COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1192

HIGHER ORDER APPROXIMATIONS FOR WALD STATISTICS IN COINTEGRATING REGRESSIONS

Zhijie Xiao and Peter C. B. Phillips

August 1998

Higher Order Approximations for Wald Statistics in Cointegrating Regressions*

Zhijie Xiao
Department of Economics
University of Illinois at Urbana-Champaign

and

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University

June 16, 1998

Abstract

Asymptotic expansions are developed for Wald test statistics in cointegrating regression models. These expansions provide an opportunity to reduce size distortion in testing by suitable bandwidth selection, and automated rules for doing so are calculated. Band spectral regression methods and tests are also considered. In such cases, it is shown how the effects of nonstationarity that dominate low frequency limit behaviour also carry over to high frequency asymptotics, with consequential effects on bandwidth rules.

*This paper was keyboarded by the authors in Scientific Word 2.5. We would like to thank Roger Koenker for interesting discussion. Phillips thanks the NSF for research support under grant no. SBR94-22922 and Xiao thanks the Cowles Foundation for hospitality during a visit in May - June 1998.

1. Introduction

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

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

where u_t is a stationary process with zero mean and continuous spectral density $f_{uu}(\lambda)$. The regressor x_t may be a vector of either stationary or unit root nonstationary time series. Results for the stationary case are given here, but the main focus of the paper is on the nonstationary case. In this event, if x_t is an integrated process of order one (or 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).

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

$$w_y(\lambda_s) = \beta' w_x(\lambda_s) + w_u(\lambda_s),\ \lambda_s = 2\pi s/T,\ s = 0, 1, ..., n-1, \tag{1.2}$$

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 $\lambda_s = 2\pi s/n$, $s = 0, 1, ..., n-1$,defined by $w_a(\lambda_s) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda_s}$. If the residual process in (1.1), u_t, is stationary with continuous spectrum bounded away from the origin, the residuals of regression (1.2), $w_u(\lambda_t)$, 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 semiparametric 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.

Regression analysis in the frequency domain was introduced by Hannan (1963), following ideas of Whittle (1953), and extended to nonlinear models by Hannan (1971) and Robinson (1972). 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 Phillips, Ouliaris and Corbae (1997) has explored spectral regression methods in the presence of deterministic trends.

This 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; Llush 1971; and Deaton 1974 among others). 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; and Bera, Byron and Jarque 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 1976; Phillips 1978; Phillips, 1983; Phillips and Park 1988; Rothenberg 1984a,b; and Linton 1995a, b among others). 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 useable 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. 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 spectral density $f_{uu}(\lambda)$. 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 are also given 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 in cointegrated systems 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 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 proofs are given in the Appendix in Section 9.

2. Background and Assumptions

To develop higher order asymptotics for the Wald statistic in regression (1.1), it is convenient for us to make the following assumptions on x_t and u_t :

ASSUMPTION 1: *x_t is an I(1) process satisfying $\Delta x_t = v_t$, initialized at $t = 0$ by any $O_p(1)$ random variables.*
ASSUMPTION 2: *v_t and u_t are independent stationary and ergodic $k-$ vector and scalar time series with zero mean, finite second moments and continuous spectral densities $f_{vv}(\lambda)$ and $f_{uu}(\lambda)$, which are positive at the origin. The vector x_t and partial sums of u_t both satisfy invariance principles with independent limit processes, so that, as $n \to \infty$, $n^{-1/2}x_{[Tr]} \Rightarrow B_x(r) \equiv BM(2\pi f_{vv}(0))$, a vector Brownian motion of dimension k with covariance matrix $2\pi f_{vv}(0)$, and $n^{-1/2}\sum_{t=1}^{[Tr]} u_t \Rightarrow B_u(r) \equiv BM(2\pi f_{uu}(0))$.*

The matrix representation of regression (1.2) can be written as

$$W_y = W_x\beta + W_u, \tag{2.1}$$

where W_y, W_x and W_u are vectors of $w_y(\lambda_s)$, $w_x(\lambda_s)$, $w_u(\lambda_s)$, and can be written as Uy, Ux, and Uu, with $U = \exp[i(\frac{2\pi}{n})\tau g']/\sqrt{n}$ being the $n \times n$ Fourier matrix, $\tau' = [0, 1, ..., n-1]$, $g' = [1, 2, ..., n]$, $y = (y_1, ..., y_n)'$, $x = (x_1, ..., x_n)', u = (u_1, ..., u_n)'$, and $U^*U = I_n$, where the affix * indicates transposition combined with complex conjugation. Under Assumption 2, the covariance matrix of the residual in (2.1) is asymptotically diagonal since the discrete Fourier transforms of u_t are asymptotically uncorrelated. Thus, the frequency domain efficient estimator of β can be obtained based on weighted averages of periodogram estimates at the fundamental frequencies, viz

$$\begin{aligned}\widehat{\beta} &= \left[W_x^* \widehat{\Sigma}^{-1} W_x\right]^{-1} \left[W_x^* \widehat{\Sigma}^{-1} W_y\right] \\ &= \left[\sum_s I_{xx}(\lambda_s) \widehat{f}_{uu}^{-1}(\lambda_s)\right]^{-1} \left[\sum_s I_{xy}(\lambda_s) \widehat{f}_{uu}^{-1}(\lambda_s)\right],\end{aligned}$$

where $I_{xx}(\lambda_s) = w_x(\lambda_s) w_x(\lambda_s)^*$, $I_{xy}(\lambda_s) = w_x(\lambda_s) w_y(\lambda_s)^*$, $\widehat{\Sigma}^{-1} = diag[\cdots\cdot, \widehat{f}_{uu}(\lambda_s)^{-1}, \cdots]$, and $\widehat{f}_{uu}(\lambda_s)^{-1}$ are nonparametric spectral density estimators.

We are interested in testing the linear hypothesis

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

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

$$W = (R\widehat{\beta} - r)' \left[R(W_x^* \widehat{\Sigma}^{-1} W_x)^{-1} R'\right]^{-1} (R\widehat{\beta} - r). \tag{2.3}$$

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 summarizes the limit theory of the estimator $\widehat{\beta}$ and the corresponding Wald statistic (2.3) in the nonstationary case.

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

(1) $n(\widehat{\beta} - \beta) \to \left[\int B_x B_x'\right]^{-1} \left[\int B_x dB_u\right] = MN\left(0, 2\pi f_{uu}(0) \left[\int B_x B_x'\right]^{-1}\right).$

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

Remark 1: As the Lemma shows, the asymptotic distribution of the regression estimator is mixed normal with matrix mixing variate $2\pi f_{uu}(0) \left[\int B_x B_x'\right]^{-1}$ that depends on 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 $W_x^* \Sigma^{-1} W_x$. 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, and this helps to partially 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 nonparametric estimator for the spectral density of the residual process u_t, viz

$$\begin{aligned}
\widehat{f}_{uu}(\lambda_s) &= \frac{1}{m}\sum_{\lambda_j\in B(\lambda_s),\lambda_j\neq\lambda_s} K(\lambda_j-\lambda_s)\widehat{w}_u(\lambda_j)\widehat{w}_u(\lambda_j)^* \\
&= \frac{1}{m}\sum K(\lambda_j-\lambda_s)\widehat{I}_{uu}(\lambda_j) \\
&= \sum_{j\neq s}\omega_{sj}\widehat{I}_{uu}(\lambda_j),
\end{aligned}$$

where $B(\lambda_s) = \{\omega : \lambda_s - \frac{\pi}{2M} < \omega \leq \lambda_s + \frac{\pi}{2M}\}$ is a frequency band of width π/M centered on $\lambda_s = 2\pi s/T$. Let $m = [T/2M]$, where $[\cdot]$ signifies integer part. Then, each band B_t contains m fundamental frequencies λ_s. $K(\cdot)$ is a spectral window and satisfies conventional properties of being a real, even function with $\frac{1}{m}\sum_{\lambda_j\in B(\omega)} K(\lambda_j-\omega) = 1$. The corresponding lag window is $k(\frac{h}{M}) = \frac{1}{m}\sum_{\lambda_s\in B(\omega)} K(\lambda_s-\omega)e^{-ih(\lambda_s-\omega)}$. Many candidate kernel functions are available and are discussed in standard texts of spectral analysis (e.g., Hannan 1970; Brillinger 1980; and Priestley 1981). The periodogram ordinates $\widehat{I}_{uu}(\lambda_s)$ are calculated using consistent estimates of the residuals. In our analysis, we use the residuals from an OLS regression on (1.2), so that $\widehat{w}_u(\lambda_s) = w_y(\lambda_s) - \widehat{\beta}_{OLS}w_x(\lambda_s)$.

