac{1}{n^2}\sum_{s\neq t}\{\mathrm{E}I_{xx}(\lambda_s)\mathrm{E}I_{xx}(\lambda_t)\}^{1/2} f_{uu}(\lambda_s)^{-1}f_{uu}(\lambda_t)^{-1},
\end{aligned}$$

and

$$\mathrm{E}\,\frac{1}{n^2}\sum_{s\neq t} w_x(\lambda_s)^* w_x(\lambda_t) r(\lambda_s) r^*(\lambda_t) f_{uu}(\lambda_s)^{-1} f_{uu}(\lambda_t)^{-1}$$

$$\leqslant \max_{s,t}\{\mathrm{E}|r(\lambda_s)|^2|r^*(\lambda_t)|^2\}^{1/2}\frac{1}{n^2}\sum_{s\neq t}\{\mathrm{E}I_{xx}(\lambda_s)\mathrm{E}I_{xx}(\lambda_t)\}^{1/2} f_{uu}(\lambda_s)^{-1} f_{uu}(\lambda_t)^{-1},$$

so that each of the terms B_1, B_2, and B_3, tend to zero. It follows that

$$\mathrm{E}|Z_0|^2 \sim \mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1} \sim \mathrm{E}\,\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2.$$

Similarly

$$\begin{aligned}
\mathrm{E}[Z_{V1}]^2 &\sim \mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3}\frac{1}{m} f_{uu}(\lambda_t)^2 \\
&\sim \frac{1}{m}\,\mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1} \\
&\sim \frac{1}{m}\,\frac{1}{2\pi} f_{uu}(0)^{-1}\mathrm{E}\int B_x^2,
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}[Z_{P1}]^2 &\sim \mathrm{E}\,\frac{1}{n^2}\sum_{|\lambda_t|\leqslant\pi/(2M)} I_{xx}(\lambda_t) I_{uu}(\lambda_t) f_{uu}(\lambda_t)^{-2} \\
&\quad\times\left[\frac{1}{2m\pi}\left(\int \mathrm{d}B_u B_x\right)^2\left(\int B_x^2\right)^{-1}\right]^2 \\
&\sim \mathrm{E}\,\frac{1}{2m^2\pi^3} f_{uu}(0)^{-1}\left(\frac{1}{2\pi}\left[\int \mathrm{d}B_u B_x\right]^2\right)^2\left(\int B_x^2\right)^{-1} \\
&= \mathrm{O_p}(m^{-2}),
\end{aligned}$$

$$\mathrm{E}[Z_{V2}]^2 \sim \frac{1}{m^2}\,\mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1} = \mathrm{O_p}(m^{-2}),$$

$$\begin{aligned}
\mathrm{E}[Z_{B1}]^2 &\sim \mathrm{E}\,\frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} M^{-2q} k_q^2 f_q(\lambda_t)^2 \\
&\sim M^{-2q} k_q^2\,\frac{1}{2\pi}\, f_{uu}(0)^{-3} f_q(0)^2\mathrm{E}\int B_x^2 \\
&= \mathrm{O_p}(M^{-2q}).
\end{aligned}$$

Similar arguments give the result for Z_{B2}.

A.6. Proof of Theorem 1

Notice that

$$W = Z^*\hat{H}^{-1}Z, \tag{9.2}$$

and

$$\hat{H}^{-1} = H^{-1} - H^{-2}D + H^{-3}D^2 - R_2,$$

$$Z = Z_0 - Z_{V1} + Z_{V2} - Z_{B1} + Z_{B2} - Z_{P1} + R_3,$$

$$D = -D_{V1} + D_{V2} - D_{B1} + D_{B2} - D_{P1} + R_4.$$

The terms in these expansions are analyzed in Lemma A.1. Drop the higher order terms and substitute the truncated versions of Z, D, and $\hat{H}^{-1}$ into (9.2), giving

$$\begin{aligned}
W \sim\ & H^{-1}Z_0^2 \\
& + H^{-1}Z_{V1}^2 - 2H^{-1}Z_0^*Z_{P1} + H^{-2}D_{P1}Z_0^*Z_0^* + H^{-3}D_{V1}^2Z_0^*Z_0 \\
& + H^{-1}Z_{B1}^2 + H^{-3}D_{B1}^2Z_0^2 - 2H^{-2}D_{B1}Z_0Z_{B1} - H^{-2}D_{B2}Z_0^2 \\
& - 2H^{-1}Z_0Z_{B1} + H^{-2}D_{B1}Z_0^2 \\
& + H^{-2}D_{V1}Z_0^2 - 2H^{-1}Z_0Z_{V1} \\
& + 2H^{-1}Z_0Z_{V2} - H^{-2}D_{V2}Z_0^2 \\
& + 2H^{-1}Z_{B2}Z_0 \\
& + 2H^{-1}Z_{V1}Z_{B1} - 2H^{-2}D_{B1}Z_0Z_{V1}.
\end{aligned}$$

It can be verified that, besides the leading term that is of order $O_p(1)$, the above approximation contains terms of order m^{-1}, M^{-q}, and M^{-2q}, etc. The bias effect of order M^{-q} dominates the $O(M^{-2q})$ terms, giving the results in Theorem 1.

