BAND SPECTRAL REGRESSION WITH TRENDING DATA

BY

DEAN CORBAE, SAM OULIARIS
and PETER C. B. PHILLIPS

COWLES FOUNDATION PAPER NO. 1039

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

2002

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

Econometrica, Vol. 70, No. 3 (May, 2002), 1067–1109

BAND SPECTRAL REGRESSION WITH TRENDING DATA

BY DEAN CORBAE, SAM OULIARIS, AND PETER C. B. PHILLIPS[1]

Band spectral regression with both deterministic and stochastic trends is considered. It is shown that trend removal by regression in the time domain prior to band spectral regression can lead to biased and inconsistent estimates in models with frequency dependent coefficients. Both semiparametric and nonparametric regression formulations are considered, the latter including general systems of two-sided distributed lags such as those arising in lead and lag regressions. The bias problem arises through omitted variables and is avoided by careful specification of the regression equation. Trend removal in the frequency domain is shown to be a convenient option in practice. An asymptotic theory is developed and the two cases of stationary data and cointegrated nonstationary data are compared. In the latter case, a levels and differences regression formulation is shown to be useful in estimating the frequency response function at nonzero as well as zero frequencies.

KEYWORDS: Band spectral regression, deterministic and stochastic trends, discrete Fourier transform, distributed lag, integrated process, leads and lags regression, nonstationary time series, two-sided spectral BN decomposition.

1. INTRODUCTION

HANNAN'S (1963a, b) BAND-SPECTRUM REGRESSION procedure is a useful regression device that has been adopted in some applied econometric work, notably Engle (1974), where there are latent variables that may be regarded as frequency dependent (like permanent and transitory income) and where there is reason to expect that the relationship between the variables may depend on frequency. More recently, band spectral regression has been used to estimate cointegrating relations, which describe low frequency or long run relations between economic variables. In particular, Phillips (1991a) showed how frequency domain regressions that concentrate on a band around the origin are capable of producing asymptotically efficient estimates of cointegrating vectors. Interestingly, in that case the spectral methods are used with nonstationary integrated time series in precisely the same way as they were originally designed for stationary series, on which there is now a large literature (see, inter alia, Hannan (1970 Ch. VII), Hannan and Thomson (1971), Hannan and Robinson (1973), and Robinson (1991)). Related methods were used in Choi and Phillips (1993) to construct frequency domain tests for unit roots. In all of this work, the capacity of frequency domain techniques to deal nonparametrically with short memory components in

[1] Our thanks go to a co-editor and three referees for helpful comments. An early draft was written in the summer of 1994. The first complete version of the paper was written over the period 1 January–1 July 1997 while Phillips was visiting the University of Auckland. Lemmas A–D are due to Phillips and are taken from other work to make this paper self contained. Phillips thanks the NSF for research support under Grant Nos. SBR 94-22922, SBR 97-30295, and SES 0092509.

the data is exploited. Robinson (1991), who allowed for band dependent model formulations, showed how, in such semiparametric models, data based methods can be used to determine the smoothing parameter that plays a central role in the treatment of the nonparametric component. Other data based methods have been used directly in the study of cointegrating relations and testing problems by Xiao and Phillips (1998, 1999).

To the extent that relations between variables may be band dependent, it is also reasonable to expect that causal relations between variables may vary according to frequency. Extending work by Geweke (1982), both Hosoya (1991) and Granger and Lin (1995) have studied causal links from this perspective and constructed measures of causality that are frequency dependent. Most recently, frequency band methods have been suggested for the estimation of models with long memory components (e.g., Marinucci (1998), Robinson and Marinucci (1998), Kim and Phillips (1999)), where salient features of the variables are naturally evident in frequency domain formulations.

This paper studies some properties of frequency band regression in the presence of both deterministic and stochastic trends, a common feature in economic time series applications. In such cases, a seemingly minor issue relates to the manner of deterministic trend removal. In particular, should the deterministic trends be eliminated by regression in the time domain prior to the use of the frequency domain regression or not? Since spectral regression procedures were originally developed for stationary time series and since the removal of deterministic trends by least squares regression is well known to be asymptotically efficient under weak conditions (Grenander and Rosenblatt (1957, p. 244)), it may seem natural to perform the trend removal in the time domain prior to the use of spectral methods. Indeed, this is recommended in Hannan (1963a, p. 30; 1970, pp. 443–444), even though the development of spectral regression there and in Hannan (1970, Ch. VII.4) allows for regressors like deterministic trends that are susceptible to a generalized harmonic analysis, satisfying the so-called Grenander conditions (Grenander and Rosenblatt (1957, p. 233), Hannan (1970, p. 77)). Theorem 11 of Hannan (1970, p. 444) confirms that such time domain detrending followed by spectral regression is asymptotically efficient in models where all the variables are trend stationary and where the model coefficients do not vary with frequency.

In time domain regression, the Frisch–Waugh (1933) theorem assures invariance of the regression coefficients to prior trend removal or to the inclusion of trends in the regression itself. Such invariance is often taken for granted in empirical work. However, as will be shown here, invariance does not always apply for band-spectrum regression when one switches from time domain to frequency domain regressions. In particular, detrending by removing deterministic components in the time domain and then applying band-spectrum regression is not necessarily equivalent to detrending in the frequency domain and applying band-spectrum regression. The reason is that a band dependent spectral model is similar in form to a linear regression with a structural change in the coefficient vector. So, when there are deterministic trends in the model, the original deterministic regressors also need to be augmented by regressors that are relevant to the change period (here, the frequency band where the change occurs). Thus, the seemingly innocuous matter of detrending can have some nontrivial consequences in practice. In particular, detrending in the time domain yields estimates that can be biased in finite samples, and, in the case of nonstationary data, inconsistent. The appropriate procedure is to take account of the augmented regression equation prior to detrending and this can be readily accomplished by including the deterministic variables explicitly in the frequency domain regression. These issues are relevant whenever band spectral methods are applied, including those cases that involve cointegrating regressions. In the latter case, a levels and differences regression formulation is shown to be appropriate and useful in estimating the frequency response function at nonzero as well as zero frequencies.

The present paper provides an asymptotic analysis of frequency domain regression with trending data for the two cases of stationary and cointegrated nonstationary data. We consider both semiparametric and fully nonparametric formulations of the regression model, in both cases allowing the model coefficients to vary with frequency. In dealing with the nonstationary case, the paper introduces some new methods for obtaining a limit theory for discrete Fourier transforms (dft's) of integrated time series and makes extensive use of two sided BN decompositions, which lead to levels and differences model formulations. This asymptotic theory simplifies some earlier results given in Phillips (1991a) and shows that the dft's of an $I(1)$ process are spatially (i.e., frequencywise) dependent across all the fundamental frequencies, even in the limit as the sample size $n \to \infty$, due to leakage from the zero frequency. This leakage is strong enough to ensure that smoothed periodogram estimates of the spectrum away from the origin are inconsistent at all frequencies $\omega \neq 0$. The techniques and results given here should be useful in other regression contexts where these quantities arise. Some extensions of the methods to cases of data with long range dependence are provided in Phillips (1999).

Most of our notation is standard: $[a]$ signifies the largest integer not exceeding a, $>$ signifies positive definiteness when applied to matrices, a^* is the complex conjugate transpose of the matrix a, $\|a\|$ is a matrix norm of a, a^- is the Moore Penrose inverse of a, $P_a = a(a^*a)^- a^*$ is the orthogonal projector onto the range of a, $L_2[0, 1]$ is the space of square integrable functions on $[0, 1]$, $1[A]$ is the indicator of A, $\overset{d}{\sim}$ signifies 'asymptotically distributed as' and $\overset{d}{\to}$ and $\overset{p}{\to}$ are used to denote weak convergence of the associated probability measures and convergence in probability, respectively, as the sample size, $n \to \infty$; $I(1)$ signifies an integrated process of order one, $BM(\Omega)$ denotes a vector Brownian motion with covariance matrix Ω, and we write integrals like $\int_0^1 B(r)\,dr$ as $\int_0^1 B$, or simply $\int B$ if there is no ambiguity over limits; $MN(0, G)$ signifies a mixed normal distribution with matrix mixing variate G, $N_c(0, G)$ signifies a complex normal distribution with covariance matrix G, the discrete Fourier transform (dft) of $\{a_t; t = 1, \ldots, n\}$ is written $w_a(\lambda) = (1/\sqrt{n}) \sum_{t=1}^{n} a_t e^{i\lambda t}$, and $\{\lambda_s = (2\pi s/n) : s = 0, 1, \ldots, n-1\}$ are the fundamental frequencies.

2. MODELS AND ESTIMATION PRELIMINARIES

Let x_t $(t = 1, \ldots, n)$ be an observed time series generated by

$$x_t = \Pi_2' z_t + \tilde{x}_t, \tag{1}$$

where z_t is a $p+1$-dimensional deterministic sequence and $\tilde{x}_t$ is a zero mean k-dimensional time series. The series x_t therefore has both a deterministic component involving the sequence z_t and a stochastic (latent) component $\tilde{x}_t$. In developing our theory it will be convenient to allow for both stationary and nonstationary $\tilde{x}_t$. Accordingly, we introduce the following two alternative assumptions. The mechanism linking the observed variables x_t, unobserved disturbances ε_t, and a dependent variable y_t is made explicit later.

ASSUMPTION 1: *$\varsigma_t = (\varepsilon_t, \tilde{x}_t')'$ is a jointly stationary time series with Wold representation $\varsigma_t = \sum_{j=0}^{\infty} C_j \xi_{t-j}$, where $\xi_t = iid\ (0, \Sigma)$ with finite fourth moments and the coefficients C_j satisfy $\sum_{j=0}^{\infty} j^{\frac{1}{2}} \|C_j\| < \infty$. Partitioned conformably with ς_t, the spectral density matrix $f_{\varsigma\varsigma}(\lambda)$ of ς_t is*

$$f_{\varsigma\varsigma}(\lambda) = \begin{bmatrix} f_{\varepsilon\varepsilon}(\lambda) & 0 \\ 0 & f_{xx}(\lambda) \end{bmatrix},$$

with $f_{\varepsilon\varepsilon}(\lambda), f_{xx}(\lambda) > 0\ \forall\, \lambda$.

ASSUMPTION 2: *$\tilde{x}_t$ is an $I(1)$ process satisfying $\Delta\tilde{x}_t = v_t$, initialized at $t = 0$ by any $O_p(1)$ random variable. The shocks $\varsigma_t = (\varepsilon_t, v_t')'$ satisfy Assumption 1 with spectral density $f_{\varsigma\varsigma}(\lambda)$ partitioned as*

$$f_{\varsigma\varsigma}(\lambda) = \begin{bmatrix} f_{\varepsilon\varepsilon}(\lambda) & 0 \\ 0 & f_{vv}(\lambda) \end{bmatrix},$$

with $f_{\varepsilon\varepsilon}(\lambda), f_{vv}(\lambda) > 0\ \forall \lambda$.

Assumptions 1 and 2 suffice for partial sums of ς_t to satisfy the functional law $n^{-1/2} \sum_{t=1}^{[nr]} \varsigma_t \xrightarrow{d} B(r) = BM(\Omega)$, a vector Brownian motion of dimension $(k+1)$ with covariance matrix $\Omega = 2\pi f_{\varsigma\varsigma}(0)$ (e.g., Phillips and Solo (1992, Theorem 3.4)). The vector process B and matrix Ω can be partitioned conformably with ς as $B = (B_\varepsilon, B_x')'$ and

$$\Omega = \begin{bmatrix} \Omega_{\varepsilon\varepsilon} & 0 \\ 0 & \Omega_{xx} \end{bmatrix},$$

where $\Omega_{xx} > 0$, so that $\tilde{x}_t$ is a full rank vector $I(1)$ process. In all cases, $\tilde{x}_t$ is taken to be independent of ε_t. This amounts to a strict exogeneity assumption when we introduce a generating mechanism for a dependent variable y_t. The assumption is standard for consistent estimation in the stationary case (as in Hannan (1963a, b)). In the nonstationary case, it is not required for the consistent estimation of (low frequency) cointegrating regression components. However, at bands away from the origin, some version of incoherency between the regressors and errors is required for consistency and asserting incoherency enables us to examine the bias and inconsistency effects of detrending in such regressions, just as in the case of stationary $\tilde{x}_t$.

We make the following assumptions concerning the deterministic sequence z_t. We confine our treatment to time polynomials and set $z_t = (1, t, \ldots, t^p)'$, the most common case in practice, although we anticipate that our conclusions about bias and inconsistency will apply more generally, for instance to trigonometric polynomials, after some appropriate extensions of our dft formulae in Lemma A and the limit results in Lemma C below. Let $\delta_n = \text{diag}(1, n, \ldots, n^p)$, and define $d_t = \delta_n^{-1} z_t$. Then

$$d_{[nr]} = \delta_n^{-1} z_{[nr]} \to u(r) = (1, r, \ldots, r^p)', \tag{2}$$

uniformly in $r \in [0, 1]$. The limit functions $u(r)$ are linearly independent in $L_2[0, 1]$ and $n^{-1}(\sum_{t=1}^n d_t d_t') \to \int_0^1 uu' > 0$. Now let Z (respectively, D) be the $n \times p$ observation matrix of the nonstochastic regressors z_t (respectively, standardized regressors d_t), P_Z be the projector onto the range of Z, and $Q_Z = I_n - P_Z$ be the residual projection matrix (respectively, P_D and $I - P_D$). Clearly, $P_Z = P_D$ and $Q_Z = Q_D$.

The Nonparametric Model

The type of model we have in mind for the stochastic component of the data allows for the regression coefficient to vary across frequency bands. In the time domain, suppose a (latent) dependent variable $\tilde{y}_t$ is related to $\tilde{x}_t$ and ε_t according to a (possibly two-sided) distributed lag

$$\tilde{y}_t = \sum_{j=-\infty}^{\infty} \beta_j' \tilde{x}_{t-j} + \varepsilon_t = \beta(L)' \tilde{x}_t + \varepsilon_t, \tag{3}$$

where the transfer function of the filter, $b(\omega) = \beta(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \beta_j e^{ij\omega}$, is assumed to converge for all $\omega \in [-\pi, \pi]$. Under stationarity (Assumption 1), the cross spectrum $f_{yx}(\omega)$ between $\tilde{y}_t$ and $\tilde{x}_t$ satisfies

$$f_{yx}(\omega) = b(\omega)' f_{xx}(\omega), \tag{4}$$

whose complex conjugate transpose gives the complex regression coefficient

$$\beta_\omega = b(-\omega) = f_{xx}(\omega)^{-1} f_{xy}(\omega). \tag{5}$$

As argued in Hannan (1963a, b), spectral methods provide a natural mechanism for focusing attention on a particular frequency, or band of frequencies, and for performing estimation of β_ω.