It is convenient to decompose the error term in the nonparametric spectral density estimator $\widehat{f}_{uu}(\lambda_s)$ 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

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

where $B_s = \sum_{j\neq s}\omega_{sj}[f_{uu}(\lambda_j) - f_{uu}(\lambda_s)]$, $V_s = \sum_{j\neq s}\omega_{sj}[I_{uu}(\lambda_j) - f_{uu}(\lambda_s)]$, $P_s = \sum_{j\neq s}\omega_{sj}[\widehat{I}_{uu}(\lambda_j) - I_{uu}(\lambda_s)]$. 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 completeness.

Lemma 2:

$$B_t \sim -M^{-q}k_q f_q(\lambda_t),$$
$$\sqrt{m}V_t \overset{d}{\sim} N(0, \frac{1}{2}\int_{-\infty}^{\infty} k(x)^2 dx f_{uu}^2(\lambda_t)),$$

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

Most commonly used kernel functions have the property that $q = 2$ and thus the bias term is of order $O(M^{-2})$. For example, the commonly used smoothed periodogram estimate corresponds to the Daniell kernel that $k(x) = \sin(\pi x/2)/(\pi x/2)$, and

$$K(\lambda) = \begin{matrix} \pi, \text{ for } |\lambda| \leq \pi/2, \\ 0, \text{ otherwise.} \end{matrix}$$

Since our interest is primarily on the analysis of second order effect, without loss of generality, in what follows, we simply use the smoothed periodogram so that $a = \frac{1}{2}\int_{-\infty}^{\infty}k(x)^2dx = 1$, and $\sqrt{m}V_t \overset{d}{\sim} N(0, f_{uu}^2(\lambda_t))$.

In spectral regression for stationary time series, the error term coming from the preliminary estimation of $w_u(\lambda_s)$ is of smaller order of magnitude and thus can be dropped. However, different results arise in cointegrating regression models. Because the spectral power of an integrated process is heavily concentrated at the origin, errors from the preliminary estimation are amplified around zero frequency and thus enter the second order effect. The order of magnitude for P_t

is summarized in the following Lemma.

Lemma 3:

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

and

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

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 $m^{-\frac{1}{2}}$ and M^{-q}. In these expansions, terms that are of order $O_p(m^{-1})$ or $O_p(M^{-2q})$ 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, $\widehat{f}_{uu}(\lambda_t)^{-1}$, can be obtained based on (3.1) and Lemma 2 and Lemma 3. The results are given in Lemma 4 and readers are referred to Xiao and Phillips (1998) for details of the derivation.

Lemma 4:

$$\begin{aligned}\widehat{f}_{uu}(\lambda_t)^{-1} &= f_{uu}(\lambda_t)^{-1} - \overline{f}_{uu}(\lambda_t)^{-2}V_t + \overline{f}_{uu}(\lambda_t)^{-3}V_t^2 - f_{uu}(\lambda_t)^{-2}B_t \\ &\quad + f_{uu}(\lambda_t)^{-3}B_t^2 - f_{uu}(\lambda_t)^{-2}P_t + \text{ higher order terms,} \qquad (3.2)\end{aligned}$$

where $\overline{f}_{uu}(\lambda_s) = m^{-1}\sum_{j\neq s} K(\lambda_j - \lambda_s) f_{uu}(\lambda_j)$.

The following Lemma gives some useful limiting results for periodogram averages of I(1) processes and is based on results in Phillips (1991). The results will be used in the analysis of higher order asymptotics in later sections.

Lemma 5:

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

$$\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}^{-1}(\lambda_s) \xrightarrow{d} \frac{1}{2\pi}f_{uu}(0)^{-1}\int_0^1 B_x B_x',$$

$$\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}^{-2}(\lambda_s)f_q(\lambda_s) \xrightarrow{d} \frac{1}{2\pi}f_{uu}(0)^{-2}f_q(0)\int_0^1 B_x B_x',$$

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

REMARK 2: As pointed out in Phillips (1991), contributions from the zero-frequency ordinates dominate the limit behavior of these periodogram averages since the spectral power of x_t is concentrated at the origin. The limits are characterized in terms of quadratic functionals of the Brownian motion weak limit of $n^{-\frac{1}{2}}x_{[n\cdot]}$.

4. The Wald Expansion in the Univariate Case

We start the development with the scalar case to simplify the derivations and illustrate the main results. The multivariate case will be treated later in Section 6. 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 $W = \left(\widehat{\beta} - \beta_0\right)' \left[W_x^* \widehat{\Sigma}^{-1} W_x\right] \left(\widehat{\beta} - \beta_0\right)$. Under the null hypothesis that $\beta = \beta_0$,

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

and

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

where

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

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

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

If we expand $\widehat{H}^{-1}$ around H^{-1} to the third term in a geometric expansion, we get

$$\widehat{H}^{-1} = H^{-1} - H^{-2}D + H^{-3}D^2 + R_1, \tag{4.5}$$

where $D = \widehat{H} - H$, and $R_1 = \widehat{H}^{-1}H^{-3}D^3$ is the remainder term. We then decompose D and Z as follows:

$$\begin{aligned} D &= -\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f^*_{uu}(\lambda_s)^{-2}V_s + \frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f^*_{uu}(\lambda_s)^{-3}V_s^2 \\ &\quad -\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-2}B_s + \frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-3}B_s^2 \\ &\quad -\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-2}P_s + \text{higher order terms} \\ &= -D_{V1} + D_{V2} - D_{B1} + D_{B2} - D_{P1} \\ &\quad + \text{higher order terms} \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} Z &= \frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1} - \frac{1}{n}\sum_s I_{xu}(\lambda_s)f^*_{uu}(\lambda_s)^{-2}V_s \\ &\quad +\frac{1}{n}\sum_s I_{xu}(\lambda_s)f^*_{uu}(\lambda_s)^{-3}V_s^2 - \frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}B_s \\ &\quad +\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-3}B_s^2 - \frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}P_s \\ &\quad + \text{higher order terms} \\ &= Z_0 - Z_{V1} + Z_{V2} - Z_{B1} + Z_{B2} - Z_{P1} \\ &\quad + \text{higher order terms} \end{aligned} \tag{4.7}$$

The following Lemma gives the orders of magnitude of the terms in these expressions.

Lemma 6:

$$
\begin{aligned}
Z_0 &= O_p(1), Z_{V1} = O_p(m^{-1/2}), Z_{V2} = O_p(m^{-1}), \\
Z_{B1} &= O_p(M^{-q}), Z_{B2} = O_p(M^{-2q}), Z_{P1} = O_p(m^{-1}),
\end{aligned}
$$

and

$$
D_{V1} = O_p(n^{-1/2}), D_{V2} = O_p(m^{-1}), D_{B1} = O_p(M^{-q}), D_{B2} = O_p(M^{-2q}), D_{P1} = O_p(m^{-1}).
$$

Using (4.5), (4.6), and (4.7) in (4.1) and collecting terms up to $O_p(m^{-1} + M^{-2q})$, we obtain a formal[1] moment expansion for $E[W]$, which is summarized in the following Theorem.