A.7. Lemma A.2

(1) $m\mathrm{E}[H^{-1}Z_0^*Z_{V2}] \to 1$;
(2) $m\mathrm{E}[H^{-2}D_{V2}Z_0^*Z_0] \to 1$;
(3) $m\mathrm{E}[H^{-3}D_{V1}^2Z_0^*Z_0] \to 1$;
(4) $m\mathrm{E}[H^{-1}Z_{V1}^*Z_{V1}] \to 1$;
(5) $m\mathrm{E}[H^{-1}Z_0^*Z_{P1}] \to -1$;
(6) $m\mathrm{E}[H^{-2}D_{P1}Z_0^*Z_0] \to -1$;
(7) $M^q\mathrm{E}[H^{-1}Z_0^*Z_{B1}] \to -k_q f_{uu}(0)^{-1}f_q(0)$;
(8) $M^q\mathrm{E}[H^{-2}D_{B1}Z_0^*Z_0] \to -k_q f_{uu}(0)^{-1}f_q(0)$;
(9) $M^{2q}\mathrm{E}[H^{-1}Z_{B1}^*Z_{B1}] \to k_q^2 f_{uu}(0)^{-2}f_q(0)^2$;
(10) $M^{2q}\mathrm{E}[H^{-3}D_{B1}^2Z_0^*Z_0] \to k_q^2 f_{uu}(0)^{-2}f_q(0)^2$;
(11) $M^{2q}\mathrm{E}[H^{-1}Z_0^*Z_{B2}] \to k_q^2 f_{uu}(0)^{-2}f_q(0)^2$;
(12) $M^{2q}\mathrm{E}[H^{-2}D_{B1}Z_0^*Z_{B1}] \to k_q^2 f_{uu}(0)^{-2}f_q(0)^2$;
(13) $M^{2q}\mathrm{E}[H^{-2}D_{B2}Z_0^*Z_0] \to k_q^2 f_{uu}(0)^{-2}f_q(0)^2$.

Proof.

$$\mathrm{E}[H^{-1}Z_0^2] = \mathrm{E}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^2$$

$$\sim \mathrm{E}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-2}\right]$$

$$\to 1,$$

and

$$H^{-1}Z_{V1} = \left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s\right].$$

Since the expectation of this term is zero, we check the order of the second moment. Notice that

$$H \to_{\mathrm{d}} \frac{1}{2\pi f_{uu}(0)}\int B_x^2,$$

by Lemma 5, and

$$\mathrm{E}[Z_{V1}]^2 = \mathrm{E}\,\frac{1}{n^2}\sum_t\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s I_{xu}(\lambda_t)f_{uu}(\lambda_t)^{-2}V_t$$

$$\sim \mathrm{E}\,\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-4}I_{uu}(\lambda_s)V_s^2.$$

Since the leave-one out estimator is used,

$$\mathrm{E}[Z_{V1}]^2 \sim \frac{1}{n^2}\sum_s \mathrm{E}I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-4}\,\mathrm{E}[I_{uu}(\lambda_s)]\mathrm{E}[V_s]^2$$

$$\sim \mathrm{E}\frac{1}{mn^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}$$

$$\sim \mathrm{E}\left\{\frac{1}{m}\,\frac{1}{2\pi}\,f_{uu}(0)^{-1}\int B_x^2\right\},$$

$$m\mathrm{E}[H^{-1}Z_{V1}^2] = m\mathrm{E}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s\right]^2$$

$$\sim m\mathrm{E}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-4}V_s^2\right]$$

$$\sim a = 1,$$

and so

$$m\mathrm{E}[Z_{V2}^*H^{-1}Z_0] \sim m\mathrm{E}H^{-1}\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-3}V_s^2\right]$$

$$\sim m\mathrm{E}H^{-1}\left[\frac{1}{n}\sum I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-4}V_s^2\right]$$

$$\sim a = 1.$$

Similarly,

$$mE[Z_0^* H^{-2} D_{V2} Z_0] \sim a = 1,$$

$$mE[H^{-3} D_{V1}^2 Z_0^* Z_0] \sim m \left[\frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2\right]^{-3}$$

$$\left[\frac{1}{4m\pi^2}\left(\int B_x^2\right)^2 f_{uu}(0)^{-2}\right]\left[\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2\right] \sim 1.$$

$$E[H^{-1} Z_0^* Z_{P1}]$$

$$\sim E\left[\frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x B_x'\right]^{-1}\left[\frac{1}{n}\sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]$$

$$\times\left[\frac{1}{n}\sum_{|\lambda_t|\leqslant \pi/(2M)} I_{xu}(\lambda_t) f_{uu}(\lambda_t)^{-2}\left(-\frac{1}{2m\pi}\int dB_u B_x \left[\int B_x^2\right]^{-1}\int B_x\, dB_u\right)\right]$$

$$\sim -E\left[\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2\right]^{-1}\left[\frac{1}{n^2}\sum_{|\lambda_t|\leqslant \pi/(2M)} I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2}\right]$$

$$\times\left(\frac{1}{2m\pi}\int dB_u B_x\left[\int B_x^2\right]^{-1}\int B_x\, dB_u\right)$$

$$\sim -E\left[\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2\right]^{-1}\left[\frac{1}{n^2}\sum_{|\lambda_t|\leqslant \pi/(2M)} I_{xx}(\lambda_t)(f_{uu}(0)^{-2} + o(1))\right]$$

$$\left(\frac{1}{2m\pi}\int dB_u B_x\left[\int B_x^2\right]^{-1}\int B_x\, dB_u\right)$$

$$\sim -E\left[\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2\right]^{-1}\left[\frac{m}{n^2}\left(\frac{1}{m}\sum_{|\lambda_t|\leqslant \pi/(2M)} I_{xx}(\lambda_t)\right) f_{uu}(0)^{-2}\right]$$