Next, suppose that the coefficients β_j in (3) are such that $\beta(L)$ has a valid two-sided spectral BN decomposition (see part (b) of Lemma D in the Appendix), viz.

$$\beta(L) = \beta(e^{i\omega}) + \tilde{\beta}_\omega(e^{-i\omega}L)(e^{-i\omega}L - 1), \tag{6}$$

where

$$\tilde{\beta}_\omega(L) = \sum_{j=-\infty}^{\infty} \tilde{\beta}_{\omega j} L^j, \tilde{\beta}_{\omega j} = \begin{cases} \sum_{k=j+1}^{\infty} \beta_k e^{ik\omega}, & j \geq 0, \\ -\sum_{k=-\infty}^{j} \beta_k e^{ik\omega}, & j < 0, \end{cases}$$

with $\sum_{j=-\infty}^{\infty} \|\tilde{\beta}_j\|^2 < \infty$. Using (6) in (3) we have

$$\begin{aligned} \tilde{y}_t &= \beta(L)'\tilde{x}_t + \varepsilon_t \\ &= \left[b(\omega) + \tilde{\beta}_\omega(e^{-i\omega}L)(e^{-i\omega}L - 1)\right]'\tilde{x}_t + \varepsilon_t, \\ &= b(\omega)'\tilde{x}_t + (e^{-i\omega}L - 1)\tilde{a}_{\omega t} + \varepsilon_t, \end{aligned} \tag{7}$$

in which $\tilde{a}_{\omega t} = \tilde{\beta}_\omega(e^{-i\omega}L)\tilde{x}_t$. Then, setting $\omega = \lambda_s$ in (7) and taking dft's at frequency λ_s we obtain the following frequency domain form of (3):

$$w_{\tilde{y}}(\lambda_s) = b(\lambda_s)'w_{\tilde{x}}(\lambda_s) + w_\varepsilon(\lambda_s) + \frac{1}{\sqrt{n}}\left[e^{-i\lambda_s}\tilde{a}_{\lambda_s 0} - \tilde{a}_{\lambda_s n}\right]. \tag{8}$$

For stationary $\tilde{x}_t$, $\tilde{a}_{\omega t}$ is also stationary and (8) can be written as

$$w_{\tilde{y}}(\lambda_s) = b(\lambda_s)'w_{\tilde{x}}(\lambda_s) + w_\varepsilon(\lambda_s) + O_p\left(\frac{1}{\sqrt{n}}\right), \tag{9}$$

giving an approximate frequency domain version of (3).

The case of integrated $\tilde{x}_t$ (Assumption 2) can be handled in a similar way. Here it is useful to apply to (3) the two-sided BN decomposition at unity (part(a) of Lemma D), viz.

$$\beta(L) = \beta(1) + \tilde{\beta}(L)(L - 1), \quad \text{where} \tag{10}$$

$$\tilde{\beta}(L) = \sum_{j=-\infty}^{\infty} \tilde{\beta}_j L^j, \quad \tilde{\beta}_j = \begin{cases} \sum_{k=j+1}^{\infty} \beta_k, & j \geq 0, \\ -\sum_{k=-\infty}^{j} \beta_k, & j < 0, \end{cases}$$

so that (3) can be interpreted as

$$\tilde{y}_t = \beta(1)'\tilde{x}_t + \varepsilon_t - \tilde{\beta}(L)'v_t = \beta(1)'\tilde{x}_t + \tilde{\varepsilon}_t, \quad \text{say}, \tag{11}$$

which is a cointegrating equation between $\tilde{y}_t$ and $\tilde{x}_t$ with cointegrating vector $(1, -\beta(1)')$ and composite error $\tilde{\varepsilon}_t = \varepsilon_t - \tilde{\beta}(L)'v_t$. Setting $\eta_t = \tilde{\beta}(L)'v_t$ and taking the dft of (11), we obtain

$$w_{\tilde{y}}(\lambda_s) = \beta(1)' w_{\tilde{x}}(\lambda_s) + w_{\varepsilon}(\lambda_s) - w_{\eta}(\lambda_s) \tag{12}$$

$$= b(\lambda_s)' w_{\tilde{x}}(\lambda_s) + w_{\varepsilon}(\lambda_s) - w_{\eta}(\lambda_s) + [\beta(1) - b(\lambda_s)]' w_{\tilde{x}}(\lambda_s). \tag{13}$$

The final two terms of (13) are important, revealing that the approximation (9) fails in the nonstationary case and has an error of specification that arises from omitted variables.

Assuming it is valid, the spectral BN decomposition of $\tilde{\beta}(L)$ at frequency λ_s is

$$\tilde{\beta}(L) = \tilde{\beta}(e^{i\lambda_s}) + \overset{\approx}{\beta}_{\lambda_s}(e^{-i\lambda_s}L)(e^{-i\lambda_s}L - 1)$$

where $\overset{\approx}{\beta}_{\lambda_s}$ is constructed as in Lemma D(b). Then

$$\begin{aligned} w_{\eta}(\lambda_s) &= \tilde{\beta}(e^{i\lambda_s})' w_v(\lambda_s) + \frac{1}{\sqrt{n}}\left[e^{-i\lambda_s}\tilde{v}_{\lambda_s 0} - \tilde{v}_{\lambda_s n}\right] \\ &= \tilde{\beta}(e^{i\lambda_s})' w_v(\lambda_s) + O_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned} \tag{14}$$

where $\tilde{v}_{\lambda_s t} = \overset{\approx}{\beta}_{\lambda_s}(e^{-i\lambda_s}L)v_t$ is stationary. It follows that (12) has the form

$$w_{\tilde{y}}(\lambda_s) = \beta(1)' w_{\tilde{x}}(\lambda_s) - \tilde{\beta}(e^{i\lambda_s})' w_v(\lambda_s) + w_{\varepsilon}(\lambda_s) + O_p\left(\frac{1}{\sqrt{n}}\right), \tag{15}$$

giving an approximate frequency domain version of the cointegrated model that applies for all frequencies. Since $v_t = \Delta\tilde{x}_t$, it is apparent from (15) that if data on $\tilde{y}_t$ and $\tilde{x}_t$ were available, the equation could be estimated in the frequency domain by band spectral regression of $w_{\tilde{y}}(\lambda_s)$ on $w_{\tilde{x}}(\lambda_s)$ and $w_{\Delta\tilde{x}}(\lambda_s)$ for frequencies λ_s centered around some frequency ω. For λ_s centered around $\omega = 0$, this procedure was suggested originally in Phillips (1991b) and Phillips and Loretan (1991), where it is called augmented frequency domain regression, although in that work deterministic trends were excluded. In the general case where λ_s is centered on $\omega \neq 0$, we can recover the spectral coefficient $\beta_{\omega} = b(-\omega)$ from the coefficients $\beta(1)$ and $\tilde{\beta}(e^{i\lambda_s})' = \tilde{\beta}(e^{-i\lambda_s})^*$ in (15) using the fact that

$$b(-\omega) = \beta(e^{-i\omega}) = \beta(1) + \tilde{\beta}(e^{-i\omega})(e^{-i\omega} - 1). \tag{16}$$

Causal economic relationships are often formulated as in (3) with one-sided relations where $\beta_j = 0$ for $j < 0$, but two-sided relations are not excluded. They are well known to occur in situations where there is feedback in both directions between the variables $\tilde{y}$ and $\tilde{x}$. A notable recent example arises in the simplification of cointegrating systems to single equation formulations where two-sided dynamic regressions like (11) arise and where they have been used in regression analysis to obtain efficient estimates of the cointegrating vector (Phillips and Loretan (1991), Saikonnen (1991), Stock and Watson (1993)).

The mechanism linking the observed dependent variable y_t to z_t and $\tilde{y}_t$ (and, hence, $\tilde{x}_t$, and ε_t) is now given by

$$y_t = z_t'\pi_1 + \tilde{y}_t, \tag{17}$$

where $\tilde{y}_t$ follows (3) in the stationary case with $(\tilde{x}_t, \varepsilon_t)$ satisfying Assumption 1, and (11) in the cointegrated case with $(\tilde{x}_t, \varepsilon_t)$ satisfying Assumption 2. The observed series y_t and x_t therefore have deterministic trends and stochastic components that are linked by a system of distributed lags with stationary errors. The natural statistical approach would appear to be deterministic trend removal by regression of (y_t, x_t) on z_t, followed by an analysis of the system of lagged dependencies between the residuals from these regressions. The latter can be conducted in the frequency domain where the coefficient of interest is the transfer function of the filter $b(\omega)'$. This coefficient will vary according to frequency (unless $\beta_j = 0$ for $j \neq 0$) and can be estimated by nonparametric regression techniques. One of the aims of this paper is to examine the finite sample and asymptotic performance of this natural procedure against that of a procedure that seeks to perform the regression analysis in the frequency domain directly on the observed quantities (y_t, x_t, z_t) rather than on detrended data.

A Fixed Band Regression Model

A simple situation of interest occurs when the response function $b(\omega)$ is constant, let us say on two fixed frequency bands. This case captures the essential features of the problem that we intend to discuss, is easy to analyze, and will be considered at some length here and in the asymptotic theory. It has the advantage that $b(\omega)$ is real and parametric and can be estimated at parametric rates when the bands are of fixed positive length. This case is analogous to a regression model with a structural break and that analogy helps to explain why time and frequency domain detrending have different effects. The nonparametric model can be interpreted as a limiting case in which one of the bands shrinks to a particular frequency as the sample size increases. To promote the analogy with a structural break model we will develop a conventional regression model formulation for the data.

To this end, we postulate the dual bands $\mathscr{B}_A = [-\omega_0, \omega_0]$ and $\mathscr{B}_A^c = [-\pi, -\omega_0) \cup (\omega_0, \pi]$ for some given frequency $\omega_0 > 0$ and define the frequency dependent coefficient vector

$$b(\omega) = \beta_A 1[\omega \in \mathscr{B}_A] + \beta_{A^c} 1[\omega \in \mathscr{B}_A^c], \tag{18}$$

where β_A is a k-vector of parameters pertinent to the band $\mathscr{B}_A$, and β_{A^c} is the corresponding k-vector of parameters pertinent to the band $\mathscr{B}_A^c$. This formulation of $b(\omega)$ enables us to separate low frequency ($|\omega| \leq \omega_0$) from high frequency

($|\omega| > \omega_0$) responses in a regression context. The coefficients β_j in the Fourier representation, $b(\omega) = \sum_{j=-\infty}^{\infty} \beta_j e^{ij\omega}$, of (18) are

$$\beta_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} b(\omega)e^{-ij\omega}\,dw = \begin{cases} \dfrac{\beta_A\omega_0}{\pi} + \dfrac{\beta_{A^c}}{\pi}(\pi-\omega_0), & j=0, \\ \dfrac{\sin j\omega_0}{\pi j}(\beta_A - \beta_{A^c}), & j\neq 0, \end{cases} \tag{19}$$

so that the filter in (3) is symmetric and two-sided. This is an interesting case where the coefficients β_j decay slowly and do not satisfy the sufficient condition $\sum_{-\infty}^{\infty} |j|^{\frac{1}{2}}|\beta_j| < \infty$ given in Lemma D for the validity of the BN decomposition of $\beta(L) = \sum_{j=-\infty}^{\infty} \beta L^j$. Nevertheless, the BN decomposition of $\beta(L)$ is still valid in this case, as shown in Remark (i) following Lemma D in the Appendix.

While data for $\tilde{y}_t$ can be generated by (3) and (19) by truncating the filter, an alternate frequency domain approach that uses (18) is possible. This mechanism has the advantage of revealing the impact of the shift in (18) in terms of a standard regression model with a structural change. Like x_t, the dependent variable y_t is assumed to have both deterministic and stochastic components. Its stochastic component $\tilde{y}_t$ is generated from $\tilde{x}_t$ and ε_t by means of a triangular array formulation which we explain as follows.

Let $\widetilde{X} = [\tilde{x}_1, \dots, \tilde{x}_n]'$ be the $n \times k$ matrix of observations of the exogenous regressors $\tilde{x}_t$, let U be the $n \times n$ unitary matrix with (j,k) element $e^{2\pi ijk/n}/\sqrt{n}$, and let $W = E_{1n}U$, where

$$E_{1n} = \begin{bmatrix} 0 & 1 \\ I_{n-1} & 0 \end{bmatrix}$$

is an $n \times n$ permutation matrix that reorders the rows of U so that the last row becomes the first and succeeding rows are simply advanced by one position. Then, $W^*W = I_n$, and the matrix $W\widetilde{X}$ has kth row given by the dft $w_{\tilde{x}}(\lambda_s)'$ with $s = k-1$.

We introduce the following matrices that are used to provide a frequency dependent structure to the model for y_t. Define:

- $A = n \times n$ selector matrix that zeroes out frequencies in $W\widetilde{X}$ that are not relevant to the primary band of interest, say $\mathcal{B}_A$. Then, $A^cW = [I-A]W$ extracts the residual frequencies over $\mathcal{B}_A^c$. Note that $A^cA = AA^c = 0$.
- $\Psi = W^*AW$ and $\Psi^c = W^*A^cW = I - \Psi$.

Then, for each n, $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_n)$ is generated in the frequency domain in dft form by the system

$$AW\tilde{y} = AW\widetilde{X}\beta_A + AW\varepsilon, \tag{20}$$

$$A^cW\tilde{y} = A^cW\widetilde{X}\beta_{A^c} + A^cW\varepsilon, \tag{21}$$

where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$. Equations (20) and (21) generate observations in the frequency domain that correspond to the model (8) in the nonparametric case with the final term of (8) omitted. In the stationary case, this omitted term is $O_p(n^{-\frac{1}{2}})$ as seen in (9), so the difference between the generating mechanisms is negligible asymptotically. In the nonstationary case, however, the omitted term is nonnegligible, as is apparent from (13) and (15). In consequence, the two generating mechanisms differ in the nonstationary case and this leads to a difference in the asymptotic theory between the two models that will figure in our discussion in Section 5.

Adding (20) and (21) and multiplying by W^* gives

$$\tilde{y} = \Psi \widetilde{X} \beta_A + \Psi^c \widetilde{X} \beta_{A^c} + \varepsilon, \tag{22}$$

which is the time domain generating mechanism for $\tilde{y}$. By construction, the matrices Ψ and Ψ^c have elements that depend on n, and it follows that $\{\tilde{y}_t : t = 1, \ldots, n\}$ has a triangular array structure, although we do not emphasize this by using an additional subscript and we will have no need of triangular array limit theory in our asymptotic development.

In (22), $(\varepsilon_t, \tilde{x}_t')$ satisfies Assumption 1 in the stationary case and $(\varepsilon_t, \Delta\tilde{x}_t')$ satisfies Assumption 2 in the nonstationary case. Thus, (22) is a variable (band dependent) coefficient time series regression with exogenous regressors and stationary errors under Assumption 1, and a cointegrating regression with exogenous regressors and band dependent coefficients under Assumption 2. Note that model (22) is semiparametric: parametric in the regression coefficients β_A and β_{A^c} and nonparametric in the regression errors ε.

We now suppose that the observed series y_t has deterministic components like x_t in (1) and is related to the unobserved component $\tilde{y}_t$ by means of

$$y = Z\pi_1 + \tilde{y}. \tag{23}$$

Using (22) and (23), it follows that the observed data satisfy the model

$$y = Z(\pi_1 - \Pi_2\beta_A) + X\beta_A + \Psi^c Z\Pi_2(\beta_A - \beta_{A^c}) - \Psi^c X(\beta_A - \beta_{A^c}) + \varepsilon, \tag{24}$$

or, equivalently,

$$y = Z(\pi_1 - \Pi_2\beta_{A^c}) + \Psi Z\Pi_2(\beta_{A^c} - \beta_A) + X\beta_{A^c} - \Psi X(\beta_{A^c} - \beta_A) + \varepsilon, \tag{25}$$

or

$$y = \Psi Z(\pi_1 - \Pi_2\beta_A) + \Psi^c Z(\pi_1 - \Pi_2\beta_{A^c}) + \Psi X\beta_A + \Psi^c X\beta_{A^c} + \varepsilon. \tag{26}$$

These models extend (22) to cases where the observed data contain deterministic trends. Observe that the trend regressors z_t now appear in the model with band dependent coefficients, just like the exogenous regressors x_t.

The formulations (24)–(26) make it clear that detrending the data in the time domain is not a simple matter of applying the projection matrix Q_Z, as might be expected immediately from (1) and (23). In fact, correct trend removal is accomplished by the use of the operator $Q_V = I - P_V$, where $V = [Z, \Psi^c Z]$ or, equivalently, in view of (26) $V = [\Psi Z, \Psi^c Z]$. Methods that rely on prefiltering by means of Q_Z do not fully remove the trends and this can have important consequences, like biased and inconsistent estimates of the regression coefficients β_A and β_{A^c}. Of course, when $A = I_n$ the coefficient is invariant across bands ($\beta_A = \beta$, say), we have $\Psi = I_n$ and Ψ^c is null, so that $V = Z$ and usual detrending by Q_Z is appropriate. In that case (26) reduces to

$$y = Z(\pi_1 - \Pi_2\beta) + X\beta + \varepsilon.$$

Put another way, conventional detrending by Q_Z implicitly assumes that there is no variation in the coefficient across frequency bands. If there are such variations, then correct detrending needs to take into account the effects of such variations on the trends, as in (26). Otherwise, there will be misspecification bias from omitted variables in the deterministic detrending.