Theorem 7:

$$
\begin{aligned}
E[W] &= Q_0 + m^{-1}Q_1 + M^{-q}Q_2 + M^{-2q}Q_3 + o_p(m^{-1} + M^{-2q}) \\
&= \overline{W} + o_p(m^{-1} + M^{-2q}),
\end{aligned}
$$

where Q_0, Q_1, Q_2 and Q_3 are $O(1)$ quantities defined as

$$
\begin{aligned}
Q_0 &= E[H^{-1}Z_0^* Z_0], \\
Q_1 &= Q_{11} + Q_{12}, \\
Q_{11} &= m\{E[H^{-1}Z_{V1}^* Z_{V1}] + 2E[H^{-1}Z_0^* Z_{V2}] - E[H^{-2}D_{V2}Z_0^* Z_0]\}, \\
Q_{12} &= m\{E[H^{-1}D_{P1}Z_0^* Z_0] - 2E[H^{-1}Z_0^* Z_{P1}]\} \\
Q_2 &= M^q\{E[H^{-2}D_{B1}Z_0^* Z_0] - 2E[H^{-1}Z_0^* Z_{B1}]\}, \\
Q_3 &= M^{2q}\{E[H^{-1}Z_{B1}^* Z_{B1}] + E[H^{-3}D_{B1}^2 Z_0^* Z_0] + 2E[H^{-1}Z_0^* Z_{B2}] \\
&\quad -2E[H^{-2}D_{B1}Z_0^* Z_{B1}] - E[H^{-2}D_{B2}Z_0^* Z_0]\},
\end{aligned}
$$

and $\overline{W} = Q_0 + m^{-1}Q_1 + M^{-q}Q_2 + M^{-2q}Q_3$ is the truncated expectation of the Wald statistic.

[1]Here and elsewhere in the paper we proceed in a conventional way with formal moment approximations. The expressions given are formally moments of the approximating expansions rather than formal asymptotic expansions of the moments of the statistics.

REMARK 3: 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 4: The higher order effects (up to $O_p(m^{-1} + M^{-2q})$) include: Q_{11}, a variance effect; Q_{12}, second order effect coming from the preliminary estimation; Q_2, a bias effect; and Q_3, a squared bias term. Here, the bias effect Q_2 plays an important role in the second order term and dominates the squared bias term Q_3 in order of magnitude.

REMARK 5: 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 $f_{uu}(\omega)$. The fourth term, Q_3, 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 $M^{-q}\left|Q_2\right|$, 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 the second order efficiency.

When Q_2 is positive, both terms in the second order effect of $E[W]$ are positive. As a result, at least to the second order, the Wald statistic (2.3) 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

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

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

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

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

Using the results in Lemma 4, we have the following asymptotics:

LEMMA 8:

(1) $mE[H^{-1} Z_0^* Z_{V2}] \to 1$;

(2) $mE[H^{-2} D_{V2} Z_0^* Z_0] \to 1$;

(3) $mE[H^{-1} Z_{V1}^* Z_{V1}] \to 1$;

(4) $mE[H^{-1} Z_0^* Z_{P1}] \to -1$;

(5) $mE[H^{-2} D_{P1} Z_0^* Z_0] \to -1$;

(4) $M^q E[H^{-1} Z_0^* Z_{B1}] \to k_q f_{uu}(0)^{-1} f_q(0)$;

(5) $M^q E[H^{-2} D_{B1} Z_0^* Z_0] \to k_q f_{uu}(0)^{-1} f_q(0)$;

(6) $M^{2q} E[H^{-1} Z_{B1}^* Z_{B1}] \to k_q^2 f_{uu}(0)^{-2} f_q(0)^2$;

(7) $M^{2q} E[H^{-3} D_{B1}^2 Z_0^* Z_0] \to k_q^2 f_{uu}(0)^{-2} f_q(0)^2$;

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

(9) $M^{2q}E[H^{-2}D_{B1}Z_0^*Z_{B1}] \to k_q^2 f_{uu}(0)^{-2} f_q(0)^2$;

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

These results lead us to the following approximation for the truncated expected Wald statistic.

THEOREM 9:

$$\overline{W} \sim 1 + \frac{3}{m} + \frac{k_q}{M^q} f_u(0)^{-1} f_q(0) + \frac{k_q^2}{M^{2q}} f_{uu}(0)^{-2} f_q(0)^2.$$

Since the effect at zero frequency dominates, the second order effect in $\overline{W}$ is asymptotically determined by the spectral density of the residual process at the origin. As a result, the size of the hypothesis 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.

$$M = \left(\frac{q k_q f_{uu}(0)^{-1} f_q(0)}{6} \right)^{1/(q+1)} n^{1/(q+1)}. \tag{4.8}$$

If we plug consistent estimates $\widehat{f}_{uu}(0)$ and $\widehat{f}_q(0)$ 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:

$$W_{SO} = W - n^{-q/(q+1)} \left[q^{1/(q+1)} + q^{-q/(q+1)} \right] (6)^{q/(q+1)} \left[k_q \widehat{f}_q(0) / \widehat{f}_{uu}(0) \right]^{1/(q+1)}. \tag{4.9}$$

REMARK 6: If x_t is a stationary process, the higher order asymptotics are different. It can be shown in this case that

$$E[W] \sim 1 + \frac{2a}{m} + \frac{k_q}{M^q} \Omega^{-1}\Gamma_2 + \frac{k_q^2}{M^{2q}} \left[2\Omega^{-1}\Gamma_1 - \Omega^{-2}\Gamma_2^2 \right],$$

where

$$\Omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{xx}(\omega) f_{uu}(\omega)^{-1} d\omega,$$
$$\Gamma_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{xx}(\omega) f_{uu}(\omega)^{-3} f_q(\omega)^2 d\omega,$$
$$\Gamma_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{xx}(\omega) f_{uu}(\omega)^{-2} f_q(\omega) d\omega,$$

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

$$M = \left(\frac{q k_q \Omega^{-1}\Gamma_2}{4a}\right)^{1/(q+1)} n^{1/(q+1)}.$$

Such a formula depends on integrated forms of the spectral density functions, rather than their values at a single point. As a result, we can not simply substitute consistent nonparametric density estimates for a single point into the formula to obtain a bandwidth choice rule. Instead, the usual procedure is to prespecify a parametric model for the error process u_t, estimate these 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. Engle (1974) first applied the band spectral regression method to test the permanent income hypothesis, while Corbae, Ouliaris and Phillips (1994) extended it to cases with stochastic trends, and, in a later work, Phillips, Ouliaris and Corbae (1997) studied band spectral regression with deterministic trends.

Corresponding to a frequency band B_i, we may hypothesize a spectral regression model of the form

$$w_y(\lambda_s) = \beta_i w_x(\lambda_s) + w_u(\lambda_s), \text{ if } \lambda_s \in B_i. \tag{5.1}$$

In this band spectral model, the value of the regression coefficient vector β_i is now frequency dependent. In what follows, we consider a simple band spectral regression model where the regression coefficients have different values on two symmetric frequency bands: a low frequency band $A = [-\omega_0, \omega_0]$, which includes the zero frequency, and a high frequency band $B = [-\pi, \pi] - [-\omega_0, \omega_0]$. As a result, the full model has the form

$$w_y(\lambda_s) = \beta_A w_x(\lambda_s) + w_u(\lambda_s), \text{ if } \lambda_s \in A, \tag{5.2}$$

and

$$w_y(\lambda_s) = \beta_B w_x(\lambda_s) + w_u(\lambda_s), \text{ if } \lambda_s \in B. \tag{5.3}$$

As we have seen in previous sections, due to the dominating effects of the zero frequency components, regression (5.2) (which includes the zero frequency) and (5.3) (which excludes the zero frequency) lead to estimators with differing rates of convergence and limit distributions. It can therefore be anticipated that test statistics arising from these two different regressions will have different second order effects.

The following lemma summarizes the limit distributions for efficient band spectral regression estimators.

LEMMA 10: I*f*

$$\widehat{\beta}_A = \left[\sum_{\lambda_s \in A} I_{xx}(\lambda_s)\widehat{f}_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\sum_{\lambda_s \in A} I_{xy}(\lambda_t)\widehat{f}_{uu}(\lambda_t)^{-1}\right]$$

and

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

then

$$n(\widehat{\beta}_A - \beta_A) \xrightarrow{d} \left[\int B_x B_x'\right]^{-1} \left[\int B_x dB_u\right] = MN\left(0, 2\pi f_{uu}(0)\left[\int B_x B_x'\right]^{-1}\right),$$

$$\sqrt{n}(\widehat{\beta}_B - \beta_B) \xrightarrow{d} N\left(0, \left[\int_B h_x(\omega) f_{uu}(\omega)^{-1} d\omega\right]^{-1}\right), \qquad (5.4)$$

where

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

The asymptotics for $\widehat{\beta}_A$ and its corresponding Wald statistic are the same as those of $\widehat{\beta}$ in Section 4, allowing for the fact that the full band is replaced by the partial band A in Remark 6 (the stationary case) - see Remark 9 below. Thus, the following discussion focuses on $\widehat{\beta}_B$ and its related regression Wald test, for which case the zero frequency is not included. Notice that in the frequency band B, the errors coming from preliminary estimation of $w_u(\lambda_t)$ can be dropped by Lemma 3, and thus do not enter the second order effect. Some useful asymptotic results are summarized in Lemma 11. Because of the absence of the zero frequency, rates of convergence of these periodogram averages are different from those of Section 4.