$$\times\left(\frac{1}{2m\pi}\int dB_u B_x\left[\int B_x^2\right]^{-1}\int B_x\, dB_u\right)$$

$$\sim -E\left[\frac{1}{2\pi} f_{uu}(0)^{-1}\int B_x^2\right]^{-1}\left[\frac{m}{n^2}\hat{f}_{xx}(0) f_{uu}(0)^{-2}\right]$$

$$\times\left(\frac{1}{2m\pi}\int dB_u B_x\left[\int B_x^2\right]^{-1}\int B_x\, dB_u\right)$$

$$\sim -\mathrm{E}\left[\frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2\right]^{-1} \left[\frac{1}{2\pi} f_{uu}(0)^{-2} \int B_x^2\right]$$

$$\times \left(\frac{1}{2m\pi} \int \mathrm{d}B_u B_x \left[\int B_x^2\right]^{-1} \int B_x \, \mathrm{d}B_u\right)$$

$$\sim -\frac{1}{m} \mathrm{E}\left(\frac{1}{2\pi} f_{uu}(0)^{-1} \int \mathrm{d}B_u B_x \left[\int B_x^2\right]^{-1} \int B_x \, \mathrm{d}B_u\right).$$

Notice that $B_u(r) = BM(2\pi f_{uu}(0)) =_{\mathrm{d}} [2\pi f_{uu}(0)]^{1/2} W_u(r)$, where $W_u(r)$ is a standard Brownian motion independent of $B_x(r)$. Thus,

$$\mathrm{E}\left(\frac{1}{2\pi} f_{uu}(0)^{-1} \int \mathrm{d}B_u B_x \left[\int B_x^2\right]^{-1} \int B_x \, \mathrm{d}B_u\right) = 1,$$

and

$$\mathrm{E}[H^{-1} Z_0^* Z_{P1}] \sim -m^{-1}.$$

$$\mathrm{E}[Z_0^* H^{-2} D_{B1} Z_0] \sim \mathrm{E}H^{-2} D_{B1} \left[\frac{1}{\sqrt{n}} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]$$

$$\times \left[\frac{1}{\sqrt{n}} \sum I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]$$

$$\sim -M^{-q} k_q f_{uu}(0)^{-1} f_q(0),$$

$$\mathrm{E}[Z_{B1}^* H^{-1} Z_0] = \mathrm{E}H^{-1} \left[\frac{1}{n} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s\right] \left[\frac{1}{\sqrt{n}} \sum I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]$$

$$\sim -M^{-q} k_q f_{uu}(0)^{-1} f_q(0),$$

$$\mathrm{E}[Z_{B1}^* H^{-1} Z_{B1}] = \mathrm{E}H^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum_j I_{xx}(\lambda_j) I_{uu}(\lambda_j) f_q(\lambda_j)^2 f_u(\lambda_j)^{-4}$$

$$\sim \mathrm{E}H^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum_j I_{xx}(\lambda_j) f_q(\lambda_j)^2 f_u(\lambda_j)^{-3}$$

$$= \mathrm{E}H^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_q(\lambda_t)^2 f_u(\lambda_t)^{-3}$$

$$= \mathrm{E}M^{-2q} k_q^2 H^{-1} \frac{1}{n^2} \sum_t \left[\frac{1}{2\pi} \sum_r C_{xx}(r) \mathrm{e}^{-\mathrm{i}r\lambda_t}\right]$$

$$\times \left[\frac{1}{(2\pi)^2} \sum_{l,j} C_{uu}(l) C_{uu}(j) |j|^q |l|^q \mathrm{e}^{-\mathrm{i}(l+j)\lambda_t}\right]$$

$$\times \left[\frac{1}{(2\pi)^3} \sum_{g,h,p} D_g D_h D_{\mathrm{p}} \mathrm{e}^{\mathrm{i}(g+h+p)\lambda_t}\right]$$

$$
\begin{aligned}
&= \mathrm{E}M^{-2q}k_q^2H^{-1}\frac{1}{n^2}\frac{1}{(2\pi)^5}\sum_{g,h,p}\sum_{l,j}C_{uu}(j)C_{uu}(l)|j|^q|l|^qD_gD_hD_p \\
&\quad \times\sum_r C_{xx}(r)\left(\sum_{t=-n/2+1}^{n/2}\mathrm{e}^{\mathrm{i}(g+h+p-j-l-r)\lambda_t}\right) \\
&= \mathrm{E}M^{-2q}k_q^2H^{-1}\frac{1}{n^2} \\
&\quad \times\left[\frac{1}{(2\pi)^5}\sum_{g,h,p}\sum_{l,j}C_{uu}(j)C_{uu}(l)|j|^q|l|^qD_gD_hD_pC_{xx}(\underline{r})n\right] \\
&\sim M^{-2q}k_q^2H^{-1}\frac{1}{2\pi}f_{uu}(0)^{-3}f_q(0)^2\int B_xB_x' \\
&\sim M^{-2q}k_q^2f_{uu}(0)^{-2}f_q(0)^2.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathrm{E}[Z_0^*H^{-3}D_{B1}^2Z_0] &\sim \mathrm{E}H^{-3}k_q^2M^{-2q}\left[\frac{1}{n^2}\sum_j I_{xx}(\lambda_j)f_q(\lambda_j)f_u(\lambda_j)^{-2}\right]^2 \\
&\quad \times\left[\frac{1}{n^2}\sum_j I_{xx}(\lambda_j)I_{uu}(\lambda_j)f_u(\lambda_j)^{-1}\right] \\
&\sim \mathrm{E}H^{-4}k_q^2M^{-2q}\left[\frac{1}{2\pi}f_{uu}(0)^{-2}f_q(0)\int B_x^2\right]^2 \\
&\quad \times\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x^2\right] \\
&= k_q^2M^{-2q}f_{uu}(0)^{-2}f_q(0)^2,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}[Z_{B1}^*H^{-2}D_{B1}Z_0] &\sim \mathrm{E}H^{-2}\left[\frac{1}{n}\sum_j I_{ux}(\lambda_j)B_jf_u(\lambda_j)^{-2}\right] \\
&\quad \times\left[\frac{1}{n}\sum_j I_{xx}(\lambda_j)B_jf_u(\lambda_j)^{-2}\right]\left[\frac{1}{n}\sum_j I_{xu}(\lambda_j)f_u(\lambda_j)^{-1}\right] \\
&\sim k_q^2M^{-2q}f_{uu}(0)^{-2}f_q(0)^2.
\end{aligned}
$$