Fixed Band Regressions and Misspecification Bias

A simple illustration with the Hannan (1963a) inefficient band-spectrum regression estimator shows the effects of such misspecification. This estimator can be constructed for the band $\mathscr{B}_A$ and then has the form

$$\hat{\beta}_A = (X'Q_Z\Psi Q_Z X)^{-1}(X'Q_Z\Psi Q_Z y), \tag{27}$$

with a corresponding formula for $\hat{\beta}_{A^c}$, the estimator over the band $\mathscr{B}_A^c$. In forming $\hat{\beta}_A$ and $\hat{\beta}_{A^c}$, the data are filtered by a trend removal regression via the projection Q_Z before performing the band-spectrum regression. This procedure follows Hannan's (1963a, b) recommendation for dealing with deterministic trends and, as we have discussed in the introduction, is the conventional approach in this context. Using (24) and (27), we find

$$\begin{aligned} \hat{\beta}_A &= \beta_A + \{X'Q_Z\Psi Q_Z X\}^{-1}\{X'Q_Z\Psi Q_Z[\Psi^c Z\Pi_2(\beta_A - \beta_{A^c}) \\ &\qquad - \Psi^c X(\beta_A - \beta_{A^c}) + \varepsilon]\} \\ &= \beta_A - \{X'Q_Z\Psi Q_Z X\}^{-1}\{X'Q_Z\Psi Q_Z[\Psi^c \widetilde{X}(\beta_A - \beta_{A^c}) - \varepsilon]\} \\ &= \beta_A - \{\widetilde{X}'Q_Z\Psi Q_Z \widetilde{X}\}^{-1}\{\widetilde{X}'Q_Z\Psi Q_Z[\Psi^c \widetilde{X}(\beta_A - \beta_{A^c}) - \varepsilon]\}, \end{aligned} \tag{28}$$

and the corresponding formula for $\hat{\beta}_{A^c}$ is

$$\hat{\beta}_{A^c} = \beta_{A^c} - \{\widetilde{X}'Q_Z\Psi^c Q_Z \widetilde{X}\}^{-1}\{\widetilde{X}'Q_Z\Psi^c Q_Z[\Psi \widetilde{X}(\beta_{A^c} - \beta_A) - \varepsilon]\}. \tag{29}$$

It follows that

$$E(\hat{\beta}_A|X) = \beta_A - \{X'Q_Z\Psi Q_Z X\}^{-1}\{X'Q_Z\Psi Q_Z\Psi^c \widetilde{X}(\beta_A - \beta_{A^c})\}, \tag{30}$$

and

$$E(\hat{\beta}_{A^c}|X) = \beta_{A^c} - \{X'Q_Z\Psi^c Q_Z X\}^{-1}\{X'Q_Z\Psi^c Q_Z\Psi \widetilde{X}(\beta_{A^c} - \beta_A)\}.$$

Hence, band-spectrum regression yields biased estimates of the coefficients when $\beta_A \neq \beta_{A^c}$ and trend removal regression is conducted via the projection Q_Z.

To some extent, the problem with $\hat{\beta}_A$ is a finite sample one. In fact, for stationary $\tilde{x}_t$, the bias disappears as $n \to \infty$. But, as we shall see in Section 5, when $\tilde{x}_t$ is integrated the bias in $\hat{\beta}_{A^c}$ does not always disappear in the limit.

Narrow Band Regressions

In the nonparametric case we estimate β_ω in (5) at a particular frequency ω, rather than over a fixed discrete band. In that case the band spectral estimator has the same form as (27) but the selector matrix $A = A_\omega$ (respectively, $\Psi = \Psi_\omega$) chooses frequencies in a shrinking band about ω. We write the estimator as

$$\hat{\beta}_\omega = (X'Q_Z\Psi_\omega Q_Z X)^{-1}(X'Q_Z\Psi_\omega Q_Z y). \tag{31}$$

From the form of (15) it is apparent that the narrow band regression leading to (31) omits the term involving the dft, $w_v(\lambda_s)$, of the differences $v_t = \Delta\tilde{x}_t$. So this narrow band regression also involves omitted variable misspecification.

An alternate nonparametric procedure is to use an augmented narrow band regression in which the differences are included. To fix ideas, we detrend the data using Q_Z and then perform a band spectral regression of the form

$$w_{y.z}(\lambda_s) = \tilde{a}'_{1\omega} w_{x.z}(\lambda_s) + \tilde{a}'_{2\omega} w_{\Delta x.z}(\lambda_s) + \text{residual} \tag{32}$$

for frequencies λ_s in a shrinking band around ω. In equation (32) the affix '$.z$' on the subscript (e.g., in $w_{x.z}(\lambda_s)$) signifies that the data have been detrended before taking dft's. This approach is analogous to the augmented spectral regression approach of Phillips and Loretan (1991) for estimating a cointegrating vector. However, the narrow band regression (32) now applies for frequencies ω that are nonzero as well as zero and prior detrending of the data has occurred in the time domain. Define $\underline{X} = [X, \Delta X]$ and $\tilde{a}'_\omega = (\tilde{a}'_{1\omega}, \tilde{a}'_{2\omega})$. Then

$$\tilde{a}_\omega = (\underline{X}'Q_Z\Psi_\omega Q_Z\underline{X})^{-1}(\underline{X}'Q_Z\Psi_\omega Q_Z y). \tag{33}$$

In view of the relation (16), an estimate of $\beta_\omega = \beta(e^{-i\omega})$ may be recovered from $\tilde{a}_\omega$ using the linear combination

$$\tilde{\beta}_\omega = \begin{cases} \tilde{a}_{1,-\omega} - (e^{-i\omega} - 1)\tilde{a}_{2,-\omega}, & \omega \neq 0, \\ \tilde{a}_{10}, & \omega = 0. \end{cases} \tag{34}$$

The asymptotic theory for the narrow band estimates $\hat{\beta}_\omega$ and $\tilde{\beta}_\omega$ is developed in Section 5.

3. DFT RECURSION FORMULAE

Here we develop some useful formulae for dft's of the deterministic sequence z_t. The following lemma gives general recursion formulae for the quantity $W_k(\lambda_s) = \sum_{t=1}^{n} t^k e^{i\lambda_s t}$. The case for $\lambda_s = 0$ is well known, of course, but the result for $\lambda_s \neq 0$ appears to be new.

LEMMA A: (a) *For* $s=0$,

$$W_k(\lambda_s)=\sum_{t=1}^{n}t^k=\frac{1}{k+1}\left\{n^{k+1}-\binom{k+1}{1}B_1n^k+\binom{k+1}{2}B_2n^{k-1}\right.$$
$$\left.+\binom{k+1}{3}B_3n^{k-2}+\cdots+\binom{k+1}{k}B_{k-1}n\right\},$$

where B_j *are the Bernoulli numbers.*

(b) *For* $s\neq 0$, $W_k(\lambda_s)$ *satisfies the recursion*

$$W_k(\lambda_s)=n^k\frac{e^{i\lambda_s}}{e^{i\lambda_s}-1}+\frac{1}{e^{i\lambda_s}-1}\sum_{j=1}^{k}\binom{k}{j}(-1)^jW_{k-j}(\lambda_s) \tag{35}$$

with initialization

$$W_0(\lambda_s)=\sum_{t=1}^{n}e^{i\lambda_s t}=\begin{cases}n, & s=0,\\ 0, & s\neq 0.\end{cases} \tag{36}$$

Using (35), the standardized quantities $\underline{W}_k(\lambda_s)=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\frac{t}{n}\right)^k e^{i\lambda_s t}$ satisfy the recursion

$$\underline{W}_k(\lambda_s)=\frac{1}{\sqrt{n}}\frac{e^{i\lambda_s}}{e^{i\lambda_s}-1}+\frac{1}{e^{i\lambda_s}-1}\sum_{j=1}^{k}\frac{1}{n^j}\binom{k}{j}(-1)^j\underline{W}_{k-j}(\lambda_s), \tag{37}$$

and then the dft of d_t is simply

$$w_d(\lambda_s)=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}d_te^{i\lambda_s t}=\left(\underline{W}_0(\lambda_s),\ldots,\underline{W}_p(\lambda_s)\right)'.$$

The main cases of interest are low order polynomials, where explicit expressions for the discrete Fourier transforms $\underline{W}_k(\lambda_s)$ are easily obtained from Lemma A. Thus, when $k=0,1,2$ we have

$$\underline{W}_0(\lambda_s)=\begin{cases}\sqrt{n}, & s=0,\\ 0, & s\neq 0,\end{cases} \tag{38}$$

$$\underline{W}_1(\lambda_s)=\begin{cases}\dfrac{n+1}{2n^{1/2}}, & s=0,\\ \dfrac{1}{n^{1/2}}\dfrac{e^{i\lambda_s}}{e^{i\lambda_s}-1}, & s\neq 0,\end{cases} \tag{39}$$

$$\underline{W}_2(\lambda_s)=\begin{cases}\dfrac{(n+1)(2n+1)}{6n^{3/2}}, & s=0,\\ \dfrac{1}{n^{1/2}}\dfrac{e^{i\lambda_s}}{e^{i\lambda_s}-1}-\dfrac{2}{n^{3/2}}\dfrac{e^{i\lambda_s}}{(e^{i\lambda_s}-1)^2}, & s\neq 0.\end{cases} \tag{40}$$

In case (38), it is apparent that eliminating the zero frequency will demean the data and leave the model unchanged for $\lambda_s \neq 0$. Then, $\Psi^c Z = 0$ and so $\Psi Q_Z \Psi^c = \Psi \Psi^c - \Psi P_z \Psi^c = 0$, and $Q_Z = Q_V$. It follows that $X' Q_Z \Psi Q_Z \Psi^c \widetilde{X} = 0$ in (30) and therefore $\hat{\beta}_A$ is unbiased in this case of simple data demeaning.

On the other hand, when $z_t' = (1, t)$ and $d_t' = (1, t/n)$, it follows from (38) and (39) that

$$w_d(\lambda_s)' = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} d_t' e^{i\lambda_s t} = \begin{cases} \left(\sqrt{n}, \dfrac{n+1}{2n^{1/2}}\right), & s = 0, \\ \left(0, \dfrac{1}{n^{1/2}} \dfrac{e^{i\lambda_s}}{e^{i\lambda_s} - 1}\right), & s \neq 0, \end{cases} \tag{41}$$

and the second component of $w_d(\lambda_s)$ has nonzero elements for $s \neq 0$. Hence, $\Psi^c Z \neq 0$ and $Q_Z \neq Q_V$, so that $\hat{\beta}_A$ is biased in this case. The same applies for higher order trends.

Frequencies in the band $\mathscr{B}_A^c$ satisfy $|\lambda_s| > \omega_0 > 0$. It therefore follows from Lemma A and (37) that $\underline{W}_k(\lambda_s) = O(n^{-\frac{1}{2}})$ uniformly for $\lambda_s \in \mathscr{B}_A^c$. Hence,

$$w_d(\lambda_s) = O(n^{-1/2}) \qquad \text{for all } \lambda_s \in \mathscr{B}_A^c \tag{42}$$

and, thus, $\Psi^c D = W^* A^c W D$ has elements that are of $O(n^{-1/2})$.

4. FREQUENCY DOMAIN DETRENDING

We consider first the fixed band regression model (22) and (23). The problem of the omitted variable bias evident in (28) can be dealt with either in the frequency domain or in the time domain. In the time domain, one simply detrends using the projection operator $Q_V = I - P_V$, where $V = [\Psi Z, \Psi^c Z]$, as is apparent from the form of (26). In the frequency domain, the alternative is to leave the detrending until the regression is performed in the frequency domain.

To do frequency domain detrending, we simply apply the discrete Fourier transform operator W to (24) and then perform the band spectrum regression. The transformed model is

$$\begin{aligned} Wy &= WZ(\pi_1 - \Pi_2 \beta_A) + W\Psi^c Z \Pi_2 (\beta_A - \beta_{A^c}) + WX\beta_A \\ &\quad - W\Psi^c X(\beta_A - \beta_{A^c}) + W\varepsilon \\ &= WZ(\pi_1 - \Pi_2 \beta_A) + A^c WZ \Pi_2 (\beta_A - \beta_{A^c}) + WX\beta_A \\ &\quad - A^c WX(\beta_A - \beta_{A^c}) + W\varepsilon. \end{aligned}$$

The resulting band spectral estimator for the band $\mathscr{B}_A$ is equivalent to a regression on

$$AWy = AWZ(\pi_1 - \Pi_2 \beta_A) + AWX\beta_A + AW\varepsilon, \tag{43}$$

since $AA^c = 0$, and therefore this estimator has the form

$$\begin{aligned} \hat{\beta}_A^f &= (X'W^* A Q_{AWZ} AWX)^{-1} (X'W^* A Q_{AWZ} AWy) \\ &= \beta_A + (X'W^* A Q_{AWZ} AWX)^{-1} (X'W^* A Q_{AWZ} AW\varepsilon). \end{aligned} \tag{44}$$

Clearly, $E(\hat{\beta}_A^f|X) = \beta_A$, and the estimator is unbiased. A similar result holds for the corresponding estimator $\hat{\beta}_{A^c}^f$ of β_{A^c}.

In this frequency domain approach to detrending, the so called Frisch–Waugh (1933) theorem clearly holds, i.e., the regression coefficient $\hat{\beta}_A^f$ on the variable AWX in (43) is invariant to whether the regressor AWZ is included in the regression or whether all the data have been previously detrended in the frequency domain by regression on AWZ.

Following (32), the natural narrow band approach is to use an augmented regression model that includes the dft of the trend in the regression, viz.

$$w_y(\lambda_s) = \tilde{c}_{1\omega}' w_x(\lambda_s) + \tilde{c}_{2\omega}' w_{\Delta x}(\lambda_s) + \tilde{c}_{3\omega}' w_z(\lambda_s) + \text{residual}, \tag{45}$$

for frequencies λ_s centered on ω. This narrow band regression leads to the estimate

$$\tilde{\beta}_\omega^f = \begin{cases} \tilde{c}_{1,-\omega} - (e^{-i\omega} - 1)\tilde{c}_{2,-\omega}, & \omega \neq 0, \\ \tilde{c}_{10}, & \omega = 0, \end{cases} \tag{46}$$

similar to (34).

5. ASYMPTOTIC THEORY

We derive a limit distribution theory for the detrended band spectral regression estimates and consider what happens to the bias as $n \to \infty$. We start by introducing some notation and making the framework for the limit theory more precise.

We start with the parametric case where there are the two discrete bands $\mathcal{B}_A$ and $\mathcal{B}_{A^c}$. Let $n_a = \#\{\lambda_s \in \mathcal{B}_A\}$ and $n_c = \#\{\lambda_s \in \mathcal{B}_{A^c}\}$ be the number of fundamental frequencies in the bands $\mathcal{B}_A$ and $\mathcal{B}_{A^c}$. It is convenient to subdivide $[-\pi, \pi]$ into subbands $\mathcal{B}_j$ of equal width (say, π/J) that center on frequencies $\{\omega_j = \pi j/J : j = -J+1, \ldots, J-1\}$. Let $m = \#(\lambda_s \in \mathcal{B}_j)$ and suppose that J_a of these bands lie in $\mathcal{B}_A$. It follows that n and n_a can be approximated by $n = 2mJ$ and $n_a = 2mJ_a$, respectively.