LEMMA 11:

$$\frac{1}{n}\sum_{\lambda_s \in B} I_{xx}(\lambda_s) \xrightarrow{d} \int_B h_x(\omega) d\omega,$$
$$\frac{1}{n}\sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1} \xrightarrow{d} \Sigma_B,$$
$$\frac{1}{n}\sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} f_q(\lambda_s)^2 \xrightarrow{d} \Lambda_1,$$
$$\frac{1}{n}\sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} f_q(\lambda_s) \xrightarrow{d} \Lambda_2,$$

where

$$\Sigma_B = \int_B h_x(\omega) f_{uu}(\omega)^{-1} d\omega,$$

$$\Lambda_1 = \int_B h_x(\omega) f_{uu}(\omega)^{-3} f_q(\omega)^2 d\omega, \textit{ and}$$
$$\Lambda_2 = \int_B h_x(\omega) f_{uu}(\omega)^{-2} f_q(\omega) d\omega.$$

The Wald statistic for H_0 ($\beta = \beta_0$) in this band spectral regression can be written as $W_B = \Psi^* \widehat{H}_B^{-1} \Psi$ where

$$\widehat{H}_B = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) \widehat{f}_{uu}(\lambda_s)^{-1},$$
$$H_B = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}, \text{ and}$$
$$\Psi = \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) \widehat{f}_{uu}(\lambda_s)^{-1}.$$

The following lemma verifies the orders of magnitude for some higher order components and is useful in the expansion that follows.

LEMMA 12:

$$\Delta_{V1} = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s = O_p(n^{-1/2}),$$
$$\Delta_{V2} = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = O_p(m^{-1}),$$
$$\Delta_{B1} = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = O_p(M^{-q}),$$
$$\Delta_{B2} = \frac{1}{n} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = O_p(M^{-2q}),$$

and

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

$$
\begin{aligned}
\Psi_{V1} &= \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s = O_p(m^{-1/2}), \\
\Psi_{V2} &= \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = O_p(m^{-1}), \\
\Psi_{B1} &= \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = O_p(M^{-q}), \\
\Psi_{B2} &= \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = O_p(M^{-2q}).
\end{aligned}
$$

Performing a similar expansion to that in Section 4, we get the following result for the band spectral regression Wald statistic.

THEOREM 13:

$$
\begin{aligned}
E[W_B] &= \Phi_0 + m^{-1}\Phi_1 + M^{-q}\Phi_2 + M^{-2q}\Phi_3 + o_p(m^{-1} + M^{-2q}) \\
&= \overline{W}_B + o_p(m^{-1} + M^{-2q}),
\end{aligned}
$$

where Φ_0, Φ_1, Φ_2, and Φ_3 are $O(1)$ quantities defined as

$$
\begin{aligned}
\Phi_0 &= E[H_B^{-1}\Psi_0^*\Psi_0], \\
\Phi_1 &= m\{E[H_B^{-1}\Psi_{V1}^*\Psi_{V1}] + 2E[H_B^{-1}\Psi_0^*\Psi_{V2}] - E[H_B^{-2}\Delta_{V2}\Psi_0^*\Psi_0]\}, \\
\Phi_2 &= M^q\{E[H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_0] - 2E[H_B^{-1}\Psi_0^*\Psi_{B1}]\}, \text{ and} \\
\Phi_3 &= M^{2q}\{E[H_B^{-1}\Psi_{B1}^*\Psi_{B1}] + E[H_B^{-3}\Delta_{B1}^2\Psi_0^*\Psi_0] + 2E[H_B^{-1}\Psi_0^*\Psi_{B2}] \\
&\quad -2E[H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_{B1}] - E[H_B^{-2}\Delta_{B2}\Psi_0^*\Psi_0]\},
\end{aligned}
$$

and $\overline{W}_B = \Phi_0 + m^{-1}\Phi_1 + M^{-q}\Phi_2 + M^{-2q}\Phi_3$ is the truncated expected Wald statistic.

Remark 7: This expansion is similar to that in Section 4. However, the quantities in the formula have different definitions and different asymptotics. As discussed before, for hypothesis testing (2.2) and the corresponding 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:

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

Lemma 14 provides some useful asymptotic results parallel to those of Lemma 8.

Lemma 14:

(1) $mH_B^{-1}\Psi_0^*\Psi_{V2} \xrightarrow{d} a,$

(2) $mH_B^{-2}\Delta_{V2}\Psi_0^*\Psi_0 \to a,$

(3) $mH_B^{-1}\Psi_0^*\Psi_0 \to a,$

(4) $M^q H_B^{-1}\Psi_0^*\Psi_{B1} \xrightarrow{d} k_q\Sigma_B^{-1}\Lambda_2,$

(5) $M^q H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_0 \xrightarrow{d} k_q\Sigma_B^{-1}\Lambda_2,$

(6) $M^{2q} H_B^{-1}\Psi_{B1}^*\Psi_{B1} \xrightarrow{d} k_q^2\Sigma_B^{-1}\Lambda_1,$

(7) $M^{2q} H_B^{-3}\Delta_{B1}^2\Psi_0^*\Psi_0 \xrightarrow{d} k_q^2\Sigma_B^{-2}\Lambda_2^2,$

(8) $M^{2q} H_B^{-1}\Psi_0^*\Psi_{B2} \xrightarrow{d} k_q^2\Sigma_B^{-1}{\Lambda_1}^2,$

(9) $M^{2q} H_B^{-2}\Delta_{B1}\Psi_0^*\Psi_{B1} \xrightarrow{d} k_q^2\Sigma_B^{-2}\Lambda_2^2,$

(10) $M^{2q}H_B^{-2}\Delta_{B2}\Psi_0^*\Psi_0 \xrightarrow{d} k_q^2\Sigma_B^{-1}\Lambda_1.$

The following approximation for the truncated expected Wald statistic can then be obtained based on Lemma 14:

THEOREM 15:
$$\overline{W}_B \sim 1 + \frac{2a}{m} + \frac{k_q}{M^q}\Sigma_B^{-1}\Lambda_2 + \frac{k_q^2}{M^{2q}}\left[2\Sigma_B^{-1}\Lambda_1 - \Sigma_B^{-2}\Lambda_2^2\right].$$

REMARK 8: Notice that the frequency band does not include the zero frequency and thus the second order effects in $\overline{W}_B$ are functions of integrals of the spectral density. As a result, we can not obtain a bandwidth selection formula as we did in Section 4 by simply using spectral density estimates at the origin. Instead, a plug-in method or cross validation procedure has to be used to find a bandwidth in practice.

REMARK 9: If x_t is a stationary time series, the expansion of the expected Wald statistic in the band spectral regression model is very similar to the full spectral regression model. The asymptotics for the higher order components have the same formula but integrations are now taken over the frequency band B. Specifically,
$$E[W_B] \sim 1 + \frac{2a}{m} + \frac{k_q}{M^q}\Omega_B^{-1}\Gamma_{B2} + \frac{k_q^2}{M^{2q}}\left[2\Omega_B^{-1}\Gamma_{B1} - \Omega_B^{-2}\Gamma_{B2}^2\right],$$
where
$$\begin{aligned}
\Omega_B &= \frac{1}{2\pi}\int_B f_{xx}(\omega)f_{uu}(\omega)^{-1}d\omega, \\
\Gamma_{B1} &= \frac{1}{2\pi}\int_B f_{xx}(\omega)f_{uu}(\omega)^{-3}f_q(\omega)^2 d\omega, \text{ and} \\
\Gamma_{B2} &= \frac{1}{2\pi}\int_B f_{xx}(\omega)f_{uu}(\omega)^{-2}f_q(\omega)d\omega.
\end{aligned}$$

6. The Multivariate Case

In this section, we consider 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 an $p \times 1$ vector and R is an $p \times k$ matrix. 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

$$W = (\widehat{\beta} - \beta)' R' \left[R(W_x^* \widehat{\Sigma}^{-1} W_x)^{-1} R' \right]^{-1} R(\widehat{\beta} - \beta).$$