Finally,

$$\mathrm{E}[Z_{B2}^*H^{-1}Z_0] \sim k_q^2M^{-2q}f_{uu}(0)^{-2}f_q(0)^2,$$

and

$$\mathrm{E}[Z_0^*H^{-2}D_{B2}Z_0] \sim k_q^2M^{-2q}f_{uu}(0)^{-2}f_q(0)^2. \qquad \square$$

A.8. Proof of Lemma 6

We have

$$\sqrt{2\ell}(\hat{\beta}_B - \beta) = \left[\frac{1}{2\ell}\sum_{\lambda_s \in B} I_{xx}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s \in B} I_{xu}(\lambda_s)\hat{f}_{uu}(\lambda_s)^{-1}\right],$$

and, following the proof of Lemma 1, we perform a geometric expansion on $\hat{f}_{uu}(\lambda_s)^{-1}$, and analyze the leading terms. In particular, we need

$$\frac{1}{2\ell}\sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1} \to_{\mathrm{d}} f_{uu}(\omega)^{-1} h_x(\omega), \tag{9.3}$$

and we also need the following central limit theorem:

$$\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} \to_{\mathrm{d}} MN(0, f_{uu}(\omega)^{-1} h_x(\omega)). \tag{9.4}$$

We start with (9.4) and prove (9.3) en route. For convenience in proving this CLT, we assume that u_t follows a linear process (Phillips and Solo, 1992):

Assumption A.1. u_t is independent with v_t and has a linear process representation: $u_t = C(L)\varepsilon_t$, $C(L) = \sum_{j=0}^{\infty} c_j L^j$, where L is the lag operator defined as $L\varepsilon_t = \varepsilon_{t-1}$, $C(1) \neq 0$, $\sum_{j=1}^{\infty} j^{\rho}|c_j| < \infty$, for some $\rho \geqslant 1$. ε_t is an i.i.d. process with zero mean and finite 4th moments.

The linear process assumption is satisfied by a wide class of time series models and has a long history in frequency-domain time series analysis—see, for example, Hannan (1960), Walker (1965) and Priestley (1981). Under the summability condition in Assumption A.1, it is known (e.g. Phillips and Solo, 1992) that partial sums of u_t satisfy the invariance principle $n^{-1/2}\sum_{t=1}^{[Tr]} u_t \Rightarrow B_u(r) \equiv BM(2\pi f_{uu}(0))$. In addition, we have the following spectral decomposition of the operator:

$$C(L) = C(\mathrm{e}^{\mathrm{i}\lambda}) + \tilde{C}_{\lambda}(L)(L\mathrm{e}^{-\mathrm{i}\lambda} - 1),$$

where

$$\tilde{C}_{\lambda}(L) = \sum_{j=0}^{\infty} \tilde{c}_j L^j, \quad \tilde{c}_j = \mathrm{e}^{-\mathrm{i}j\lambda}\left(\sum_{l=j+1}^{\infty} c_l \mathrm{e}^{\mathrm{i}l\lambda}\right).$$

Thus, as in Phillips and Solo (1992), we get the decomposition

$$u_t = C(L)\varepsilon_t = C(\mathrm{e}^{\mathrm{i}\lambda})\varepsilon_t + \mathrm{e}^{-\mathrm{i}\lambda}\tilde{\varepsilon}_{\lambda t-1} - \tilde{\varepsilon}_{\lambda t}$$

where

$$\tilde{\varepsilon}_{\lambda t} = \tilde{C}_{\lambda}(L)\varepsilon_t.$$

Then,

$$\begin{aligned} w_u(\lambda_s) &= \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} u_t \mathrm{e}^{\mathrm{i}t\lambda_s} = C(\mathrm{e}^{\mathrm{i}\lambda_s}) w_{\varepsilon}(\lambda_s) + \frac{1}{\sqrt{2\pi n}}(\tilde{\varepsilon}_{\lambda_s 0} - \mathrm{e}^{-\mathrm{i}n\lambda_s}\tilde{\varepsilon}_{\lambda_s n}) \\ &= C(\mathrm{e}^{\mathrm{i}\lambda_s}) w_{\varepsilon}(\lambda_s) + r_n(\lambda_s). \end{aligned} \tag{9.5}$$

Notice that the ε_t are i.i.d. $(0, \sigma^2)$, and the $w_\varepsilon(\lambda_s)$ form an uncorrelated sequence. It is readily seen from (9.5) that $r_n(\lambda_s)$ is $\mathrm{O_p}(1/\sqrt{n})$, uniformly in λ_s.