In the nonparametric case, we focus on a single frequency ω and consider a shrinking band $\mathcal{B}_\omega$ of width π/J centered on ω. Again, we let $m = \#(\lambda_s \in \mathcal{B}_\omega)$.

The following condition will be useful in the development of the asymptotics and will be taken to hold throughout the remainder of the paper. Additional requirements will be stated as needed.

ASSUMPTION 3: (a) $n_a/n \to \theta$ *and* $n_c/n \to 1 - \theta$ *for some fixed number* $\theta \in [0, 1]$ *as* $n \to \infty$.
(b) $m, J \to \infty$, *and* $J/\sqrt{n} \to 0$ *as* $n \to \infty$.

For the bias in $\hat{\beta}_A$ to vanish asymptotically, the deviation that depends on the term

$$\{X'Q_Z\Psi Q_Z X\}^{-1}\{X'Q_Z\Psi Q_Z\Psi^c\widetilde{X}(\beta_A - \beta_{A^c})\}$$

in (28) needs to disappear as $n \to \infty$. We will distinguish the two cases of stationary and nonstationary (integrated) $\tilde{x}_t$ corresponding to Assumptions 1 and 2 in the following discussion.

5.1. *The Stationary Case*

Here, the bias in $\hat{\beta}_A$ will disappear when

$$\left(\frac{X'Q_Z\Psi Q_Z X}{n}\right)^{-1}\left(\frac{X'Q_Z\Psi Q_Z\Psi^c\widetilde{X}}{n}\right) \xrightarrow{p} 0. \tag{47}$$

A similar requirement, obtained by interchanging Ψ and Ψ^c in (47), holds for the bias in $\hat{\beta}_{A^c}$. We have the following result.

THEOREM 1 (Semiparametric Case): *If $\tilde{x}_t$ and ε_t are zero mean, stationary, and ergodic time series satisfying Assumption* 1, *and $\tilde{y}_t$ is generated by* (22), *band spectral regression with detrending in the time domain or in the frequency domain is consistent for both β_A and β_{A^c}. The common limit distribution of $\hat{\beta}_A$ and $\hat{\beta}^f_A$ is given by*

$$\sqrt{n}(\hat{\beta}_A-\beta_A), \sqrt{n}(\hat{\beta}^f_A-\beta_A) \xrightarrow{d} N(0, V_a),$$

where

$$V_a = \left(\int_{\mathscr{B}_A} f_{xx}(\omega)\,d\omega\right)^{-1}\left(2\pi\int_{\mathscr{B}_A} f_{xx}(\omega)f_{\varepsilon\varepsilon}(\omega)\,d\omega\right)\left(\int_{\mathscr{B}_A} f_{xx}(\omega)\,d\omega\right)^{-1}, \tag{48}$$

with an analogous result for $\hat{\beta}_{A^c}$ and $\hat{\beta}^f_{A^c}$.

In the stationary case, therefore, the bias in band spectral regression from time domain detrending disappears as $n \to \infty$ and there is no difference between the two bands $\mathscr{B}_A$ and $\mathscr{B}^c_A$ in terms of the limit theory. It is therefore irrelevant whether the main focus of interest is high or low frequency regression. The form of the asymptotic covariance matrix V_a is a band spectral version of the familiar sandwich formula for the robust covariance matrix in least squares regression. V_a can be estimated by replacing the spectra in the above formula with corresponding consistent estimates and averaging over the band $\mathscr{B}_A$. For the full band case where $\mathscr{B}_A = [-\pi, \pi]$, the matrix V_a is the well known formula for the asymptotic covariance matrix of the least squares regression estimator in a time series regression (c.f. Hannan (1970, p. 426)). A formula related to V_a was given in Hannan (1963b, equation (16)) for band spectral estimates in the context of models with measurement error.

A similar result holds in the nonparametric case, but with a different convergence rate.

THEOREM 1′ (Nonparametric Case): *If* $\tilde{x}_t$ *and* ε_t *are zero mean, stationary, and ergodic time series satisfying Assumption* 1, *and* $\tilde{y}_t$ *is generated by* (3), *nonparametric band spectral regression with detrending in the time domain or in the frequency domain is consistent for* $\beta_\omega = b(-\omega)$. *The common limit distribution of* $\hat{\beta}_\omega$ *and* $\hat{\beta}^f_\omega$ *is given by*

$$\sqrt{m}(\hat{\beta}_\omega - \beta_\omega), \sqrt{m}(\hat{\beta}^f_\omega - \beta_\omega) \xrightarrow{d} N_c(0, V_\omega),$$

where

$$V_\omega = f_{\varepsilon\varepsilon}(\omega) f_{xx}(\omega)^{-1}.$$

5.2. *The Nonstationary Case*

Here, the distinction between low frequency regression and regression at other frequencies becomes important. In the band regression model (22), there are only two bands $\mathcal{B}_A$ and $\mathcal{B}^c_A$. Over $\mathcal{B}_A$, which includes the zero frequency, the estimator is known to be n-consistent when $\beta_A = \beta_{A^c}$ (see Phillips (1991a)) since in that case, the regression equation is a conventional cointegrating relation. When $\beta_A \neq \beta_{A^c}$, the same result continues to hold over the band $\mathcal{B}_A$, as shown in Theorem 2 below. In this case, the bias in (28) disappears when

$$\left(\frac{X'Q_Z\Psi Q_Z X}{n^2}\right)^{-1}\left(\frac{X'Q_Z\Psi Q_Z\Psi^c\widetilde{X}}{n^2}\right) \xrightarrow{p} 0. \tag{49}$$

Note that in (49) the moment matrices are standardized by n^2, because the data nonstationarity is manifest in bands like $\mathcal{B}_A$ that include the zero frequency. Over frequency bands like $\mathcal{B}^c_A$ that exclude the zero frequency the rate of convergence of the moment matrices is slower and the bias in $\hat{\beta}_{A^c}$ will disappear when

$$\left(\frac{X'Q_Z\Psi^c Q_Z X}{n}\right)^{-1}\left(\frac{X'Q_Z\Psi^c Q_Z\Psi\widetilde{X}}{n}\right) \xrightarrow{p} 0. \tag{50}$$

Our starting point in the $I(1)$ case is to provide some limit theory for discrete Fourier transforms of $I(1)$ and detrended $I(1)$ processes. The following two lemmas do this and provide a limit theory for periodogram averages of such processes. The remainder of the limit theory then follows in a fairly straightforward way from these lemmas. To make the derivations simpler we confine our attention here to the case of a linear trend.

LEMMA B: *Let* $\tilde{x}_t$ *be an* $I(1)$ *process satisfying Assumption* 2. *Then*, *the discrete Fourier transform of* $\tilde{x}_t$ *for* $\lambda_s \neq 0$ *is given by*

$$w_{\tilde{x}}(\lambda_s) = \frac{1}{1-e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}}. \tag{51}$$

Equation (51) shows that the discrete Fourier transforms of an $I(1)$ process are not asymptotically independent across fundamental frequencies. Indeed, they are frequency-wise dependent by virtue of the component $n^{-1/2}\tilde{x}_n$, which produces a common leakage into all frequencies $\lambda_s \neq 0$, even in the limit as $n \to \infty$. As the next lemma shows, this leakage is strong enough to ensure that smoothed periodogram estimates of the spectral density $f_{xx}(\omega) = |1-e^{i\omega}|^{-2} f_{vv}(\omega)$ are inconsistent at frequencies $\omega \in \mathscr{B}_A^c$. Lemma C(f) shows that the leakage is still manifest when the data are first detrended in the time domain. On the other hand, it is apparent from (41) that we can write (51) as

$$w_{\tilde{x}}(\lambda_s) = \frac{1}{1-e^{i\lambda_s}} w_v(\lambda_s) + w_{\frac{t}{n}}(\lambda_s)[\tilde{x}_n - \tilde{x}_0], \tag{52}$$

from which it is clear that frequency domain detrending (i.e., using residuals from regressions of the frequency domain data on $w_{t/n}(\lambda_s)$ or $w_d(\lambda_s)$) will remove the second term of (51) and thereby eliminate the common leakage from the low frequency.

LEMMA C: *Let $\tilde{x}_t$ be an $I(1)$ process satisfying Assumption 2 and $d_t = (1, t/n)$. Define $\tilde{x}'_{.d,t} = \tilde{x}'_t - d'_t(D'D)^{-1}(D'\widetilde{X})$ and let $w_{x.d}(\lambda)$ be the discrete Fourier transform of $\tilde{x}_{.d,t}$. Then, as $n \to \infty$:*

(a) $n^{-2}\sum_{\lambda_s \in \mathscr{B}_A} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_x B_x'$;
(b) $n^{-3/2}\sum_{\lambda_s \in \mathscr{B}_A} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \xrightarrow{d} \int_0^1 B_x u'$;
(c) $n^{-2}\sum_{\lambda_s \in \mathscr{B}_A} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u} B_{x.u}'$;
(d) $n^{-1/2}\sum_{\lambda_s \in \mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \xrightarrow{d} -\frac{1}{2\pi} B_x(1) \int_{\mathscr{B}_A^c} (e^{i\lambda} f_1(\lambda)^*/(1-e^{i\lambda}))\, d\lambda$;
(e) $n^{-1}\sum_{\lambda_s \in \mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \xrightarrow{d} \int_{\mathscr{B}_A^c} \left[f_{xx}(\omega) + \frac{1}{2\pi}(1/(|1-e^{i\omega}|^2)) B_x(1) B_x(1)'\right] d\omega$;
(f) $n^{-1}\sum_{\lambda_s \in \mathscr{B}_A^c} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* \xrightarrow{d} \int_{\mathscr{B}_A^c} \left[f_{xx}(\omega) + (2\pi)^{-1} g(\omega, B_x) g(\omega, B_x)^*\right] d\omega$;
(g) $n^{-1/2}\sum_{\lambda_s \in \mathscr{B}_A^c} w_{x.d}(\lambda_s) w_d(\lambda_s)^* \xrightarrow{d} -\frac{1}{2\pi} \int_{\mathscr{B}_A^c} g(\omega, B_x) f_1(\omega)^*\, d\omega$;
(h) $\sum_{\lambda_s \in \mathscr{B}_A^c} w_d(\lambda_s) w_d(\lambda_s)^* \to \frac{1}{2\pi} \int_{\mathscr{B}_A^c} f_1(\omega) f_1(\omega)^*\, d\omega$;
(i) $n^{-1}\sum_{\lambda_s \in \mathscr{B}_A} w_d(\lambda_s) w_d(\lambda_s)^* \to \int_0^1 uu'$;
(j) $n^{-2}\sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_x B_x'$;
(k) $n^{-2}\sum_{\lambda_s \in \mathscr{B}_0} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u} B_{x.u}'$;
(l) $\frac{1}{n}\sum_{\lambda_s \in \mathscr{B}_0} w_d(\lambda_s) w_d(\lambda_s)^* \to \int_0^1 uu'$;
(m) $n^{-3/2}\sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \xrightarrow{d} \int_0^1 B_x u'$;
(n) $(n/m)\sum_{\lambda_s \in \mathscr{B}_\omega} w_d(\lambda_s) w_d(\lambda_s)^* \to f_1(\omega) f_1(\omega)^*$, $\omega \neq 0$;
(o) $\sqrt{n} m^{-1}\sum_{\lambda_s \in \mathscr{B}_\omega} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \xrightarrow{d} (e^{i\omega}/(1-e^{i\omega})) B(1) f_1(\omega)^*$, $\omega \neq 0$;

where $f_{xx}(\omega) = |1 - e^{i\omega}|^{-2} f_{vv}(\omega)$, $f_1(\omega) = (0, e^{i\omega}/(e^{i\omega} - 1))'$ *for* $\omega \neq 0$, *and*

$$B_{x.u}(r) = B_x(r) - \left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1} u(r),$$

and

$$g(\omega, B_x) = \frac{e^{i\omega}}{1 - e^{i\omega}} B_x(1) + \left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1} f_1(\omega).$$

Joint convergence applies in (a)–(o).

With this result in hand, the $I(1)$ band regression model can be analyzed. The limit theory for time domain detrended band spectral regression is as follows.

THEOREM 2: *Suppose* $(\varepsilon_t, \tilde{x}_t)$ *satisfies Assumption* 2 *and* $\tilde{y}_t$ *is generated by* (22). *Then*

$$n(\hat{\beta}_A - \beta_A) \xrightarrow{d} \left(\int_0^1 B_{x.u} B'_{x.u}\right)^{-1} \left(\int_0^1 B_{x.u} dB_\varepsilon\right) = MN\left(0, \left(\int_0^1 B_{x.u} B'_{x.u}\right)^{-1} 2\pi f_{\varepsilon\varepsilon}(0)\right), \tag{53}$$

and

$$\hat{\beta}_{A^c} \xrightarrow{d} \beta_{A^c} - \Xi^{-1} \xi (\beta_{A^c} - \beta_A), \tag{54}$$

where

$$\Xi = \left[\int_{\mathcal{B}_A^c} [2\pi f_{xx}(\omega) + g(\omega, B_x) g(\omega, B_x)^*]\, d\omega\right],$$

and

$$\xi = \left[\left(\int_{\mathcal{B}_A^c} g(\omega, B_x) f_1(\omega)^*\, d\omega\right)\left(\int_0^1 uu'\right)^{-1} \int_0^1 u B'_x\right]. \tag{55}$$

So, when the regressors are $I(1)$, band spectral regression is inconsistent in frequency bands that exclude the origin when detrending is performed in the time domain prior to frequency domain regression. The bias is random and is linear in the differential, $\beta_{A^c} - \beta_A$, between the coefficients in the frequency bands. In the case of a linear trend $z_t = t$, the limit function $f_1(\omega) = e^{i\omega}/(e^{i\omega} - 1)$ and then the bias depends on

$$g(\omega, B_x) = \frac{e^{i\omega}}{1 - e^{i\omega}}\left[B_x(1) - \left(\int_0^1 B_x r\right)\left(\int_0^1 r^2\right)^{-1}\right].$$

When $f_{xx}(\omega) = (1/2\pi)|1-e^{i\omega}|^{-2}$ (i.e., when v_t is iid(0, 1)), a simple calculation reveals that the probability limit (54) of $\hat{\beta}_{A^c}$ simplifies to

$$(56) \qquad \beta_{A^c} + (\beta_A - \beta_{A^c}) \frac{\left[\left(\int_0^1 B_x r\right)\left(\int_0^1 r^2\right)^{-1} - B_x(1)\right]\left(\int_0^1 r^2\right)^{-1}\left(\int_0^1 B_x r\right)}{\left(1 + \left[B_x(1) - \left(\int_0^1 B_x r\right)\left(\int_0^1 r^2\right)^{-1}\right]^2\right)}.$$

So, in this case, the bias is not dependent on the width of the band β_{A^c}.

The following theorem gives the corresponding results for the nonparametric estimators $\hat{\beta}_\omega$ and $\tilde{\beta}_\omega$.

THEOREM 2′: *Suppose* $(\varepsilon_t, \tilde{x}_t)$ *satisfies Assumption* 2 *and* $\tilde{y}_t$ *is generated by* (11) *where* $\beta(L)$ *and* $\tilde{\beta}(L)$ *both have valid BN decompositions.*

(a) *For* $\omega = 0$,

$$(57) \qquad n(\hat{\beta}_0 - \beta_0) \xrightarrow{d} \left[\int_0^1 B_{x.u} B_{x.u}'\right]^{-1}\left[\int_0^1 B_{x.u} dB_\varepsilon - \left\{\int_0^1 B_{x.u} dB_x' + \Delta_x\right\}\tilde{\beta}(1)\right]$$

where $\beta_0 = \beta(1)$ *and* $\Delta_x = \sum_{j=-\infty}^{\infty} E(v_0 v_j')$. *For* $\omega \neq 0$,

$$(58) \qquad \hat{\beta}_\omega \to_d \beta_\omega + \left[2\pi \Xi_\omega^{-1} f_{xx}(\omega) - 1\right]\tilde{\beta}(e^{-i\omega})(e^{-i\omega} - 1),$$

where $\beta_\omega = \beta(e^{-i\omega}) = b(-\omega)$,

$$\Xi_\omega = 2\pi f_{xx}(\omega) + g(\omega, B_x) g(\omega, B_x)^*,$$

and $f_{xx}(\omega) = |1-e^{i\omega}|^{-2} f_{vv}(\omega)$.