For simplicity, we still use notations $\widehat{H}, H$, and Z, which are defined by formulae (4.2), (4.3), (4.4), for their matrix counterpart, then the Wald statistic can be written as

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

The expansion for W in this section follows similar ideas as those in Section 4, and the higher order asymptotics are similar to those in Theorem 9. However, the expansion is now much more complicated because of the multivariate nature of the model. The detailed expansions are given in the appendix and we only state the major steps of the expansions here. Denote $R\widehat{H}^{-1}R'$ by $\widehat{J}_0$, and $RH^{-1}R'$ by J_0, we expand $\widehat{J}_0^{-1}$ around J_0 to the third order and get the following expansion for W,

$$Z^* \widehat{H}^{-1} R' \left[J_0^{-1} - J_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1} + J_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1} + R_2 \right]^{-1} R \widehat{H}^{-1} Z, \tag{6.1}$$

where $R_2 = \widehat{J}_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1}(\widehat{J}_0 - J_0)J_0^{-1}$. Notice that $\widehat{J}_0 - J_0 = R(\widehat{H}^{-1} - H^{-1})R'$, $\widehat{H}^{-1}$ can be expanded around H as $\widehat{H}^{-1} = H^{-1} - H^{-1}DH^{-1} + H^{-1}DH^{-1}DH^{-1} + R_1$, 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 $O_p(m^{-1} + M^{-q})$, we get the following expansion for the expected Wald statistic.

Theorem 16:

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

where S_0, S_1 and S_2 are $O(1)$ quantities defined as

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

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

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

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

Remark 10: 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 goes into both the first and the second order asymptotics. The following Lemma gives asymptotic results for the higher order terms in the expansion.

Lemma 17:

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

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

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

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

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

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

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

Thus, the truncated expected Wald statistic $\overline{W}_G$ has the following approximation, which is similar to that in Theorem 9.

THEOREM 18:

$$\overline{W}_G \sim p + \frac{3p}{m} + \frac{pk_q}{M^q} f_{uu}^{-1}(0) f_q(0).$$

REMARK 11: 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 12: When x_t is a stationary process, the zero frequency no longer has the dominant effect and the value of spectral density elsewhere goes into the asymptotics. As a result, certain terms can not 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 expected value of Wald statistic in this case has the following expansion

$$E[W] \sim p + \frac{2pa}{m} + \frac{k_q}{M^q} tr \left\{ \left(R\Omega^{-1} R' \right)^{-1} R\Omega^{-1} \Gamma_2 \Omega^{-1} R' \right\},$$

where Ω and Γ_2 are $k \times k$ matrices of integration of spectral densities defined by the same formulae as those in Remark 6.

REMARK 13: Expansions for the Wald statistic in band spectral regression models can be derived in a similar way.

7. Monte Carlo Results

We conducted a small Monte Carlo experiment to evaluate the second order theory for Wald tests in cointegrating regressions. 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

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

where v_t and ε_t are both iid N(0,1) variates and are independent of each other. Two values of autoregressive coefficient, $\alpha = 0.3$ and $\alpha = 0.6$ 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, $T = 64$, $T = 128$, $T = 256$, were used in the experiment. The number of replications was 10,000 for each case. The Daniell window was used in the nonparametric density estimation. The size properties of the 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.

We calculated the empirical sizes of these test statistics from the Monte Carlo rejection rates based on asymptotic critical value at the 5% level. The results are reported in Table 1. The empirical critical values of these tests are also calculated as quantiles in the simulations under the null hypothesis. Table 2 gives these values.

The optimal bandwidth selection criterion yields relatively satisfactory finite sample 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 those 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 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, we can see from these tables that the empirical size decreases and empirical critical values are closer to the asymptotic critical values, corroborating the asymptotic approximations. We also find more size distortion when the autocorrelation in the residual process u_t increases.

Table 1: Empirical Size (nominal size = 5%)

		W_0	W_1	W_2	W_{SO}
$\alpha = 0.3$	$T = 64$	0.1474	0.1587	0.1598	0.1069
$\alpha = 0.3$	$T = 128$	0.1127	0.1213	0.1215	0.0879
$\alpha = 0.3$	$T = 256$	0.0821	0.0849	0.0855	0.0688
$\alpha = 0.6$	$T = 64$	0.1827	0.2005	0.2017	0.1400
$\alpha = 0.6$	$T = 128$	0.1352	0.1445	0.1458	0.1104
$\alpha = 0.6$	$T = 256$	0.0932	0.1041	0.1047	0.0799

Table 2: Empirical Critical Values

		W_0	W_1	W_2	W_{SO}
$\alpha = 0.3$	$T = 64$	7.4858	8.0124	8.0596	6.5098
$\alpha = 0.3$	$T = 128$	5.9016	6.2244	6.2279	5.2868
$\alpha = 0.3$	$T = 256$	4.7959	4.9326	4.9488	4.4086
$\alpha = 0.6$	$T = 64$	9.6044	10.347	10.349	8.6285
$\alpha = 0.6$	$T = 128$	6.9657	7.3700	7.4274	6.3509
$\alpha = 0.6$	$T = 256$	5.3542	5.5971	5.5990	4.9669

8. Conclusion

This paper develops higher order expansions for Wald test statistics in efficient, frequency domain semi-parametric cointegrating regression models. 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) process, 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 band spectral regression models. When the frequency band under consideration does not contain the 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 integrals of the spectral density. 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.

9. Appendix: Proofs

9.1. Lemma 1:

As in Phillips (1991), we find

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

$$n^{-1}\sum_s I_{xu}(\lambda_s)\widehat{f}_{uu}(\lambda_s)^{-1} \sim n^{-1}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1} \xrightarrow{d} [2\pi f_{uu}(0)]^{-1}\int B_x dB_u.$$

9.2. Lemma 2:

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

9.3. Lemma 3:

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

$$P_t = -2(\widehat{\beta}-\beta)\sum_{s\neq t}\omega_{ts}I_{xu}(\lambda_s) + (\widehat{\beta}-\beta)^2\sum_{s\neq t}\omega_{ts}I_{xx}(\lambda_s)$$

and

$$n(\widehat{\beta}-\beta) \Rightarrow \left[\int B_x B_x'\right]^{-1} \int B_x dB_u.$$

It can be verified that

$$\begin{aligned}
& \sum_s \omega_{ts} I_{xu}(\lambda_s) \\
= \ & \widehat{f}_{xu}(\lambda_t) \\
= \ & \frac{1}{2\pi} \sum_h k\left(\frac{h}{M}\right) C_{xu}(h) e^{i\lambda_t h} \\
\sim \ & \frac{1}{2\pi} \sum_h k\left(\frac{h}{M}\right) e^{i\lambda_t h} \int B_x dB_u \\
= \ & M \left\{\frac{1}{2\pi M} \sum_h k\left(\frac{h}{M}\right) e^{i\lambda_t h}\right\} \int B_x dB_u \\
= \ & M \cdot K(M\lambda_t) \int B_x dB_u \\
= \ & \begin{matrix} O_p(M), \text{ if } |\lambda_t| \leq \pi/(2M), \\ 0, \text{ otherwise.} \end{matrix}
\end{aligned}$$

and

$$\begin{aligned}
& \sum_s \omega_{ts} I_{xx}(\lambda_s) \\
= \ & \widehat{f}_{xx}(\lambda_t) \\
= \ & \frac{1}{2\pi} \sum_h k\left(\frac{h}{M}\right) C_{xx}(h) e^{i\lambda_t h} \\
\sim \ & \frac{n}{2\pi} \sum_h k\left(\frac{h}{M}\right) e^{i\lambda_t h} \int B_x B_x' \\
= \ & MnK(M\lambda_t) \int B_x B_x' \\
= \ & \begin{matrix} O_p(nM), \text{ if } |\lambda_t| \leq \pi/(2M), \\ 0, \text{ otherwise} \end{matrix}
\end{aligned}$$

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

$$P_t = -2(\widehat{\beta}-\beta) \sum_{s \neq t} \omega_{ts} I_{xu}(\lambda_s) + (\widehat{\beta}-\beta)^2 \sum_{s \neq t} \omega_{ts} I_{xx}(\lambda_s)$$