By definition, the discrete transform of the $I(1)$ process x_t is

$$
\begin{aligned}
w_x(\lambda_s) &= n^{-1/2}\sum_{t=1}^{n} x_{t-1}\mathrm{e}^{\mathrm{i}\lambda_s t} + w_v(\lambda_s) \\
&= \mathrm{e}^{\mathrm{i}\lambda_s}[w_x(\lambda_s) - n^{-1/2}(\mathrm{e}^{\mathrm{i}\lambda_s n}x_n - x_0)] + w_v(\lambda_s),
\end{aligned}
$$

so that, as shown in Corbae et al. (2002),

$$
w_x(\lambda_s) = \frac{1}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\, w_v(\lambda_s) - \frac{\mathrm{e}^{\mathrm{i}\lambda_s}}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\,\frac{[x_n - x_0]}{n^{1/2}}. \tag{9.6}
$$

Using this result and the continuity of f_{uu} it is a straightforward matter to show that (9.3) holds. We also have

$$
\begin{aligned}
&\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} \\
&\quad = \frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1}\left[\frac{1}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\, w_v(\lambda_s) w_u^*(\lambda_s)\right] \\
&\qquad - \left[\frac{x_n - x_0}{n^{1/2}}\right]\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1}\left[\frac{\mathrm{e}^{-\mathrm{i}\lambda_s}}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\right] w_u^*(\lambda_s).
\end{aligned}
$$

From (9.5)

$$
\begin{aligned}
&\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} \\
&\quad = \frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1} w_x(\lambda_s) w_\varepsilon^*(\lambda_s) C(\mathrm{e}^{\mathrm{i}\lambda_s})^* \\
&\qquad + \frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1} w_x(\lambda_s) r_n^*(\lambda_s),
\end{aligned} \tag{9.7}
$$

and by (9.6),

$$
w_x(\lambda_s) r_n^*(\lambda_s) = \left[\frac{1}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\, w_v(\lambda_s) r_n^*(\lambda_s) - \frac{[x_n - x_0]}{n^{1/2}}\,\frac{\mathrm{e}^{-\mathrm{i}\lambda_s}}{1-\mathrm{e}^{\mathrm{i}\lambda_s}} r_n^*(\lambda_s)\right].
$$

Note that $w_v(\lambda_s) r^*(\lambda_s)$ has zero mean, variance of $\mathrm{O}(n^{-1})$ and is asymptotically uncorrelated over frequencies because $w_v(\lambda_s)$ is asymptotically uncorrelated across frequencies and is independent of $r_n^*(\lambda_t)$ for all λ_t. It follows that

$$
\mathrm{E}\left|\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1} w_v(\lambda_s) r_n^*(\lambda_s)\right|^2 = \mathrm{O}\left(\frac{1}{n}\right).
$$

Also, since $\mathrm{E}|r_n^*(\lambda_s)| = \mathrm{O}(n^{-1/2})$ uniformly in λ_s we have

$$
\mathrm{E}\left|\frac{[x_n - x_0]}{n^{1/2}}\,\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1}\,\frac{\mathrm{e}^{-\mathrm{i}\lambda_s}}{1-\mathrm{e}^{\mathrm{i}\lambda_s}}\, r^*(\lambda_s)\right| = \mathrm{O}\left(\sqrt{\frac{l}{n}}\right) = \mathrm{o}(1),
$$

since $\ell/n \to 0$, and we deduce that

$$\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1}w_x(\lambda_s)r_n^*(\lambda_s) = o_p(1).$$

Thus,

$$\begin{aligned}
&\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\\
&\quad=\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} f_{uu}(\lambda_s)^{-1}w_x(\lambda_s)w_\varepsilon^*(\lambda_s)C(e^{i\lambda_s})^* + o_p(1)\\
&\quad= f_{uu}(\omega)^{-1}\left[\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B} w_x(\lambda_s)w_\varepsilon^*(\lambda_s)C(e^{i\lambda_s})^*\right] + o_p(1)\\
&\quad= f_{uu}(\omega)^{-1}\left[\frac{2}{\sqrt{2\ell}}\operatorname{Re}\left\{\sum_{\lambda_s\in B_\omega} w_x(\lambda_s)w_\varepsilon^*(\lambda_s)C(e^{i\lambda_s})^*\right\}\right] + o_p(1).
\end{aligned}$$

Since B_ω is a shrinking band, the $w_\varepsilon(\lambda_s)$ are asymptotically independent[2] $N_c(0,\sigma^2/2\pi)$ and are independent of $w_x(\lambda_s)$. Thus,

$$\begin{aligned}
\frac{1}{\sqrt{2\ell}}\sum_{\lambda_s\in B_\omega} w_x(\lambda_s)w_\varepsilon^*(\lambda_s)C(e^{i\lambda_s})^* &\to_d MN_c(0, h_x(\omega)\sigma^2 C(e^{i\omega})C(e^{i\omega})^*)\\
&= MN_c(0, h_x(\omega)f_{uu}(\omega)),
\end{aligned}$$

and since the real and imaginary parts of this limit distribution have conditional variance $\frac{1}{2}h_x(\omega)f_{uu}(\omega)$, it follows that

$$\sqrt{\frac{2}{\ell}}\operatorname{Re}\left\{\sum_{\lambda_s\in B_\omega} w_x(\lambda_s)w_\varepsilon^*(\lambda_s)C(e^{i\lambda_s})^*\right\} \to_d MN(0, h_x(\omega)f_{uu}(\omega)),$$

giving the required result (9.4). □

By an analysis similar to that of Lemmas A1 and A2, we obtain the following results.