(b) *For* $\omega = 0$,

$$n(\tilde{\beta}_0 - \beta_0) \xrightarrow{d} \left[\int_0^1 B_{x.u} B_{x.u}'\right]^{-1}\left[\int_0^1 B_{x.u} dB_\varepsilon\right].$$

For $\omega \neq 0$,

$$\sqrt{m}(\tilde{\beta}_\omega - \beta_\omega) \to_d N_c\left(0, \frac{f_{\varepsilon\varepsilon}(\omega)}{f_{xx}(\omega)}\right).$$

REMARKS: (i) In Theorem 2′(a) it is apparent that the limit distribution of $\hat{\beta}_0$ has a second order bias term involving

$$\left\{\int_0^1 B_{x.u}\, dB_x' + \Delta_x\right\}\tilde{\beta}(1),$$

which is nonzero except when $\tilde{\beta}(1) = 0$, a circumstance that arises when the filter $\beta(L)$ is symmetric (see Remark (ii) following Lemma D in the Appendix) as it is in the case (19). In that case, the limit (57) reduces to the same mixed normal distribution as (53). Part (a) shows also that the estimator $\hat{\beta}_\omega$ is inconsistent at all frequencies $\omega \neq 0$ except when $\tilde{\beta}(e^{-i\omega}) = 0$. The bias in $\hat{\beta}_0$ and inconsistency in $\hat{\beta}_\omega$ are due to omitted variable misspecification in the regression.

(ii) Part (b) of Theorem 2′ shows that the estimator $\tilde{\beta}_\omega$ based on the augmented regression (32) is consistent for both $\omega = 0$, and $\omega \neq 0$. The same mixed normal distribution as (53) applies when $\omega = 0$. When $\omega \neq 0$ the limit distribution is the same as in the stationary case given in Theorem 1′.

The following results give the limit theory for the frequency domain detrended estimators in the fixed band and narrow band cases for nonstationary data.

THEOREM 3: *Under the conditions of Theorem* 2

$$n(\hat{\beta}^f_A - \beta_A) \xrightarrow{d} \left(\int_0^1 B_{x.u}B'_{x.u}\right)^{-1}\left(\int_0^1 B_{x.u}\, dB_\varepsilon\right) \tag{59}$$

and

$$\sqrt{n}(\hat{\beta}^f_{A^c} - \beta_{A^c}) \xrightarrow{d} N\Bigg(0, \left[\int_{\mathscr{B}^c_A} f_{xx}(\omega)\, d\omega\right]^{-1}\left[2\pi\int_{\mathscr{B}^c_A} f_{xx}(\omega) f_{\varepsilon\varepsilon}(\omega)\, d\omega\right] \times \left[\int_{\mathscr{B}^c_A} f_{xx}(\omega)\, d\omega\right]^{-1}\Bigg). \tag{60}$$

In fixed band regression models there is usually some advantage to be gained by averaging over the band and using weighted regression, as shown in Hannan's (1963a, b) original work. Efficient regression is based on a weighted band spectral regression that uses a preliminary regression to obtain estimates of the equation errors and a corresponding estimate of the error spectrum, say $\hat{f}_{\varepsilon\varepsilon}(\omega)$, that is uniformly consistent so that $\sup_\omega |\hat{f}_{\varepsilon\varepsilon}(\omega) - f_{\varepsilon\varepsilon}(\omega)| \to_p 0$ (e.g., Hannan (1970, p. 488)). When such weighted regression is performed with frequency domain detrending, the resulting estimates have optimal properties for both the nonstationary component and the stationary component. For β_A, the limit distribution is the same as (59) and is optimal in the sense of Phillips (1991b). For β_{A^c}, the limit distribution of the efficient estimate is normal with variance matrix $2\pi[\int_{\mathscr{B}^c_A} f_{xx}(\omega) f_{\varepsilon\varepsilon}(\omega)^{-1}\, d\omega]^{-1}$ and therefore attains the usual efficiency bound in time series regression (e.g., Hannan (1970, eqn. (3.4), p. 427)) adjusted here for band limited regression (Hannan (1963, eqn. (16))). Details are omitted and are available in an earlier version of the present paper.

THEOREM 3′: *Under the conditions of Theorem* 2′, *at* $\omega = 0$

$$n(\tilde{\beta}^f_0 - \beta_0) \xrightarrow{d} \left(\int_0^1 B_{x.u}B'_{x.u}\right)^{-1}\left(\int_0^1 B_{x.u}\, dB_\varepsilon\right)$$

and for $\omega \neq 0$

$$\sqrt{m}(\tilde{\beta}^f_\omega - \beta_\omega) \xrightarrow{d} N_c(0, f_{\varepsilon\varepsilon}(\omega) f_{xx}(\omega)^{-1}). \tag{61}$$

So, $\hat{\beta}^f_A$ and $\tilde{\beta}^f_0$ are consistent and have the same mixed normal limit distribution as that of $\hat{\beta}_A$ in (53). The limit distribution makes asymptotic inference about β_A and β_0 straightforward, using conventional regression Wald tests adjusted in the usual fashion so that a consistent estimate of the spectrum of ε_t is used (based on regression residuals) in place of a variance estimate.

The frequency domain detrended estimators $\hat{\beta}^f_{A^c}$ and $\hat{\beta}^f_\omega$ are also consistent, unlike the time detrended estimator $\hat{\beta}_\omega\,(\omega \neq 0)$ in the nonstationary case. The limit distribution of $\hat{\beta}^f_{A^c}$ is the same as it is in the case of trend stationary regressors (Theorem 1) and has the same $\sqrt{n}$ rate. The nonparametric estimator $\tilde{\beta}^f_\omega$ is $\sqrt{m}$ consistent and is asymptotically equivalent to the augmented regression estimator $\tilde{\beta}_\omega$. As far as the model (11) is concerned, it is therefore apparent from Theorems 2′ and 3′ that if the correct augmented regression model is used in estimation, it does not matter asymptotically whether detrending is done in the time domain or the frequency domain.

6. CONCLUSION

It is natural to eliminate deterministic trends in the time domain by simple least squares regression because the Grenander–Rosenblatt (1957, p. 244) theorem shows that such regression is asymptotically efficient when the time series are trend stationary (although this conclusion does not hold when there are stochastic as well as deterministic trends—see Phillips and Lee (1996)). In a similar way, it seems natural to eliminate deterministic trends in band spectral regressions by detrending in the time domain prior to the use of spectral methods because these methods were originally intended for application to stationary time series. However, this paper shows that such time domain detrending will lead to biased coefficient estimates in models where the coefficients are frequency dependent. In models that have both deterministic and stochastic trends, time domain detrending can lead to inconsistent estimates of the coefficients at frequency bands away from the origin. The inconsistency, which arises from omitted variable effects, can be substantial and has been confirmed in simulations that are not reported here (Corbae, Ouliaris, and Phillips (1997)).

The bias and inconsistency arise from omitted variable misspecification and are managed by use of an appropriate augmented regression model. The situation is analogous to a structural break model, but here the coefficients change across frequency rather than over time. An alternate approach that is suitable in practice is to model the data and run regressions, including detrending regressions, in the frequency domain. In effect, discrete Fourier transforms of all the variables in the model, including the deterministic trends, are taken and band regression is performed. When nonparametric estimation is being conducted, the same principle is employed but one uses a shrinking band that is local to a particular frequency. In the nonstationary case, it turns out to be particularly important to specify the model in terms of levels and differences as in (15) leading to the fitted regressions (32) and (45). An estimate of the frequency response coefficient at a particular frequency is then recovered from a linear combination of the coefficients in the regression as in (34) and (46). This approach, which can be regarded as a frequency domain version of leads and lags dynamic regression, provides a convenient single equation method of estimating a long run relationship in the presence of deterministic trends and short run dynamics.

Dept. of Economics, University of Texas at Austin, Austin, TX 78712, U.S.A.; corbae@eco.utexas.edu; http://www.eco.utexas.edu/~corbae

School of Business, National University of Singapore, Singapore 117591; and Research Dept., IMF, Washington DC, U.S.A.; souliaris@nus.edu.sg

and

Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520-8281, U.S.A.; and University of Auckland and University of York; peter.phillips@yale.edu; http://Korora.econ.yale.edu

Manuscript received September, 1997; final revision received June, 2001.

APPENDIX

LEMMA D (Two Sided BN Decompositions): *If* $C(L) = \sum_{j=-\infty}^{\infty} c_j L^j$ *and* $\sum_{j=-\infty}^{\infty} |j|^{\frac{1}{2}} \|c_j\| < \infty$, *then*:

(a) $C(L) = C(1) + \tilde{C}(L)(1-L)$, *where* $\widetilde{C}(L) = \sum_{j=-\infty}^{\infty} \tilde{c}_j L^j$, *with*

$$\tilde{c}_j = \begin{cases} \sum_{k=j+1}^{\infty} c_k, & j \geq 0, \\ -\sum_{k=-\infty}^{j} c_k, & j < 0; \end{cases}$$

(b) $C(L) = C(e^{iw}) + \widetilde{C}_w(e^{-iw}L)(e^{-iw}L - 1)$, *where* $\widetilde{C}_w(L) = \sum_{j=-\infty}^{\infty} \tilde{c}_{wj} L^j$, *with*

$$\tilde{c}_{wj} = \begin{cases} \sum_{k=j+1}^{\infty} c_k e^{ikw}, & j \geq 0, \\ -\sum_{k=-\infty}^{j} c_k e^{ikw}, & j < 0. \end{cases}$$

PROOF OF LEMMA D: The proof is along the same lines as the proof in Phillips and Solo (1992, Lemma 2.1) of the one sided BN decomposition.

REMARKS: (i) The condition

$$\sum_{j=-\infty}^{\infty} |j|^{\frac{1}{2}} \|c_j\| < \infty \tag{62}$$

is sufficient for (a) and (b) and ensures that $\sum_{j=-\infty}^{\infty} \tilde{c}_j^2 < \infty$ but it is not necessary for the latter. For instance, the series $\beta(L) = \sum_{j=-\infty}^{\infty} \beta_j L^j$, whose coefficients β_j are given by (19), fails (62) but still has a valid BN decomposition with $\sum_{j=-\infty}^{\infty} \tilde{\beta}_j^2 < \infty$. To see this, let $c_j = e^{ijx}/j$ and then β_j is a constant times the imaginary part of c_j. Define $S_n = \sum_{j=1}^{n} e^{ijx} = (e^{ix}/(e^{ix}-1))(e^{inx}-1)$. Partial summation gives

$$P_n = \sum_{j=1}^{n} \frac{e^{ijx}}{j} = \frac{1}{n} S_n + \frac{e^{ix}}{e^{ix}-1} \sum_{j=2}^{n} \frac{e^{i(j-1)x}-1}{j(j-1)}.$$

It follows that

$$P_n - P_m = \frac{S_n}{n} - \frac{S_m}{m} + \frac{e^{ix}}{e^{ix}-1}\sum_{j=m+1}^{n}\frac{e^{i(j-1)x}-1}{j(j-1)},$$

and letting $n \to \infty$ we have

$$\tilde{c}_m = \sum_{j=m+1}^{\infty}\frac{e^{ijx}}{j} = -\frac{S_m}{m} + \frac{e^{ix}}{e^{ix}-1}\sum_{j=m+1}^{\infty}\frac{e^{i(j-1)x}-1}{j(j-1)} = O\left(\frac{1}{m}\right),$$

with a similar result for $\tilde{c}_{-m}$. We deduce that $\sum_{-\infty}^{\infty}|\tilde{c}_m|^2 < \infty$. Hence, $\beta(L) = \sum_{j=-\infty}^{\infty}\beta_j L^j$ with β_j given by (19) has a valid BN decomposition with $\sum_{j=-\infty}^{\infty}\tilde{\beta}_j^2 < \infty$.

(ii) The case of a symmetric filter with $c_j = c_{-j}$ for $j \neq 0$ is an important specialization, which includes the series with coefficients β_j given by (19). In this case, $-\tilde{c}_{-m} = \sum_{j=m}^{\infty} c_{-j} = \sum_{j=m}^{\infty} c_j = c_m + \tilde{c}_m$. Then, $\widetilde{C}(1) = \sum_{j=1}^{\infty}(\tilde{c}_j + \tilde{c}_{-j}) + \tilde{c}_0 = \sum_{j=1}^{\infty}(\tilde{c}_j - \tilde{c}_j - c_j) + \sum_{j=1}^{\infty} c_j = 0$. If $\widetilde{C}(L)$ has a valid BN decomposition, we deduce that

$$C(L) = C(1) + \tilde{\tilde{C}}(L)(1-L)^2, \quad \tilde{\tilde{C}}(L) = \sum_{j=-\infty}^{\infty}\tilde{\tilde{c}}_j L^j, \tag{63}$$

where

$$\tilde{\tilde{c}}_j = \begin{cases} \displaystyle\sum_{k=j+1}^{\infty}\tilde{c}_k & j \geq 0, \\ \displaystyle -\sum_{k=-\infty}^{j}\tilde{c}_k & j < 0. \end{cases}$$

A sufficient condition for the validity of (63) is $\sum_{j=-\infty}^{\infty}|j|^{\frac{1}{2}}\|\tilde{c}_j\| < \infty$ or $\sum_{j=-\infty}^{\infty}|j|^{\frac{3}{2}}\|c_j\| < \infty$ in terms of the original coefficients.

PROOF OF LEMMA A: Part (a) is well known (e.g., Gradshteyn and Ryzhik (1965, formula 0.121)). For part (b) we use partial summation to give

$$\begin{aligned}\Delta\left(\sum_{t=1}^{n} t^k e^{i\lambda_s t}\right) &= \sum_{t=1}^{n}(\Delta t^k)e^{i\lambda_s t} + \sum_{t=1}^{n}(t-1)^k\left(\Delta e^{i\lambda_s t}\right) \\ &= \sum_{t=1}^{n}\left\{-\sum_{j=1}^{k}\binom{k}{j}t^{k-j}(-1)^j\right\}e^{i\lambda_s t} + \sum_{t=1}^{n}(t-1)^k e^{i\lambda_s(t-1)}\left(e^{i\lambda_s}-1\right),\end{aligned}$$

or

$$n^k = \sum_{t=1}^{n}\left\{-\sum_{j=1}^{k}\binom{k}{j}t^{k-j}(-1)^j\right\}e^{i\lambda_s t} + \sum_{t=1}^{n}(t-1)^k e^{i\lambda_s(t-1)}\left(e^{i\lambda_s}-1\right).$$

Using this formula in the identity

$$\left(\sum_{t=1}^{n} t^k e^{i\lambda_s t}\right)\left(e^{i\lambda_s}-1\right) = n^k\left(e^{i\lambda_s}-1\right) + \sum_{t=1}^{n}(t-1)^k e^{i\lambda_s(t-1)}\left(e^{i\lambda_s}-1\right),$$

we get

$$\begin{aligned}\left(\sum_{t=1}^{n} t^k e^{i\lambda_s t}\right)\left(e^{i\lambda_s}-1\right) &= n^k\left(e^{i\lambda_s}-1\right) + n^k - \sum_{t=1}^{n}\left\{-\sum_{j=1}^{k}\binom{k}{j}t^{k-j}(-1)^j\right\}e^{i\lambda_s t} \\ &= n^k e^{i\lambda_s} + \sum_{j=1}^{k}\binom{k}{j}(-1)^j\left(\sum_{t=1}^{n} t^{k-j}e^{i\lambda_s t}\right),\end{aligned}$$

giving the recursion

$$W_k(\lambda_s) = n^k \frac{e^{i\lambda_s}}{e^{i\lambda_s}-1} + \frac{1}{e^{i\lambda_s}-1}\sum_{j=1}^{k}\binom{k}{j}(-1)^j W_{k-j}(\lambda_s),$$

which holds for $s = 1, \dots, n$ (i.e., $e^{i\lambda_s} \neq 1$).