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

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

$$\begin{aligned} P_t &= -2(\widehat{\beta}-\beta)\sum_{s\neq t}\omega_{ts}I_{xu}(\lambda_s) + (\widehat{\beta}-\beta)^2\sum_{s\neq t}\omega_{ts}I_{xx}(\lambda_s) \\ &\sim -\frac{1}{2m}K(M\lambda_t)\int dB_u B_x' \left[\int B_x B_x'\right]^{-1}\int B_x dB_u. \end{aligned}$$

■

9.4. Lemma 5:

We provide the proof for the third result, proofs of the other results being very similar. The approach follows that of Phillips (1991).

$$\begin{aligned} & n^{-2}\sum_j I_{xx}(\lambda_j) f_{uu}(\lambda_j)^{-2} f_q(\lambda_j) \\ = \; & n^{-2}\sum_j I_{xx}(\lambda_j)\left[\frac{1}{2\pi}\sum_{g=-\infty}^{\infty} D_g e^{ig\lambda_j}\right]^2 \left[\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}|h|^q\gamma(h)e^{-ih\lambda_j}\right] \\ = \; & n^{-2}\sum_j I_{xx}(\lambda_j)\left[\frac{1}{4\pi^2}\sum_{g,p} D_g D_p e^{ig\lambda_j}e^{ip\lambda_j}\right]\left[\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}|h|^q\gamma(h)e^{-ih\lambda_j}\right] \\ = \; & \frac{1}{8\pi^3}n^{-2}\sum_{h,g,p}|h|^q\gamma(h)D_g D_p\sum_j I_{xx}(\lambda_j)e^{i(p+g-h)\lambda_j} \\ = \; & \frac{1}{8\pi^3}n^{-2}\sum_{h,g,p}|h|^q\gamma(h)D_g D_p\sum_j\left[\frac{1}{2\pi}\sum_r C_{xx}(r)e^{i\lambda_j r}\right]e^{i(p+g-h)\lambda_j} \\ = \; & \frac{1}{16\pi^4}n^{-2}\sum_{h,g,p}|h|^q\gamma(h)D_g D_p\sum_r C_{xx}(r)\sum_j e^{i(r+p+g-h)\lambda_j} \end{aligned}$$

Denote $r+p+g-h=b, \sum e^{i\lambda_j b} = \sum e^{ib(2\pi j/n)} = n$, if $b=0$ or $b=n*l$ for some l, otherwise $\sum e^{i\lambda_j b}=0$. Thus, setting $\underline{r} = h-p-g$ in what follows, we have

$$n^{-2}\sum_j I_{xx}(\lambda_j) f_{uu}(\lambda_j)^{-2} f_q(\lambda_j)$$

$$
\begin{aligned}
&= \frac{1}{16\pi^4} n^{-2} \sum_{h,g,p} |h|^q \gamma(h) D_g D_p C_{xx}(\underline{r}) n \\
&= \frac{1}{16\pi^4} \sum_{h,g,p} |h|^q \gamma(h) D_g D_p \left[n^{-1} C_{xx}(\underline{r}) \right] \\
&\to \frac{1}{2\pi} f_q(0) f_{uu}(0)^{-2} \int B_x B_x'.
\end{aligned}
$$

9.5. Lemma 6

Proof.

$$
\begin{aligned}
D_{V1} &= \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}^*(\lambda_t)^{-2} \sum_{s \neq t} \omega_{ts} I_{uu}(\lambda_s)^c \\
&= \frac{1}{n} \sum_s \left[\sum_t \omega_{ts} \left(\frac{1}{n} I_{xx}(\lambda_t) f_{uu}^*(\lambda_t)^{-2} \right) \right] I_{uu}(\lambda_s)^c \\
&= O_p(n^{-1/2}),
\end{aligned}
$$

Note that $\left[\sum_t \omega_{ts} \left(\frac{1}{n} I_{xx}(\lambda_t) f_{uu}^*(\lambda_t)^{-2}\right)\right] = O_p(1)$, and the $I_{uu}(\lambda_s)^c$ are asymptotically uncorrelated mean zero variates. Thus, conditional on x, D_{V1} is a weighted average of n uncorrelated variates. Next,

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

$$
\begin{aligned}
D_{B1} &= \frac{1}{n^2} \sum_{t=} I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} B_t \\
&\sim \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} M^{-q} k_q f_q(\lambda_t)
\end{aligned}
$$

$$
\begin{aligned}
&\sim M^{-q}k_q \frac{1}{n^2}\sum_t \left[\frac{1}{2\pi}\sum_r C_{xx}(r)e^{-ir\lambda_t}\right]\times \\
&\left[\frac{1}{(2\pi)^2}\sum_{g,h} D_g D_h e^{i(g+h)\lambda_t}\right]\left[\frac{1}{2\pi}\sum_l C_{uu}(l)\mid l\mid^q e^{-il\lambda_t}\right] \\
&\sim M^{-q}k_q \frac{1}{n^2}\left[\frac{1}{(2\pi)^4}\sum_{g,h}\sum_l\sum_r C_{xx}(r)C_{uu}(l)\mid l\mid^q D_g D_h\left(\sum_{t=-n/2+1}^{n/2} e^{i(g+h-l-r)\lambda_t}\right)\right] \\
&\sim M^{-q}k_q \frac{1}{n^2}\left[\frac{1}{(2\pi)^4}\sum_{g,h}\sum_l C_{uu}(l)\mid l\mid^q D_g D_h \sum_r C_{xx}(r)\left(\sum_{t=-n/2+1}^{n/2} e^{i(g+h-l-r)\lambda_t}\right)\right] \\
&\sim M^{-q}k_q \frac{1}{n^2}\left[\frac{1}{(2\pi)^4}\sum_{g,h}\sum_l C_{uu}(l)\mid l\mid^q D_g D_h C_{xx}(\underline{r})n\right] \\
&\sim M^{-q}k_q \left[\frac{1}{(2\pi)^4}\sum_{g,h}\sum_l C_{uu}(l)\mid l\mid^q D_g D_h \int B_x B_x'\right] \\
&= M^{-q}k_q \frac{1}{2\pi} f_{uu}(0)^{-2} f_q(0)\int B_x B_x' \\
&= O_p(M^{-q}),
\end{aligned}
$$

and similarly for D_{B2}.

For $Z_0, Z_{V1}, etc.$, we need to check their second moments. Notice that the $w_u(\lambda_t)$ are asymptotically uncorrelated, and

$$
\begin{aligned}
E[\ Z_0]^2 &= E\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)I_{uu}(\lambda_t)f_{uu}(\lambda_t)^{-2} \\
&\sim E\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)f_{uu}(\lambda_t)^{-1} \\
&\sim E\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x B_x',
\end{aligned}
$$

$$
\begin{aligned}
E[Z_{V1}]^2 &= E\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)I_{uu}(\lambda_t)f_{uu}(\lambda_t)^{-4}V_t^2 \\
&\sim E\frac{1}{n^2}\sum_t I_{xx}(\lambda_t)f_{uu}(\lambda_t)^{-3}\frac{1}{m}f_{uu}(\lambda_t)^2
\end{aligned}
$$

$$\sim \quad \frac{1}{m} E \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1}$$
$$\sim \quad \frac{1}{m} \frac{1}{2\pi} f_{uu}(0)^{-1} E \int B_x B_x',$$

$$\begin{aligned} E[Z_{V2}]^2 &= E\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) I_{uu}(\lambda_t) f_{uu}^*(\lambda_t)^{-6} E V_t^4 \\ &\sim \frac{1}{m^2} E \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1} \\ &= O_p(m^{-2}), \end{aligned}$$

$$\begin{aligned} E[Z_{B1}]^2 &= E\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) I_{uu}(\lambda_t) f_{uu}(\lambda_t)^{-4} B_t^2 \\ &\sim E\frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} M^{-2q} k_q^2 f_q(\lambda_t)^2 \\ &\sim M^{-2q} k_q^2 \frac{1}{2\pi} f_{uu}(0)^{-3} f_q(0)^2 E \int B_x B_x \\ &= O_p(M^{-2q}). \end{aligned}$$

Similarly for Z_{B2}.