A.9. Lemma A.3

$$\Delta_{V1} = \frac{1}{2\ell}\sum_{\lambda_s\in B} I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s = O_p(m^{-1/2}),$$

[2] Hannan (1973, Theorem 3) showed that a finite collection of $w_\varepsilon(\lambda_s)$ are asymptotically independent and satisfy a central limit theorem. Phillips (2000, Theorem 3.2) recently showed that an asymptotically infinite collection of $w_\varepsilon(\lambda_s)$ satisfy a central limit theorem and are asymptotically independent for frequencies λ_s in the neigbourhood of the origin provided the number of frequencies $m = o(n^{-1/2-1/p})$ where $E(|\varepsilon_t|^p) < \infty$ for some $p > 2$, i.e. provided the number of frequencies does not go to infinity too fast. The Phillips result can be extended to frequency bands away from the origin, although a proof was not given in that paper. In the present context, the collection $\{w_\varepsilon(\lambda_s)\colon \lambda_s \in B_\omega\}$ is asymptotically infinite and the number of frequencies in B_ω is ℓ. We need to choose ℓ such that $1/\ell + \ell/n^{1/4} \to 0$ in the present case since 4th moments of ε_t are finite.

$$\Delta_{V2} = \frac{1}{2\ell} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = \mathrm{O_p}(m^{-1}),$$

$$\Delta_{B1} = \frac{1}{2\ell} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = \mathrm{O_p}(M^{-q}),$$

$$\Delta_{B2} = \frac{1}{2\ell} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = \mathrm{O_p}(M^{-2q}),$$

and

$$\Psi_0 = \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} = \mathrm{O_p}(1),$$

$$\Psi_{V1} = \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s = \mathrm{O_p}(m^{-1/2}),$$

$$\Psi_{V2} = \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = \mathrm{O_p}(m^{-1}),$$

$$\Psi_{B1} = \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = \mathrm{O_p}(M^{-q}),$$

$$\Psi_{B2} = \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = \mathrm{O_p}(M^{-2q}). \qquad \square$$

A.10. Lemma A.4

(1) $m\mathrm{E}[H_B^{-1}\Psi_0^*\Psi_{V2}] \to 1$,
(2) $m\mathrm{E}[H_B^{-2}\Delta_{V2}\Psi_0^*\Psi_0] \to 1$,
(3) $m\mathrm{E}[H_B^{-1}\Psi_0^*\Psi_0] \to 1$,
(4) $m\mathrm{E}[H_B^{-3}\Delta_{V1}^2\Psi_0^*\Psi_0] \to 1$,
(5) $M^q H_B^{-1}\Psi_0^*\Psi_{B1} \xrightarrow{\mathrm{d}} -k_q f_{uu}(\omega)^{-1} f_q(\omega)$,
(6) $M^q H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_0 \xrightarrow{\mathrm{d}} -k_q f_{uu}(\omega)^{-1} f_q(\omega)$.

A.11. Proof of Theorem 4

Notice that

$$W = Z^*\hat{H}^{-1}R'[J_0^{-1} - J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1} + J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1}$$
$$(\hat{J}_0 - J_0)J_0^{-1} + R_6]^{-1}R\hat{H}^{-1}Z,$$

and $\hat{H}^{-1} = H^{-1} - H^{-1}DH^{-1} + H^{-1}DH^{-1}DH^{-1} + R_1$. Denote $H^{-1}DH^{-1} = A_1$, $H^{-1}DH^{-1}DH^{-1} = A_2$, $RH^{-1}DH^{-1}R' = J_1$, and $RH^{-1}DH^{-1}DH^{-1}R' = J_2$, and drop higher order terms in the covariance matrix expansion, giving the following approximation:

$$W \sim Z^*(H^{-1} - A_1 + A_2)R'[J_0^{-1} + J_0^{-1}J_1J_0^{-1} - J_0^{-1}J_2J_0^{-1} + J_0^{-1}J_1J_0^{-1}J_1J_0^{-1}]^{-1}$$
$$R(H^{-1} - A_1 + A_2)Z$$
$$\sim Z^*H^{-1}R'J_0^{-1}RH^{-1}Z + Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RH^{-1}Z - 2Z^*H^{-1}R'J_0^{-1}RA_1Z$$
$$-Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RH^{-1}Z + Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}J_1J_0^{-1}RH^{-1}Z$$
$$-2Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RA_1Z + 2Z^*H^{-1}R'J_0^{-1}RA_2Z + Z^*A_1R'J_0^{-1}RA_1Z.$$

D and Z follow similar expansions as (4.6), (4.7), Substituting these preliminary expansions into the above formula, and collecting terms up to $O_p(m^{-1}+M^{-q})$, we obtain the following expansion for the expected Wald statistic.

$$E[W] \sim E[Z_0^*H^{-1}R'(RH^{-1}R')^{-1}RH^{-1}Z_0] + E[Z_{V1}^*H^{-1}R'J_0^{-1}RH^{-1}Z_{V1}]$$
$$+ E[Z_0^*H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0]$$
$$- E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0]$$
$$+ E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0]$$
$$- 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0]$$
$$+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}Z_0]$$
$$+ E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V2}H^{-1}R'J_0^{-1}RH^{-1}Z_0]$$
$$+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_{V2}] - 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V2}H^{-1}Z_0]$$
$$+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{B1}H^{-1}Z_0] - 2E[Z_{B1}^*H^{-1}R'J_0^{-1}RH^{-1}Z_0]$$
$$- E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{B1}H^{-1}R'J_0^{-1}RH^{-1}Z_0]$$
$$- 2Z_{P1}^*H^{-1}R'J_0^{-1}RH^{-1}Z_0 + 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{P1}H^{-1}Z_0]$$
$$- E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{P1}H^{-1}R'J_0^{-1}RH^{-1}Z_0].$$