The initial condition

$$W_0(\lambda_s) = \sum_{t=1}^{n} e^{i\lambda_s t} = \begin{cases} n, & s = 0, \\ 0, & s \neq 0, \end{cases}$$

follows by elementary calculation. For higher order trends with $k = 1, 2, 3$ we get

$$W_1(\lambda_s) = \begin{cases} \dfrac{n(n+1)}{2}, & s = 0, \\ n\dfrac{e^{i\lambda_s}}{e^{i\lambda_s}-1}, & s \neq 0, \end{cases}$$

$$W_2(\lambda_s) = \begin{cases} \dfrac{n(n+1)(2n+1)}{6}, & s = 0, \\ n\dfrac{e^{i\lambda_s}}{e^{i\lambda_s}-1}\left[n - \dfrac{2}{e^{i\lambda_s}-1}\right], & s \neq 0, \end{cases}$$

and

$$W_3(\lambda_s) = \begin{cases} \left(\dfrac{n(n+1)}{2}\right)^2, & s = 0, \\ n\dfrac{e^{i\lambda_s}}{e^{i\lambda_s}-1}\left[n^2 - 3(n-1)\dfrac{1}{(e^{i\lambda_s}-1)} + 6\dfrac{1}{(e^{i\lambda_s}-1)^2}\right], & s \neq 0, \end{cases}$$

which lead to the expressions for $\underline{W}_k(\lambda_s)$, $k = 0, 1, 2$, that are given in Section 2 following Lemma A.

Proof of Theorem 1: This follows standard lines and is omitted (see Corbae, Ouliaris, and Phillips (1997)).

Proof of Theorem 1′: First, observe that

$$\hat{\beta}_\omega = \{m^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\widetilde{X}\}^{-1}\{m^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\tilde{y}\}. \tag{64}$$

Next, from (8) and (9) we have

$$w_{\tilde{y}}(\lambda_s) = b(\lambda_s)'w_{\tilde{x}}(\lambda_s) + w_\varepsilon(\lambda_s) + \frac{1}{\sqrt{n}}[e^{-i\lambda_s}\tilde{a}_{\lambda_s 0} - \tilde{a}_{\lambda_s n}] \tag{65}$$

$$= b(\omega)'w_{\tilde{x}}(\lambda_s) + w_\varepsilon(\lambda_s) + O_p\left(\frac{1}{\sqrt{n}}\right) \tag{66}$$

in the stationary case. We now find that

$$m^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\widetilde{X} = m^{-1}\sum_{\lambda_s\in\mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* + o_p(1) \to_p 2\pi f_{xx}(\omega), \tag{67}$$

and using (66) we have

$$\begin{aligned} m^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\tilde{y} &= m^{-1}\widetilde{X}'\Psi_\omega\tilde{y} + o_p(1) \\ &= m^{-1}\sum_{\lambda_s\in\mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* b(-\omega) + m^{-1}\sum_{\lambda_s\in\mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s)w_\varepsilon(\lambda_s)^* + o_p(1). \end{aligned} \tag{68}$$

Then, from (64), (67), and (68) we have the expression

$$\sqrt{m}(\hat{\beta}_\omega - \beta_\omega) = \left\{ m^{-1} \sum_{\lambda_s \in \mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \right\}^{-1} \left\{ m^{-1/2} \sum_{\lambda_s \in \mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s) w_\varepsilon(\lambda_s)^* \right\} + o_p(1). \tag{69}$$

The family $\{w_\varepsilon(\lambda_s) : \lambda_s \in \mathcal{B}_\omega\}$ are known to satisfy a central limit theorem (with limit $N_c(0, 2\pi f_{\varepsilon\varepsilon}(\omega))$) for dft's of stationary processes and are independently distributed[2] as $n \to \infty$. It follows from this result and (67) that

$$m^{-1/2} \sum_{\lambda_s \in \mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s) w_\varepsilon(\lambda_s)^* \to_d N_c(0, (2\pi)^2 f_{\varepsilon\varepsilon}(\omega) f_{xx}(\omega)), \tag{70}$$

which leads to the stated limit theorem for $\hat{\beta}_\omega$. A similar argument gives the result for $\hat{\beta}^f_\omega$.

PROOF OF LEMMA B: Take dft's of the equation $\Delta \tilde{x}_t = \nu_t$, giving

$$w_{\tilde{x}}(\lambda_s) = n^{-1/2} \sum_{t=1}^{n} \tilde{x}_{t-1} e^{i\lambda_s t} + w_v(\lambda_s) = e^{i\lambda_s} \left[w_{\tilde{x}}(\lambda_s) - n^{-1/2} \left(e^{i\lambda_s n} \tilde{x}_n - \tilde{x}_0 \right) \right] + w_v(\lambda_s).$$

Then, transposing and solving yields the stated formula

$$w_{\tilde{x}}(\lambda_s) = \frac{1}{1 - e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}}. \tag{71}$$

PROOF OF LEMMA C: *Part (a)*: This result may be proved as in Phillips (1991a). A new and substantially simpler proof uses part (e) and is as follows.

$$\begin{aligned} n^{-2} \sum_{\lambda_s \in \mathcal{B}_A} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* &= n^{-2} \sum_{\lambda_s \in [-\pi, \pi]} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* - n^{-2} \sum_{\lambda_s \in \mathcal{B}_A^c} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \\ &= n^{-2} \sum_{t=1}^{n} \tilde{x}_t \tilde{x}_t' + O_p(n^{-1}) \\ &\xrightarrow{d} \int_0^1 B_x B_x', \end{aligned}$$

where the error magnitude in the second line follows directly from part (e) below.

Part (b): First note from (41) that $w_d(0)' = \left(\sqrt{n}, (n+1)/2n^{1/2}\right)$, so that $n^{-1/2} w_d(0)' \to \left(1, \frac{1}{2}\right) = f_0' = \int_0^1 u'$, and

$$w_d(\lambda_s)' = \left(0, \frac{1}{n^{1/2}} \frac{e^{i\lambda_s}}{e^{i\lambda_s - 1}}\right) = \frac{1}{\sqrt{n}} f_1(\lambda_s)', \quad s \neq 0.$$

It follows that as $n \to \infty$

$$w_d(\lambda_s)' = \left(0, \frac{\sqrt{n}}{2\pi s i} [1 + o(1)]\right), \tag{72}$$

[2] Hannan (1973, Theorem 3) showed that a finite collection of $w_\varepsilon(\lambda_s)$ satisfy a central limit theorem and are independent. Here the collection $\{\lambda_s \in \mathcal{B}_\omega\}$ has m members and is asymptotically infinite. Phillips (2000, Theorem 3.2) 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 neighborhood of the origin provided the number of frequencies $m = o(n^{-\frac{1}{2} + \frac{1}{p}})$ where $E(|\varepsilon_t|^p) < \infty$ for some $p > 2$, i.e., provided the number of frequencies does not go to infinity too fast. This result can be extended to frequency bands away from the origin, although a proof was not given in that paper.

a formula that holds for both s fixed and for $s \to \infty$ with $(s/n) \to 0$. On the other hand, when $\lambda_s \to \lambda \neq 0$ as $n \to \infty$ we have

$$w_d(\lambda_s)' = \frac{1}{\sqrt{n}}\left(0, \frac{e^{i\lambda}}{e^{i\lambda}-1}\right) + o\left(\frac{1}{\sqrt{n}}\right) = \frac{1}{\sqrt{n}} f_1(\lambda)' + o\left(\frac{1}{\sqrt{n}}\right). \tag{73}$$

Write the summation over $\mathcal{B}_A$ as follows: $\sum_{\lambda_s \in \mathcal{B}_A} = \sum_{\lambda_s \in [-\pi, \pi]} - \sum_{\lambda_s \in \mathcal{B}_A{}^c}$. Then

$$\begin{aligned} n^{-3/2}\left(\sum_{\lambda_s \in \mathcal{B}_A} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^*\right) &= n^{-3/2} \sum_{\lambda_s \in [-\pi, \pi]} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* - n^{-3/2} \sum_{\lambda_s \in \mathcal{B}_A{}^c} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \\ &= n^{-3/2} \sum_{\lambda_s \in [-\pi, \pi]} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* - n^{-2} \sum_{\lambda_s \in \mathcal{B}_A{}^c} w_{\tilde{x}}(\lambda_s) f_1(\lambda_s)^*. \end{aligned}$$

First,

$$\begin{aligned} n^{-2} \sum_{\lambda_s \in \mathcal{B}_A{}^c} w_{\tilde{x}}(\lambda_s) f_1(\lambda_s)^* &= n^{-2} \sum_{\lambda_s \in \mathcal{B}_A{}^c} \left[\frac{1}{1-e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}}\right] f_1(\lambda_s)^* \\ &= O_p(n^{-1}). \end{aligned}$$

Second,

$$\begin{aligned} n^{-3/2} \sum_{\lambda_s \in [-\pi, \pi]} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* &= n^{-2} \sum_{t=1}^{n} \tilde{x}_t \sum_{\lambda_s \in [-\pi, \pi]} e^{i\lambda_s t} w_d(\lambda_s)^* \\ &= n^{-3/2} \sum_{t=1}^{n} \tilde{x}_t d_t' = n^{-1} \sum_{t=1}^{n} \frac{\tilde{x}_t}{\sqrt{n}} d_t' \\ &\xrightarrow{d} \int_0^1 B_x(r) u(r)'\, dr. \end{aligned}$$

It follows that

$$n^{-3/2}\left(\sum_{\lambda_s \in \mathcal{B}_A} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^*\right) \xrightarrow{d} \int_0^1 B_x(r) u(r)'\, dr.$$

Part (c): First, observe that

$$n^{-1/2} \tilde{x}'_{.d,[nr]} \xrightarrow{d} B_x(r)' - u(r)'\left(\int_0^1 uu'\right)^{-1} \int_0^1 u B_x' = B_{x.u}(r)'.$$

Then

$$\frac{1}{n^2}\left(\sum_{\lambda_s \in \mathcal{B}_A} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^*\right) \xrightarrow{d} \int_0^1 B_{x.u} B_{x.u}',$$

as in part (a).

Part (d): From Lemma B we have

$$w_{\tilde{x}}(\lambda_s) = \frac{1}{1-e^{i\lambda_s}} w_v(\lambda_s) - \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}}.$$

As in the proof of part (b), we have $n^{-1/2}\tilde{x}_n \xrightarrow{d} B_x(1)$, and, if $\lambda_s \to \lambda \neq 0$ as $n \to \infty$,

$$\begin{aligned} w_{\tilde{x}}(\lambda_s) &\xrightarrow{d} \frac{1}{1-e^{i\lambda}} N_c(0, 2\pi f_{vv}(\lambda)) - \frac{e^{i\lambda}}{1-e^{i\lambda}} N(0, 2\pi f_{vv}(0)) \\ &= N_c\left(0, 2\pi \frac{f_{vv}(\lambda) + f_{vv}(0)}{|1-e^{i\lambda}|^2}\right). \end{aligned}$$

It follows that

$$\begin{aligned} n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^* &= n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)f_1(\lambda_s)^* \\ &= n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c}\left(\frac{1}{1-e^{i\lambda_s}}\right)w_\nu(\lambda_s)f_1(\lambda_s)^* - n^{-1}\frac{[\tilde{x}_n-\tilde{x}_0]}{n^{1/2}}\sum_{\lambda_s\in\mathscr{B}_A^c}\left(\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\right)f_1(\lambda_s)^* \\ &= \text{term I} + \text{term II, say.} \end{aligned}$$

Since the $w_\nu(\lambda_s)$ are asymptotically independent $N_c(0, 2\pi f_{\nu\nu}(\lambda_s))$, term I $\xrightarrow{p} 0$. For term II, since $\tilde{x}_0 = O_p(1)$, we have

$$\begin{aligned} n^{-1}\frac{[\tilde{x}_n-\tilde{x}_0]}{n^{1/2}}\sum_{\lambda_s\in\mathscr{B}_A^c}\left(\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\right)f_1(\lambda_s)^* &\sim \frac{1}{2\pi}\frac{\tilde{x}_n}{n^{1/2}}\frac{2\pi}{n}\sum_{\lambda_s\in\mathscr{B}_A^c}\left(\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\right)f_1(\lambda_s)^* \\ &\xrightarrow{d} \frac{1}{2\pi}\left(B_x(1)\int_{\mathscr{B}_A^c}\frac{e^{i\lambda}f_1(\lambda)^*}{1-e^{i\lambda}}\,d\lambda\right) \end{aligned}$$

so that

$$\text{term II} \xrightarrow{d} -\frac{1}{2\pi}\left(B_x(1)\int_{\mathscr{B}_A^c}\frac{e^{i\lambda}f_1(\lambda)^*}{1-e^{i\lambda}}\,d\lambda\right), \tag{74}$$

giving part (d).

Part (e): From Lemma B we get

$$\begin{aligned} n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* &\overset{d}{\sim} n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c}\frac{1}{|1-e^{i\lambda_s}|^2}\left[w_v(\lambda_s)w_v(\lambda_s)^* + \frac{\tilde{x}_n}{n^{1/2}}\frac{\tilde{x}_n'}{n^{1/2}}\right] \\ &= \frac{1}{2J}\sum_{J_a\le|j|\le J}\frac{1}{m}\sum_{\lambda_s\in\mathscr{B}_j}\frac{1}{|1-e^{i\lambda_s}|^2}\left[w_v(\lambda_s)w_v(\lambda_s)^* + \frac{\tilde{x}_n}{n^{1/2}}\frac{\tilde{x}_n'}{n^{1/2}}\right] + o_p(1) \\ &= \frac{1}{2J}\sum_{J_a\le|j|\le J}\frac{1}{|1-e^{i\omega_j}|^2}2\pi\hat{f}_{vv}(\omega_j) + n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c}\frac{1}{|1-e^{i\lambda_s}|^2}\frac{\tilde{x}_n}{n^{1/2}}\frac{\tilde{x}_n'}{n^{1/2}} + o_p(1), \\ &\xrightarrow{d} \int_{\mathscr{B}_A^c}\frac{f_{vv}(\omega)}{|1-e^{i\omega}|^2}\,d\omega + \frac{1}{2\pi}\int_{\mathscr{B}_A^c}\frac{1}{|1-e^{i\omega}|^2}\,d\omega\, B_x(1)B_x(1)', \end{aligned}$$

as required. In the penultimate line above, $2\pi\hat{f}_{vv}(\omega_j) = (1/m)\sum_{\lambda_s\in\mathscr{B}_j} w_\nu(\lambda_s)w_\nu(\lambda_s)^*$ and $\hat{f}_{vv}(\omega) - f_{vv}(\omega) \to_p 0$.