9.6. Theorem 7

$$W = Z^* \widehat{H}^{-1} Z = Z[H^{-1} - H^{-2} D + H^{-3} D^2 + R] Z$$

Notice that

$$\begin{aligned} Z &= Z_0 - Z_{V1} + Z_{V2} - Z_{B1} + Z_{B2} - Z_{P1} \\ D &= -D_{V1} + D_{V2} - D_{B1} + D_{B2} - D_{P1} \end{aligned}$$

Plug these expansion into (A.1), and drop higher order terms, giving

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

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

Notice that

1. The periodograms are asymptotically uncorrelated
2. A leave-one-out estimator is used so that $w_u(\lambda_t)$ is asymptotically uncorrelated with V_t

It can be verified that $E[H^{-2}D_{V1}Z_0^2]$, $E[H^{-1}Z_0Z_{V1}]$, $E[H^{-1}Z_{V1}Z_{B1}]$, $E[H^{-2}D_{B1}Z_0Z_{V1}]$ are of smaller order of magnitude, giving the result of Theorem 6.

9.7. Lemma 8

$$
\begin{aligned}
E[H^{-1}Z_0^2] &= \left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^2 \\
&\sim \left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-2}\right] \\
&\to 1,
\end{aligned}
$$

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

$$
H \to \frac{1}{2\pi}f_{uu}(0)^{-1}\int B_xB_x.
$$

For

$$
\begin{aligned}
E[Z_{V1}]^2 &= E\frac{1}{n^2}\sum_t\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_sI_{xu}(\lambda_t)f_{uu}(\lambda_t)^{-2}V_t \\
&\sim E\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-4}I_{uu}(\lambda_s)V_s^2.
\end{aligned}
$$

Notice that $I_{uu}(\lambda_s)$ and V_s are asymptotically uncorrelated by the leave-one out property, and

$$\begin{aligned} E[Z_{V1}]^2 &\sim \frac{1}{n^2}\sum_s EI_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-4}E[I_{uu}(\lambda_s)]E[V_s]^2 \\ &\sim E\frac{a}{mn^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1} \\ &\sim E\{\frac{a}{m}\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_xB_x\}, \end{aligned}$$

$$\begin{aligned} mE[H^{-1}{Z_{V1}}^2] &= mE\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n}\sum_s I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-2}V_s\right]^2 \\ &\sim mE\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{n^2}\sum_s I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-4}V_s^2\right] \\ &\sim a=1, \end{aligned}$$

$$\begin{aligned} mE[Z_{V2}^*H^{-1}Z_0] &\sim mEH^{-1}\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-3}V_s^2\right] \\ &\sim mEH^{-1}\left[\frac{1}{n}\sum I_{xx}(\lambda_s)I_{uu}(\lambda_s)f_{uu}(\lambda_s)^{-4}V_s^2\right] \\ &\sim a=1. \end{aligned}$$

Similarly

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

$$\begin{aligned} & E[H^{-1}Z_0^*Z_{P1}] \\ \sim\ & E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_xB_x'\right]^{-1}\left[\frac{1}{n}\sum I_{ux}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right] \\ & \times\left[\frac{1}{n}\sum_{|\lambda_t|\le\pi/(2M)} I_{xu}(\lambda_t)f_{uu}(\lambda_t)^{-2}\left(-\frac{1}{2m\pi}\int dB_uB_x'\left[\int B_xB_x'\right]^{-1}\int B_xdB_u\right)\right] \\ \sim\ & -E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_xB_x'\right]^{-1}\left[\frac{1}{n^2}\sum_{|\lambda_t|\le\pi/(2M)} I_{xx}(\lambda_t)f_{uu}(\lambda_t)^{-2}\right] \end{aligned}$$

$$
\begin{aligned}
&\times\left(\frac{1}{2m\pi}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right)\\
\sim\ & -E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x B_x'\right]^{-1}\left[\frac{1}{n^2}\sum_{|\lambda_t|\le\pi/(2M)} I_{xx}(\lambda_t)\left(f_{uu}(0)^{-2}+o(1)\right)\right]\\
&\times\left(\frac{1}{2m\pi}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right)\\
\sim\ & -E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x B_x'\right]^{-1}\left[\frac{m}{n^2}\left(\frac{1}{m}\sum_{|\lambda_t|\le\pi/(2M)} I_{xx}(\lambda_t)\right)f_{uu}(0)^{-2}\right]\\
&\times\left(\frac{1}{2m\pi}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right)\\
\sim\ & -E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x B_x'\right]^{-1}\left[\frac{m}{n^2}\widehat{f}_{xx}(0)f_{uu}(0)^{-2}\right]\\
&\times\left(\frac{1}{2m\pi}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right)\\
\sim\ & -E\left[\frac{1}{2\pi}f_{uu}(0)^{-1}\int B_x B_x'\right]^{-1}\left[\frac{1}{2\pi}f_{uu}(0)^{-2}\int B_x B_x'\right]\\
&\left(\frac{1}{2m\pi}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right)\\
\sim\ & -\frac{1}{m}E\left(\frac{1}{2\pi}f_{uu}(0)^{-1}\int dB_u B_x'\left[\int B_x B_x'\right]^{-1}\int B_x dB_u\right).
\end{aligned}
$$

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

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

and

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

$$
\begin{aligned}
E[Z_0^* H^{-2} D_{B1} Z_0] &\sim EH^{-2}D_{B1}\left[\frac{1}{\sqrt{n}}\sum I_{ux}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]\\
&\sim -M^{-q}k_q f_{uu}(0)^{-1}f_q(0),
\end{aligned}
$$

$$E[Z_{B1}^* H^{-1} Z_0] = EH^{-1}\left[\frac{1}{n}\sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s\right]\left[\frac{1}{\sqrt{n}}\sum I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]$$
$$\sim -M^{-q} k_q f_{uu}(0)^{-1} f_q(0),$$

$$\begin{aligned}
E[Z_{B1}^* H^{-1} Z_{B1}] &= EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2}\sum_j I_{xx}(\lambda_j) I_{uu}(\lambda_j) f_q(\lambda_j)^2 f_u(\lambda_j)^{-4} \\
&\sim EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2}\sum_j I_{xx}(\lambda_j) f_q(\lambda_j)^2 f_u(\lambda_j)^{-3} \\
&= EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2}\sum_t I_{xx}(\lambda_t) f_q(\lambda_t)^2 f_u(\lambda_t)^{-3} \\
&= EM^{-2q} k_q^2 H^{-1} \frac{1}{n^2}\sum_t \left[\frac{1}{2\pi}\sum_r C_{xx}(r) e^{-ir\lambda_t}\right] \\
&\quad \times \left[\frac{1}{(2\pi)^2}\sum_{l,j} C_{uu}(l) C_{uu}(j) \mid j \mid^q \mid l \mid^q e^{-i(l+j)\lambda_t}\right] \\
&\quad \times \left[\frac{1}{(2\pi)^3}\sum_{g,h,p} D_g D_h D_p e^{i(g+h+p)\lambda_t}\right] \\
&= EM^{-2q} k_q^2 H^{-1} \frac{1}{n^2}\frac{1}{(2\pi)^5}\sum_{g,h,p}\sum_{l,j} C_{uu}(j) C_{uu}(l) \mid j \mid^q \mid l \mid^q D_g D_h D_p \\
&\quad \times \sum_r C_{xx}(r)\left(\sum_{t=-n/2+1}^{n/2} e^{i(g+h+p-j-l-r)\lambda_t}\right) \\
&= EM^{-2q} k_q^2 H^{-1} \frac{1}{n^2}\left[\frac{1}{(2\pi)^5}\sum_{g,h,p}\sum_{l,j} C_{uu}(j) C_{uu}(l) \mid j \mid^q \mid l \mid^q D_g D_h D_p C_{xx}(\underline{r}) n\right] \\
&\sim M^{-2q} k_q^2 H^{-1} \frac{1}{2\pi} f_{uu}(0)^{-3} f_q(0)^2 \int B_x B_x' \\
&\sim M^{-2q} k_q^2 f_{uu}(0)^{-2} f_q(0)^2.
\end{aligned}$$