A.12. Lemma A.5

$$mE[Z_{V1}^*H^{-1}R'J_0^{-1}RH^{-1}Z_{V1}] \to p,$$
$$mE[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_{V2}] \to p,$$
$$mE[Z_0^*H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0] \to p,$$
$$mE[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0] \to p,$$
$$mE[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}Z_0] \to p,$$
$$mE[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}Z_0] \to p,$$
$$mE[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}D_{V1}H^{-1}D_{V1}H^{-1}Z_0] \to p,$$

$$mE[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \to p,$$

$$mE[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} Z_0] \to p,$$

$$M^q E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0] \to -p k_q f_{uu}(0)^{-1} f_q(0),$$

$$M^q E[Z_{B1}^* H^{-1} R' J_0^{-1} R H^{-1} Z_0] \to -p k_q f_{uu}(0)^{-1} f_q(0),$$

$$M^q E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \to -p k_q f_{uu}(0)^{-1} f_q(0).$$

Proof. Proofs of the asymptotic forms for these term are similar, and we will derive the results for $Z_{V1}^* H^{-1} R' J_0^{-1} R H^{-1} Z_{V1}$ and $Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0$ here. First,

$$\begin{aligned}
&E[Z_{V1}^* H^{-1} R' J_0^{-1} R H^{-1} Z_{V1}] \\
&\quad = E\left[\frac{1}{n}\sum_s I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s\right] H^{-1} R' J_0^{-1} R H^{-1} \left[\frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s\right] \\
&\quad \sim E\left[\frac{1}{n^2}\sum_s w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2 H^{-1} R' J_0^{-1} R H^{-1} w_x(\lambda_s)\right] \\
&\quad = E\left\{tr\left[\frac{1}{n^2}\sum_s w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2 H^{-1} R' J_0^{-1} R H^{-1} w_x(\lambda_s)\right]\right\} \\
&\quad = E\left\{tr\left[H^{-1} R' J_0^{-1} R H^{-1} \frac{1}{n^2}\sum_s w_x(\lambda_s) w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2\right]\right\} \\
&\quad = E\left\{tr\left[H^{-1} R' J_0^{-1} R H^{-1} \frac{1}{n^2}\sum_s I_x(\lambda_s) I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2\right]\right\} \\
&\quad \sim \frac{1}{m} E\left\{tr\left[H^{-1} R' J_0^{-1} R H^{-1} \frac{1}{n^2}\sum_s I_x(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]\right\} \\
&\quad \sim \frac{1}{m} E\{tr[H^{-1} R' J_0^{-1} R]\} = \frac{1}{m} E\{tr[R H^{-1} R' J_0^{-1}]\} = \frac{p}{m}.
\end{aligned}$$

Similarly,

$$E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} Z_{V2}] \sim \frac{p}{m},$$

$$E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m},$$

$$E[Z_0^* H^{-1} D_{V1} H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} Z_0] \sim \frac{p}{m},$$

$$E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} D_{V1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m},$$

$$E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m},$$

$$E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} Z_0] \sim \frac{p}{m},$$

$$\mathrm{E}[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} D_{V1} H^{-1} Z_0] \sim \frac{p}{m},$$

$$\mathrm{E}[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} Z_0] \sim \frac{p}{m},$$

and

$$\begin{aligned}
&\mathrm{E}[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0] \\
&\quad = \mathrm{E}\left[\frac{1}{n}\sum_s I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right] H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} \\
&\qquad \times \left[\frac{1}{n}\sum_s I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right] \\
&\sim \ \mathrm{E}\left[\frac{1}{n^2}\sum_s w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-2} H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} w_x(\lambda_s)\right] \\
&= \mathrm{E}\left\{ tr \left[\frac{1}{n^2}\sum_s w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-2} H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} w_x(\lambda_s)\right]\right\} \\
&= \mathrm{E}\left\{ tr \left[H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} \frac{1}{n^2}\sum_s w_x(\lambda_s) w_x(\lambda_s)^* I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-2}\right] \right. \\
&\sim \ \mathrm{E}\left\{ tr \left[H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} \frac{1}{n^2}\sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]\right\} \\
&= \mathrm{E}\left\{ tr \left[H^{-1} R' J_0^{-1} R H^{-1} \left(-\frac{1}{n^2}\sum_s I_{xx}(\lambda_s) f_q(\lambda_s) f_{uu}(\lambda_s)^{-2} M^{-q} k_q\right)\right]\right\} \\
&\sim \ -\frac{k_q}{M^q} f_{uu}(0)^{-1} f_q(0) \mathrm{E}\left\{ tr \left[\left(\int B_x B_x'\right)^{-1} R' \left[R\left(\int B_x B_x'\right)^{-1} R'\right]^{-1} \right.\right. \\
&\qquad \left.\left. \times R\left(\int B_x B_x'\right)^{-1} \left(\int B_x B_x'\right)\right]\right\} \\
&\sim \ -\frac{k_q}{M^q} f_{uu}(0)^{-1} f_q(0) \mathrm{E}\left\{ tr \left[R\left(\int B_x B_x'\right)^{-1} R' \left[R\left(\int B_x B_x'\right)^{-1} R'\right]^{-1}\right]\right\} \\
&\sim \ -\frac{p k_q}{M^q} f_{uu}(0)^{-1} f_q(0),
\end{aligned}$$

and

$$\mathrm{E}[Z_{B1}^* H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim -\frac{p k_q}{M^q} f_{uu}(0)^{-1} f_q(0),$$

$$\mathrm{E}[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim -\frac{p k_q}{M^q} f_{uu}(0)^{-1} f_q(0). \qquad \square$$

References

Barten, A.P., 1969. Maximum likelihood estimation of a complete system of demand equations. European Economic Review 1, 7–73.