Part (f): Note that

$$\begin{aligned} w_{x.d}(\lambda_s)^* &= w_{\tilde{x}}(\lambda_s)^* - w_d(\lambda_s)^*(n^{-1}D'D)^{-1}(n^{-1}D'\widetilde{X}) \\ &= \frac{1}{1-e^{-i\lambda_s}}\,w_v(\lambda_s)^* - \frac{e^{-i\lambda_s}}{1-e^{-i\lambda_s}}\frac{[\tilde{x}_n-\tilde{x}_0]'}{n^{1/2}} - n^{1/2}w_d(\lambda_s)^*(n^{-1}D'D)^{-1}(n^{-3/2}D'\widetilde{X}) \\ &\xrightarrow{d} \frac{1}{1-e^{-i\lambda}}\,N_c(0, 2\pi f_{\nu\nu}(\lambda)) - \frac{e^{-i\lambda}}{1-e^{-i\lambda}}B_x(1)' - f_1(\lambda)^*\left(\int_0^1 uu'\right)^{-1}\left(\int_0^1 uB_x'\right) \\ &= \frac{1}{1-e^{-i\lambda}}\,N_c(0, 2\pi f_{\nu\nu}(\lambda)) - g(\lambda, B_x)^*, \quad \text{say} \end{aligned} \tag{75}$$

when $\lambda_s \to \lambda \neq 0$ as $n\to\infty$. Here

$$g(\lambda, B_x) = \frac{e^{i\lambda}}{1-e^{i\lambda}}B_x(1) + \left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1} f_1(\lambda),$$

and the complex normal variate $N_c(0, 2\pi f_{\nu\nu}(\lambda))$ in (75) is independent of the Brownian motion B_x. If $\omega_j = \pi j/J$ is the midpoint of the band $\mathscr{B}_j$ and $\omega_j \to \omega$ as $n \to \infty$, then

$$2\pi \hat{f}_{xx.d}(\omega_j) = \frac{1}{m} \sum_{\lambda_s \in \mathscr{B}_j} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* \xrightarrow{d} \frac{2\pi f_{\nu\nu}(\omega)}{|1-e^{i\omega}|^2} + g(\omega, B_x) g(\omega, B_x)^*$$
$$= 2\pi f_{xx}(\omega) + g(\omega, B_x) g(\omega, B_x)^*.$$

We deduce that

$$\begin{aligned} n^{-1} \sum_{\lambda_s \in \mathscr{B}_A^c} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* &= \frac{1}{2J} \sum_{J_a \le |j| \le J} \frac{1}{m} \sum_{\lambda_s \in \mathscr{B}_j} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* + o_p(1) \\ &= \frac{1}{2J} \sum_{J_a \le |j| \le J} 2\pi \hat{f}_{xx.d}(\omega_j) + o_p(1) \\ &\xrightarrow{d} \int_{\mathscr{B}_A^c} \left[f_{xx}(\omega) + (2\pi)^{-1} g(\omega, B_x) g(\omega, B_x)^* \right] d\omega, \end{aligned}$$

which gives part (f).

Part (g): Observe that

$$\begin{aligned} n^{-1/2} \sum_{\lambda_s \in \mathscr{B}_A^c} w_{x.d}(\lambda_s) w_d(\lambda_s)^* &= n^{-1} \sum_{\lambda_s \in \mathscr{B}_A^c} w_{x.d}(\lambda_s) f_1(\lambda_s)^* + o_p(1) \\ &\overset{d}{\sim} -n^{-1} \sum_{\lambda_s \in \mathscr{B}_A^c} g(\lambda_s, B_x) f_1(\lambda_s)^* \\ &\overset{d}{\sim} -\left(\frac{1}{2\pi} \right) \frac{2\pi}{n} \sum_{\lambda_s \in \mathscr{B}_A^c} g(\lambda_s, B_x) f_1(\lambda_s)^* \\ &\xrightarrow{d} -\frac{1}{2\pi} \int_{\mathscr{B}_A^c} g(\omega, B_x) f_1(\omega)^* \, d\omega, \end{aligned}$$

as stated.

Part (h): From (73) we have

$$\sum_{\lambda_s \in \mathscr{B}_A^c} w_d(\lambda_s) w_d(\lambda_s)^* \sim \frac{1}{n} \sum_{\lambda_s \in \mathscr{B}_A^c} f_1(\lambda_s) f_1(\lambda_s)^* \to \frac{1}{2\pi} \int_{\mathscr{B}_A^c} f_1(\omega) f_1(\omega)^* \, d\omega,$$

as required.

Part (i): Using part (h), we get

$$\begin{aligned} n^{-1} \sum_{\lambda_s \in \mathscr{B}_A} w_d(\lambda_s) w_d(\lambda_s)^* &= n^{-1} \sum_{\lambda_s \in [-\pi, \pi]} w_d(\lambda_s) w_d(\lambda_s)^* + O_p\left(n^{-1}\right) \\ &= n^{-1} \sum_{t=1}^{n} d_t \sum_{\lambda_s \in [-\pi, \pi]} e^{i\lambda_s t} w_d(\lambda_s)^* + O_p\left(n^{-1}\right) \\ &= n^{-1} \sum_{t=1}^{n} d_t d_t' + O_p\left(n^{-1}\right) \\ &\to \int_0^1 u(r) u(r)' \, dr, \end{aligned}$$

giving the stated result.

Part (j): Using the representation of $w_{\tilde{x}}(\lambda_s)$ in Lemma B, the fact that $|s| > J \to \infty$ when $\lambda_s \in [-\pi, \pi] - \mathscr{B}_0$, and proceeding as in part (a), we find that

$$\begin{aligned} n^{-2} \sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* &= n^{-2}\left[\sum_{\lambda_s \in [-\pi,\pi]} - \sum_{\lambda_s \in [-\pi,\pi]-\mathscr{B}_0}\right] w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* + o_p(1) \\ &= n^{-2} \sum_{\lambda_s \in [-\pi,\pi]} w_{\tilde{x}}(\lambda_s) w_{\tilde{x}}(\lambda_s)^* \\ &\quad + O_p\left(\sum_{n \geq |s| \geq J} \frac{1}{s^2}\left[w_v(\lambda_s) w_v(\lambda_s)^* + \frac{\tilde{x}_n}{n^{1/2}} \frac{\tilde{x}_n'}{n^{1/2}}\right]\right) \\ &= \left[n^{-2} \sum_{t=1}^{n} \tilde{x}_t \tilde{x}_t'\right] + o_p(1) \xrightarrow{d} \int_0^1 B_x B_x'. \end{aligned}$$

Part (k): In the same way as parts (j) and (c) we find that

$$n^{-2} \sum_{\lambda_s \in \mathscr{B}_0} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* = n^{-2} \sum_{t=1}^{n} \tilde{x}_{.d,t} \tilde{x}'_{.d,t} + o_p(1) \xrightarrow{d} \int_0^1 B_{x.u} B_{x.u}'.$$

Part (l): From part (b) we have $n^{-1/2} w_d(0)' \to f_0' = \left(1, \frac{1}{2}\right) = \int_0^1 u'$ and, for $\lambda_s \neq 0$ with $(s/n) \to 0$, we have $w_d(\lambda_s)' = \left(0, (\sqrt{n}/2\pi s i)[1 + o(1)]\right)$ from (72). Thus, for $\omega = 0$, we find that

$$\begin{aligned} \frac{1}{n} \sum_{\lambda_s \in \mathscr{B}_0} w_d(\lambda_s) w_d(\lambda_s)^* &= \frac{1}{n} w_d(0) w_d(0)' + \frac{1}{n} \sum_{\lambda_s \in \mathscr{B}_0 - \{0\}} w_d(\lambda_s) w_d(\lambda_s)^* \\ &= \frac{1}{n} w_d(0) w_d(0)' + \begin{bmatrix} 0 & 0 \\ 0 & 2\left(\frac{1}{2\pi}\right)^2 \sum_{s=1}^{m} \frac{1}{s^2} \end{bmatrix} \\ &\to f_0 f_0' + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{12} \end{bmatrix} = \int_0^1 u u'. \end{aligned}$$

Part (m): As in part (b) we find that

$$\begin{aligned} n^{-3/2}\left(\sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^*\right) &= n^{-3/2} \sum_{\lambda_s \in [-\pi,\pi]} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* - n^{-3/2} \sum_{\lambda_s \notin \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* \\ &= n^{-3/2} \sum_{t=1}^{n} \tilde{x}_t d_t' - o_p(1) \\ &\xrightarrow{d} \int_0^1 B_x(r) u(r)'\, dr, \end{aligned}$$

since

$$\begin{aligned} n^{-3/2} \sum_{\lambda_s \notin \mathscr{B}_0} w_{\tilde{x}}(\lambda_s) w_d(\lambda_s)^* &= O_p\left(\frac{1}{nm} \sum_{\lambda_s \notin \mathscr{B}_0} w_{\tilde{x}}(\lambda_s)\right) \\ &= O_p\left(\frac{1}{nm} \sum_{\lambda_s \notin \mathscr{B}_0} \left[\frac{w_v(\lambda_s) - e^{i\lambda_s} \frac{\tilde{x}_n - \tilde{x}_0}{\sqrt{n}}}{1 - e^{i\lambda_s}}\right]\right) \\ &= o_p(1). \end{aligned}$$

Part (n): For $\omega \neq 0$, using (73) we obtain

$$\frac{n}{m} \sum_{\lambda_s \in \mathscr{B}_\omega} w_d(\lambda_s) w_d(\lambda_s)^* \sim \frac{1}{m} \sum_{\lambda_s \in \mathscr{B}_\omega} f_1(\lambda_s) f_1(\lambda_s)^* \to f_1(\omega) f_1(\omega)^*.$$

Part (o): For $\omega \neq 0$, using (73) we obtain

$$\begin{aligned}\frac{\sqrt{n}}{m}\sum_{\lambda_s\in\mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^* &= \frac{1}{m}\sum_{\lambda_s\in\mathcal{B}_\omega} w_{\tilde{x}}(\lambda_s)[f_1(\lambda_s)^*+o(1)]\\ &= \frac{1}{m}\sum_{\lambda_s\in\mathcal{B}_\omega}\left[\frac{w_v(\lambda_s)-e^{i\lambda_s}\frac{\tilde{x}_n-\tilde{x}_0}{\sqrt{n}}}{1-e^{i\lambda_s}}\right]f_1(\lambda_s)^*+o_p(1)\\ &\xrightarrow{d} \frac{e^{i\omega}}{1-e^{i\omega}}B(1)f_1(\omega)^*.\end{aligned}$$

Finally, joint weak convergence in (a)–(o) applies because the component elements jointly converge and one may apply the continuous mapping theorem in a routine fashion. In particular, Assumptions 1 and 2 ensure that

$$\left(n^{-1/2}\sum_{t=1}^{[nr]}\varsigma_t', w_v(\lambda)'\right)\xrightarrow{d}(B(r)',\xi_\lambda'),\quad \text{with}\quad \xi_\lambda = N_c(0,2\pi f_{vv}(\lambda)),$$

and where ξ_λ is independent of $B(r)$ for all $\lambda\neq 0$. The required quantities in (a)–(k) are functionals of these elements and smoothed periodogram ordinates (like $(1/m)\sum_{\lambda_s\in\mathcal{B}_j} w_v(\lambda_s)w_v(\lambda_s)^*$ for $\omega_j\to\omega$ as in the case of part (f)) that converge in probability to constants.

PROOF OF THEOREM 2: Note that

$$\hat{\beta}_A-\beta_A = -\{\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}\}^{-1}\{\widetilde{X}'Q_Z\Psi Q_Z[\Psi^c\widetilde{X}(\beta_A-\beta_{A^c})-\varepsilon]\}.$$

We consider the limiting behavior of $n^{-2}\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}$ and $n^{-1}\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X}$. Define $\tilde{x}_{.d,t}' = \tilde{x}_t' - d_t'(D'D)^{-1}(D'\widetilde{X})$. Since $\tilde{x}_t$ is an $I(1)$ process and satisfies an invariance principle when standardized by $n^{-1/2}$, we have

$$\text{(76)}\qquad n^{-1/2}\tilde{x}_{.d,[nr]}'\xrightarrow{d} B_x(r)'-u(r)'\left(\int_0^1 uu'\right)^{-1}\int_0^1 uB_x' = B_{x.u}(r)',\quad \text{say.}$$

Write the discrete Fourier transform of $\tilde{x}_{.d,t}'$ as $w_{x.d}(\lambda_s)$ and then from Lemma C(c) we have

$$\text{(77)}\qquad n^{-2}\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X} = \frac{1}{n^2}\left(\sum_{\lambda_s\in\mathcal{B}_A} w_{x.d}(\lambda_s)w_{x.d}(\lambda_s)^*\right)\xrightarrow{d}\int_0^1 B_{x.u}B_{x.u}'.$$

Since $\int_0^1 B_{x.u}B_{x.u}' > 0$ (see Phillips and Hansen (1990)), $n^{-2}\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}$ has a positive definite limit as $n\to\infty$.

Next, decompose $n^{-1}\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X}$ as follows:

$$\text{(78)}\qquad \begin{aligned}\frac{\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X}}{n} &= -\frac{\widetilde{X}'Q_ZW^*AWP_ZW^*A^cW\widetilde{X}}{n}-\frac{\widetilde{X}'Q_DW^*AWP_DW^*A^cW\widetilde{X}}{n}\\ &= \frac{\widetilde{X}'P_DW^*AWP_DW^*A^cW\widetilde{X}}{n}-\frac{\widetilde{X}'W^*AWP_DW^*A^cW\widetilde{X}}{n}\\ &= \text{term } A - \text{term } B.\end{aligned}$$

Take each of these terms in turn. Factor term A as follows and consider each factor separately. Write

$$\text{(79)}\qquad \frac{\widetilde{X}'P_DW^*AWP_DW^*A^cW\widetilde{X}}{n} = \left(\frac{\widetilde{X}'P_DW^*AWD}{n^{3/2}}\right)\left(\frac{D'D}{n}\right)^{-1}\left(\frac{D'W^*A^cW\widetilde{X}}{n^{1/2}}\right).$$

The first factor is

$$\text{(80)}\qquad \frac{n^{-1/2}\widetilde{X}'P_D W^* AWD}{n} = \frac{n^{-1/2}\widetilde{X}'D}{n}\left(\frac{D'D}{n}\right)^{-1}\frac{D'W^*AWD}{n}$$
$$\xrightarrow{d}\left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1}\left(\int_0^1 uu'\right) = \int_0^1 B_x u',$$

in view of (76) and Lemma C(i). The third factor is the conjugate transpose of

$$\text{(81)}\qquad n^{-1/2}\widetilde{X}'W^*A^cWD = n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*$$
$$\xrightarrow{d} -\frac{1}{2\pi}\left(B_x(1)\int_{\mathscr{B}_A^c}\frac{e^{i\lambda}f_1(\lambda)^*}{1-e^{i\lambda}}d\lambda\right),$$

from Lemma C(d). The limit of term A now follows by combining (81) and (80) and using joint weak convergence:

$$\text{(82)}\qquad \frac{\widetilde{X}'P_DW^*AWP_DW^*A^cW\widetilde{X}}{n} = \frac{n^{-1/2}\widetilde{X}'P_DW^*AWD}{n}\left(\frac{D'D}{n}\right)^{-1}(D'W^*A^cW(n^{-1/2}\widetilde{X}))$$
$$\xrightarrow{d}\left[\int_0^1 B_x u'\right]\left(\int_0^1 uu'\right)^{-1}\left[-\left(\frac{1}{2\pi}\right)\left(\int_{\mathscr{B}_A^c}\frac{e^{-i\lambda}f_1(\lambda)}{1-e^{-i\lambda}}d\lambda\right)B_x(1)'\right].$$

Next consider term B of (78). Using Lemma C, we obtain

$$\text{(83)}\qquad \frac{\widetilde{X}'W^*AWP_DW^*A^cW\widetilde{X}}{n}$$
$$= \frac{1}{n}\left(n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\right)\left(\frac{D'D}{n}\right)^{-1}\left(n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A^c} w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\right)$$
$$\xrightarrow{d}\left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1}\left(-\left(\frac{1}{2\pi}\right)\left(\int_{\mathscr{B}_A^c}\frac{e^{-i\lambda}f_1(\lambda)}{1-e^{-i\lambda}}d\lambda B_x(1)'\right)\right).$$

Combining (82) and (83) in (78) we find[3]

$$\text{(84)}\qquad n^{-1}\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X} = o_p(1).$$