Similarly,

$$E[Z_0^* H^{-3} D_{B1}^2 Z_0] \sim EH^{-3} k_q^2 M^{-2q}\left[\frac{1}{n^2}\sum_j I_{xx}(\lambda_j) f_q(\lambda_j) f_u(\lambda_j)^{-2}\right]^2$$

$$
\begin{aligned}
&\times \left[\frac{1}{n^2} \sum_j I_{xx}(\lambda_j) I_{uu}(\lambda_j) f_u(\lambda_j)^{-1} \right] \\
\sim\;& EH^{-4} k_q^2 M^{-2q} \left[\frac{1}{2\pi} f_{uu}(0)^{-2} f_q(0) \int B_x B_x' \right]^2 \left[\frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x B_x' \right] \\
=\;& k_q^2 M^{-2q} f_{uu}(0)^{-2} f_q(0)^2,
\end{aligned}
$$

and

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

Similarly

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

and

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

9.8. Lemma 10

As shown in Phillips, Ouliaris and Corbae (1997), the discrete transform of the I(1) process x_t has the following asymptotic form

$$
w_x(\lambda_s) \sim \frac{1}{1 - e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} B_x(1),
$$

Thus

$$
\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} I_{xu}(\lambda_s) f_{uu}(\lambda_s)^{-1} \\
=\;& \frac{1}{\sqrt{n}} \sum_{\lambda_s \in B} \left[\frac{1}{1 - e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} B_x(1) \right] w_u(\lambda_s)^* f_{uu}(\lambda_s)^{-1} \\
\to\;& N\left(0, \int_B \left[f_{vv}(\omega) + \frac{1}{2\pi} \frac{B_x(1) B_x(1)'}{|1 - e^{i\omega}|^2} \right] f_{uu}(\omega)^{-1} d\omega \right),
\end{aligned}
$$

and

$$\begin{aligned} &\frac{1}{n}\sum_{\lambda_s\in B} I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1} \\ \to\ &\int_B \left[f_{vv}(\omega)+\frac{1}{2\pi}\frac{B_x(1)B_x(1)'}{|1-e^{i\omega}|^2}\right] f_{uu}(\omega)^{-1}d\omega. \end{aligned}$$

Thus

$$\begin{aligned} \sqrt{n}(\widehat{\beta}_B-\beta_B) &= \left[\frac{1}{n}\sum_{\lambda_s\in B} I_{xx}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right]^{-1}\left[\frac{1}{\sqrt{n}}\sum_{\lambda_s\in B} I_{xu}(\lambda_s)f_{uu}(\lambda_s)^{-1}\right] \\ &\to N\left(0,\left\{\int_B \left[f_{vv}(\omega)+\frac{1}{2\pi}\frac{B_x(1)B_x(1)'}{|1-e^{i\omega}|^2}\right] f_{uu}(\omega)^{-1}d\omega\right\}^{-1}\right). \end{aligned}$$

9.9. Theorem 16

Notice that

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

and $\widehat{H}^{-1}=H^{-1}-H^{-1}DH^{-1}+H^{-1}DH^{-1}DH^{-1}+R_1$. 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

$$\begin{aligned} W &= Z^*(H^{-1}-A_1+A_2)R'\left[J_0^{-1}+J_0^{-1}J_1J_0^{-1}-J_0^{-1}J_2J_0^{-1}+J_0^{-1}J_1J_0^{-1}J_1J_0^{-1}\right]^{-1} \\ &\quad \times R(H^{-1}-A_1+A_2)Z+\text{ higher order terms,} \\ &= Z^*H^{-1}R'J_0^{-1}RH^{-1}Z+Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RH^{-1}Z-2Z^*H^{-1}R'J_0^{-1}RA_1Z \\ &\quad -Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RH^{-1}Z+Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}J_1J_0^{-1}RH^{-1}Z \\ &\quad -2Z^*H^{-1}R'J_0^{-1}J_1J_0^{-1}RA_1Z+2Z^*H^{-1}R'J_0^{-1}RA_2Z+Z^*A_1R'J_0^{-1}RA_1Z \\ &\quad +\text{ higher order terms} \end{aligned}$$

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

$$
\begin{aligned}
E[W] &= E[Z_0^* H^{-1} R' (RH^{-1}R')^{-1} RH^{-1} Z_0] + E[Z_{V1}^* H^{-1} R' J_0^{-1} RH^{-1} Z_{V1}] \\
&\quad + E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{V2} H^{-1} R' J_0^{-1} RH^{-1} Z_0] \\
&\quad + 2E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} Z_{V2}] - 2E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{V2} H^{-1} Z_0] \\
&\quad + 2E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{B1} H^{-1} Z_0] - 2E[Z_{B1}^* H^{-1} R' J_0^{-1} RH^{-1} Z_0] \\
&\quad - E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{B1} H^{-1} R' J_0^{-1} RH^{-1} Z_0] \\
&\quad - 2Z_{P1}^* H^{-1} R' J_0^{-1} RH^{-1} Z_0 + 2E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{P1} H^{-1} Z_0] \\
&\quad - E[Z_0^* H^{-1} R' J_0^{-1} RH^{-1} D_{P1} H^{-1} R' J_0^{-1} RH^{-1} Z_0] + \text{higher order terms}
\end{aligned}
$$

9.10. Lemma 17

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

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

$$= \frac{pa}{m}$$

Similarly,

$$\begin{aligned}
E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} Z_{V2}] &\sim \frac{pa}{m}, \\
E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} R' J_0^{-1} R H^{-1} Z_0] &\sim \frac{pa}{m}, \\
E[Z_0^* H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} Z_0] &\sim \frac{pa}{m},
\end{aligned}$$

and

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

and

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

References

[1] Andrews, D. W. K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, Econometrica, 59, 817-858.

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

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

[4] Brillinger, D. R. (1980), Time Series: Data Analysis and Theory. New York: Holt, Rinehart and Winston.

[5] Byron, R. P., 1970, The restricted Aitken estimation of sets of demand equations, Econometrica 38, 816-830.

[6] Corbae, D., S. Ouliaris, and P.C.B. Phillips (1994), A Reexamination of the consumption Function Using Frequency Domain Regressions, Empirical Economics 19, 595-609.

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

[8] Engle, R. (1974), Band spectrum regression, International Economic Review 15(1), 1-11.

[9] Hannan, E. J. (1963a), Regression for Time Series, in M. Rosenblatt(ed.), Time Series Analysis. New York:Wiley.

[10] Hannan, E.J. (1963b), Regression for Time Series with Errors of Measurement, Biometrika, 50, 293 - 302.

[11] Hannan, E. J. (1970), Multiple Time Series. New York: Wiley.

[12] Hannan, E. J. (1971), Nonlinear Time Series Regression, Journal of Applied Probability, 8, 767-780.

[13] Laitinen, K., 1978, Why is demand homogeneity so often rejected?, Economic Letters, 1, 187 - 191.

[14] Linton, O. B. (1995a), Second Order Approximation in the Partially Linear Regression Model, Econometrica, Vol. 63, No. 5, 1079-1112.

[15] Linton, O. B. (1995b), Second Order Approximation in a Linear Regression with Heteroskedasticity of Unknown Form, Econometric Reviews *15, 1-32.*

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

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

[18] 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.

[19] Phillips, P.C.B. 1987, Asymptotic Expansions in Nonstationary Vector Autoregressions, Econometric Theory 3, 45 - 68.

[20] Phillips, P. C. B., S. Ouliaris and D. Corbae (1997), Band Spectral Regressions with Trending Data, Yale University, mimeographed.

[21] Priestley, M.B., Spectral Analysis and Time Series, 1981, Academic Press.

[22] Robinson, P. M(1987), Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form, Econometrica, 56, 875-891.

[23] Robinson, P. M. (1991), Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models, Econometrica *59, 755-786.*

[24] Rothenberg, T. J. (1984a), Approximating the distributions of Econometric Estimators and Test Statistics, Ch14 in Handbook of Econometrics, Vol. 2, ed. by Z. Griliches and M. Intriligator. New York: North Holland.

[25] Rothenberg, T. J. (1984b), Approximate Normality of Generalized Least Squares Estimators, Econometrica 52,811-825.

[26] Sargan, J. D. (1976), Econometric Estimators and the Edgeworth Expansion, Econometrica 44, 421-448.

[27] Sargan, J. D., and W. M. Mikhail (1971), A General Approximation to the Distribution of Instrumental Variables Estimates, Econometrica 39, 131-169.

[28] Xiao, Z. and P.C.B. Phillips (1998), Higher Order Approximations for Frequency Domain Time Series Regression, Journal of Econometrics (forthcoming).