Bera, A.K., Byron, R.P., Jarque, C.M., 1981. Further Evidence on Asymptotic Tests for homogeneity and symmetry in large demand systems. Economics Letters 8, 101–105.

Brillinger, D.R., 1980. Time Series: Data Analysis and Theory. Holt, Rinehart and Winston, New York.

Byron, R.P., 1970. The restricted Aitken estimation of sets of demand equations. Econometrica 38, 816–830.

Corbae, D., Ouliaris, S., Phillips, P.C.B., 1994. A reexamination of the consumption function using frequency domain regressions. Empirical Economics 19, 595–609.

Corbae, D., Ouliaris, S., Phillips. P.C.B., 2002. Band spectral regression with trending data. Econometrica, forthcoming.

Deaton, A.S., 1974. The analysis of consumer demand in the United Kingdom, 1900–1970. Econometrica 42, 341–367.

Engle, R., 1974. Band spectrum regression. International Economic Review 15 (1), 1–11.

Engle, R., Granger, C., 1987. Cointegration and error correction: representation, estimation, and testing. Econometrica 55, 251–276.

Hannan, E.J., 1960. Time Series Analysis. Methuen, New York.

Hannan, E.J., 1963a. Regression for time series. In: Rosenblatt, M. (Ed.), Time Series Analysis. Wiley, New York.

Hannan, E.J., 1963b. Regression for time series with errors of measurement. Biometrika 50, 293–302.

Hannan, E.J., 1970. Multiple Time Series. Wiley, New York.

Hannan, E.J., 1971. Nonlinear time series regression. Journal of Applied Probability 8, 767–780.

Hannan, E.J., 1973. Central limit theorems for time series regression. Zeitschrift für Wahrscheinlichkeitsheorie und verwandte Gebiete 26, 157–170.

Laitinen, K., 1978. Why is demand homogeneity so often rejected?. Economic Letters 1, 187–191.

Linton, O.B., 1995a. Second order approximation in the partially linear regression model. Econometrica 63 (5), 1079–1112.

Linton, O.B., 1995b. Second order approximation in a linear regression with heteroskedasticity of unknown form. Econometric Reviews 15, 1–32.

Lluch, C., 1971. Consumer demand functions, Spain, 1958-1964. European Economic Review 2, 277–302.

Meisner, J.F., 1979. The sad fate of the asymptotic Slutsky symmetry test for large systems. Economics Letters 2, 231–233.

Naimark, M.A., 1960. Normed Rings. Noordhoff, Groningen.

Phillips, P.C.B., 1977. A general theorem in the theory of asymptotic expansion as approximations to the finite sample distributions of econometric estimators. Econometrica 45, 463–485.

Phillips, P.C.B., 1987. Asymptotic expansions in nonstationary vector autoregressions. Econometric Theory 3, 45–68.

Phillips, P.C.B., 1991. Spectral regression for cointegrated time series. In: Barnett, W., Powell, J., Tauthen, G. (Eds.), Nonparametric and Semiparametric Methods in Econometrics and Statistics. Cambridge University Press, Cambridge.

Phillips, P.C.B., Park, J.Y., 1988. Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors. Journal of the American Statistical Association 83, 111–115.

Phillips, P.C.B., Solo, V., 1992. Asymptotics for linear processes. Annals of Statistics 20, 971–1001.

Phillips, P.C.B., 2000. Unit root log periodogram regression. Cowles Foundation Discussion Paper No. 1244, Yale University.

Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press, New York.

Robinson, P.M., 1976. Fourier estimation of continuous time models. In: Berstrom, A.R. (Ed.), Statistical Inference in Continuous Time Economic Models. North-Holland, Amsterdam, pp. 215–266.

Robinson, P.M., 1991. Automatic frequency domain inference on semiparametric and nonparametric models. Econometrica 59, 755–786.

Rothenberg, T.J., 1984a. Approximating the distributions of econometric estimators and test statistics. In: Griliches, Z., Intriligator, M. (Eds.), Handbook of Econometrics, Vol. 2. North Holland, New York (Chapter 14).

Rothenberg, T.J., 1984b. Approximate normality of generalized least squares estimators. Econometrica 52, 811–825.

Sargan, J.D., 1976. Econometric estimators and the Edgeworth expansion. Econometrica 44, 421–448.

Sargan, J.D., Mikhail, W.M., 1971. A general approximation to the distribution of instrumental variables estimates. Econometrica 39, 131–169.

Walker, A.M., 1965. Some asymptotic results for the periodogram of a stationary time series. Journal of Australian Mathematical Society 5, 107–128.

Whittle, P., 1951. Hypothesis testing and time series analysis. Almquist and Wiksell, Uppsala.

Xiao, Z., Phillips, P.C.B., 1998. Higher order approximations for frequency domain time series regression. Journal of Econometrics 86, 297–336.