As for the limit distribution of $\hat{\beta}_A$, we have

$$\text{(85)}\qquad n(\hat{\beta}_A-\beta_A) = \left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}}{n^2}\right\}^{-1}\left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\varepsilon}{n}\right\}$$
$$-\left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}}{n^2}\right\}^{-1}\left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X}}{n}(\beta_A-\beta_{A^c})\right\}.$$

From Lemma C(c) and (77) above, we have

$$\text{(86)}\qquad n^{-2}\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}\xrightarrow{d}\int_0^1 B_{x.u}B_{x.u}' = G,$$

[3] By changing the probability space (on which the random sequence $\tilde{x}_t$ is defined), we can ensure that both terms tend almost surely to the same random variable (see, for example, Theorem 4, page 47 of Shorack and Wellner (1986)). Then, in the original space the difference of the terms tends in distribution to zero giving the stated result.

and from (84)

$$\left(\frac{\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}}{n^2}\right)^{-1}\left(\frac{\widetilde{X}'Q_Z\Psi Q_Z\Psi^c\widetilde{X}}{n}\right)\xrightarrow{p}0. \tag{87}$$

Next,

$$\frac{\widetilde{X}'Q_Z\Psi Q_Z\varepsilon}{n}=n^{-1}\sum_{\lambda_s\in\mathscr{B}_A}w_{x.d}(\lambda_s)w_{\varepsilon.d}(\lambda_s)^*=\frac{1}{n}\sum_{\lambda_s\in\mathscr{B}_A}w_{x.d}(\lambda_s)w_\varepsilon(\lambda_s)^* -\left(\frac{1}{n^{\frac{3}{2}}}\sum_{\lambda_s\in\mathscr{B}_A}w_{x.d}(\lambda_s)w_d(\lambda_s)^*\right)\left(\frac{D'D}{n}\right)^{-1}\left(\frac{D'\varepsilon}{\sqrt{n}}\right). \tag{88}$$

Then, using (76) and proceeding as in the proof of Lemma C(a) above, we find

$$\begin{aligned} n^{-1}\sum_{\lambda_s\in\mathscr{B}_A}w_{x.d}(\lambda_s)w_\varepsilon(\lambda_s)^* &= n^{-1}\sum_{\lambda_s\in[-\pi,\pi]}w_{x.d}(\lambda_s)w_\varepsilon(\lambda_s)^*-n^{-1}\sum_{\lambda_s\in\mathscr{B}_{A^c}}w_{x.d}(\lambda_s)w_\varepsilon(\lambda_s)^*\\ &= n^{-1}\sum_{t=1}^{n}\tilde{x}_{.d,t}\varepsilon_t-o_p(1)\xrightarrow{d}\int_0^1 B_{x.u}dB_\varepsilon. \end{aligned} \tag{89}$$

Further, from Lemma C(b), (i) and Assumption 2, we obtain

$$\left(\frac{1}{n^{\frac{3}{2}}}\sum_{\lambda_s\in\mathscr{B}_A}w_{x.d}(\lambda_s)w_d(\lambda_s)^*\right)\left(\frac{D'D}{n}\right)^{-1}\left(\frac{D'\varepsilon}{\sqrt{n}}\right)\xrightarrow{d}\left(\int_0^1 B_{x.u}u'\right)\left(\int_0^1 uu'\right)\left(\int_0^1 u\,dB_\varepsilon\right)=0, \tag{90}$$

since $\int_0^1 B_{x.u}u'=0$. Combining (88), (89), and (90), we deduce that

$$\frac{\widetilde{X}'Q_Z\Psi Q_Z\varepsilon}{n}\xrightarrow{d}\int_0^1 B_{x.u}\,dB_\varepsilon. \tag{91}$$

The limit distribution (91) is a mixture normal distribution with mixing matrix variate $\int_0^1 B_{x.u}B'_{x.u}$.

It now follows from (86) and (91) that

$$\begin{aligned}\left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\widetilde{X}}{n^2}\right\}^{-1}\left\{\frac{\widetilde{X}'Q_Z\Psi Q_Z\varepsilon}{n}\right\}&\xrightarrow{d}\left(\int_0^1 B_{x.u}B'_{x.u}\right)^{-1}\left(\int_0^1 B_{x.u}dB_\varepsilon\right)\\ &\equiv MN\left(0,2\pi f_{\varepsilon\varepsilon}(0)\left(\int_0^1 B_{x.u}B'_{x.u}\right)^{-1}\right).\end{aligned} \tag{92}$$

Using (85), (87), and (92), we deduce that

$$n(\hat{\beta}_A-\beta_A)\xrightarrow{d}\left[\int_0^1 B_{x.u}B'_{x.u}\right]^{-1}\left[\int_0^1 B_{x.u}dB_\varepsilon\right],$$

which gives the stated result.

For the limit of $\hat{\beta}_{A^c}$, we need to examine the asymptotic behavior of the bias term in (29), which depends on the matrix quotient $(n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\widetilde{X})^{-1}(n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\Psi\widetilde{X})$. Take each of these factors in turn. First,

$$\frac{\widetilde{X}'Q_Z\Psi^cQ_Z\widetilde{X}}{n}=\frac{\widetilde{X}'Q_D\Psi^cQ_D\widetilde{X}}{n}=n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c}w_{x.d}(\lambda_s)w_{x.d}(\lambda_s)^*$$

where $w_{x.d}(\lambda_s)^*=w_{\tilde{x}}(\lambda_s)^*-w_d(\lambda_s)^*(n^{-1}D'D)^{-1}(n^{-1}D'\widetilde{X})$. From Lemma C(f) we have

$$n^{-1}\sum_{\lambda_s\in\mathscr{B}_A^c}w_{x.d}(\lambda_s)w_{x.d}(\lambda_s)^*\overset{d}{\to}\int_{\mathscr{B}_A^c}\left[f_{xx}(\omega)+(2\pi)^{-1}g(\omega,B_x)g(\omega,B_x)^*\right]d\omega,\tag{93}$$

which is a positive definite limit. Next,

$$\begin{aligned}&n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\Psi\widetilde{X}\\&\quad=-n^{-1}\widetilde{X}'Q_Z\Psi^cP_Z\Psi\widetilde{X}=-n^{-1}\widetilde{X}'Q_D\Psi^cP_D\Psi\widetilde{X}\\&\quad=-\left(n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A^c}w_{x.d}(\lambda_s)w_d(\lambda_s)^*\right)(n^{-1}D'D)^{-1}\left(n^{-3/2}\sum_{\lambda_s\in\mathscr{B}_A}w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\right).\end{aligned}$$

From Lemma C(g) and (b), we obtain

$$n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\Psi\widetilde{X}\overset{d}{\to}-\left(-\frac{1}{2\pi}\int_{\mathscr{B}_A^c}g(\omega,B_x)f_1(\omega)^*\,d\omega\right)\left(\int_0^1uu'\right)^{-1}\int_0^1uB_x'.\tag{94}$$

It follows from (93) and (94) and joint weak convergence that the asymptotic bias term for $\hat{\beta}_{A^c}$ involves

$$\begin{aligned}&(n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\widetilde{X})^{-1}(n^{-1}\widetilde{X}'Q_Z\Psi^cQ_Z\Psi\widetilde{X})\\&\quad\overset{d}{\to}\left[\int_{\mathscr{B}_A^c}\left[2\pi f_{xx}(\omega)+g(\omega,B_x)g(\omega,B_x)^*\right]d\omega\right]^{-1}\\&\qquad\times\left[\left(\int_{\mathscr{B}_A^c}g(\omega,B_x)f_1(\omega)^*\,d\omega\right)\left(\int_0^1uu'\right)^{-1}\int_0^1uB_x'\right],\end{aligned}$$

establishing the stated result.

Proof of Theorem 2′: *Part (a)*: From (64) and (11) we have

$$\hat{\beta}_\omega=\beta(1)+\{\widetilde{X}'Q_Z\Psi_\omega Q_Z\widetilde{X}\}^{-1}\{\widetilde{X}'Q_Z\Psi_\omega Q_Z\tilde{\varepsilon}\}.\tag{95}$$

Since $\tilde{\varepsilon}_t$ is a strictly stationary and ergodic sequence with mean zero and satisfies a central limit theorem, $(n^{-1}D'D)^{-1}(n^{-1}D'\tilde{\varepsilon})=O_p(n^{-1/2})$, and so

$$\tilde{\varepsilon}_t'-z_t'(n^{-1}Z'Z)^{-1}(n^{-1}Z'\tilde{\varepsilon})=\tilde{\varepsilon}_t'-z_t'\delta_n^{-1}(n^{-1}D'D)^{-1}(n^{-1}D'\tilde{\varepsilon})\overset{p}{\to}\tilde{\varepsilon}_t'.\tag{96}$$

Then, since $\tilde{\beta}(L)$ has a valid BN decomposition,

$$\begin{aligned}n^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\tilde{\varepsilon}&=n^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega}w_{\tilde{x}.d}(\lambda_s)w_{\tilde{\varepsilon}}(\lambda_s)^*+o_p(1)\\&=n^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega}w_{\tilde{x}.d}(\lambda_s)w_{\varepsilon}(\lambda_s)^*-n^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega}w_{\tilde{x}.d}(\lambda_s)w_v(\lambda_s)^*\tilde{\beta}(e^{-i\lambda_s})+o_p(1).\end{aligned}\tag{97}$$

When $\omega = 0$, we find as in (89) and Theorem 3.1 of Phillips (1991a) that

(98) $$n^{-1} \sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}.d}(\lambda_s) w_\varepsilon(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u}\, dB_\varepsilon,$$

and

(99) $$n^{-1} \sum_{\lambda_s \in \mathscr{B}_0} w_{\tilde{x}.d}(\lambda_s) w_v(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u}\, dB_x' + \Delta_x,$$

where $\Delta_x = \sum_{j=-\infty}^{\infty} E(v_0 v_j')$. We deduce that

(100) $$n^{-1} \widetilde{X}' Q_Z \Psi_\omega Q_Z \tilde{\varepsilon} \xrightarrow{d} \int_0^1 B_{x.u} dB_\varepsilon - \left\{ \int_0^1 B_{x.u} dB_x' + \Delta_x \right\} \tilde{\beta}(1).$$

Next, as in Lemma C(c) we find that

(101) $$n^{-2} \widetilde{X}' Q_Z \Psi_\omega Q_Z \widetilde{X} \xrightarrow{d} \int_0^1 B_{x.u} B_{x.u}'.$$

Combining (100) and (101), we obtain

$$n(\hat{\beta}_0 - \beta(1)) \xrightarrow{d} \left[\int_0^1 B_{x.u} B_{x.u}' \right]^{-1} \left[\int_0^1 B_{x.u}\, dB_\varepsilon - \left\{ \int_0^1 B_{x.u} dB_x' + \Delta_x \right\} \tilde{\beta}(1) \right].$$

Now consider the case where $\omega \neq 0$. First we have

$$m^{-1} \widetilde{X}' Q_Z \Psi_\omega Q_Z \widetilde{X} = m^{-1} \sum_{\lambda_s \in \mathscr{B}_\omega} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^*,$$

and in a similar fashion to (93) we find that

(102) $$m^{-1} \sum_{\lambda_s \in \mathscr{B}_\omega} w_{x.d}(\lambda_s) w_{x.d}(\lambda_s)^* \xrightarrow{d} 2\pi f_{xx}(\omega) + g(\omega, B_x) g(\omega, B_x)^* = \Xi_\omega, \quad \text{say.}$$

Next, from (15) we have

(103) $$w_{\tilde{y}}(\lambda_s) = \beta(1)' w_{\tilde{x}}(\lambda_s) - \tilde{\beta}(e^{i\lambda_s})' w_v(\lambda_s) + w_\varepsilon(\lambda_s) + O_p\left(\frac{1}{\sqrt{n}} \right).$$

Now

(104) $$w_{y.d}(\lambda_s)^* = w_{\tilde{y}}(\lambda_s)^* - w_d(\lambda_s)^* (n^{-1} D' D)^{-1} (n^{-1} D' \tilde{y}),$$

and, in view of the cointegrating relation (11), we have

(105) $$n^{-\frac{1}{2}} \tilde{y}_{[nr]} \to_d \beta(1)' B_x(r) = B_y(r),$$

and it follows as in part (f) of Lemma C that

(106) $$w_{y.d}(\lambda_s)^* \sim w_{\tilde{y}}(\lambda_s)^* - f_1(\lambda_s)^* \left(\int_0^1 uu' \right)^{-1} \left(\int_0^1 u B_x' \right) \beta(1) + o_p(1).$$

However, from the proof of Lemma C(f) we also have

(107) $$w_{x.d}(\lambda_s)^* = w_x(\lambda_s)^* - f_1(\lambda_s)^* \left(\int_0^1 uu' \right)^{-1} \left(\int_0^1 u B_x' \right) + o_p(1).$$

We deduce from (103), (106), and (107) that

$$w_{y.d}(\lambda_s)^* \sim w_{\tilde{x}}(\lambda_s)^*\beta(1) - f_1(\lambda_s)^*\left(\int_0^1 uu'\right)^{-1}\left(\int_0^1 uB_x'\right)\beta(1) + w_\varepsilon(\lambda_s)^* - w_v(\lambda_s)^*\tilde{\beta}(e^{-i\lambda_s}) + o_p(1) = w_{x.d}(\lambda_s)^*\beta(1) - w_v(\lambda_s)^*\tilde{\beta}(e^{-i\lambda_s}) + w_\varepsilon(\lambda_s)^* + o_p(1). \tag{108}$$

Then

$$m^{-1}\widetilde{X}'Q_Z\Psi_\omega Q_Z\tilde{y} = m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega} w_{x.d}(\lambda_s)w_{y.d}(\lambda_s)^* \sim m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega} w_{x.d}(\lambda_s)w_{x.d}(\lambda_s)^*\beta(1) - m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega} w_{x.d}(\lambda_s)w_v(\lambda_s)^*\tilde{\beta}(e^{-i\omega}) + o_p(1). \tag{109}$$

Using (51) we find that

$$m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega} w_{x.d}(\lambda_s)w_v(\lambda_s)^* = m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega} w_x(\lambda_s)w_v(\lambda_s)^* + o_p(1) = m^{-1}\sum_{\lambda_s\in\mathscr{B}_\omega}\frac{1}{1-e^{i\lambda_s}}w_v(\lambda_s)w_v(\lambda_s)^* + o_p(1) \to_p \frac{2\pi}{1-e^{i\omega}}f_{vv}(\omega). \tag{110}$$

It follows from (95), (102), (109), and (110) that

$$\hat{\beta}_\omega \to_d \beta(1) - \frac{2\pi}{1-e^{i\omega}}\Xi_\omega^{-1}f_{vv}(\omega)\tilde{\beta}(e^{-i\omega}) = \beta(1) + \frac{2\pi}{|1-e^{i\omega}|^2}\Xi_\omega^{-1}f_{vv}(\omega)\tilde{\beta}(e^{-i\omega})(e^{-i\omega}-1).$$

The true coefficient is

$$\beta_\omega = b(-\omega) = \beta(e^{-i\omega}) = \beta(1) + \tilde{\beta}(e^{-i\omega})(e^{-i\omega}-1),$$

so we have

$$\hat{\beta}_\omega \to_d \beta_\omega + \left[\frac{2\pi}{|1-e^{i\omega}|^2}\Xi_\omega^{-1}f_{vv}(\omega) - 1\right]\tilde{\beta}(e^{-i\omega})(e^{-i\omega}-1)$$

which gives the stated result since $f_{xx}(\omega) = |1-e^{i\omega}|^{-2}f_{vv}(\omega)$.

Part (b): The estimator $\tilde{\beta}_\omega$ is derived from the augmented spectral regression (32) and the relation (34). In (32), y_t and x_t are first detrended in the time domain and then dft's are taken of the detrended data and differences of the detrended data. Since $w_{\Delta x.d}(\lambda_s) \sim w_v(\lambda_s) + o_p(1)$, (108) can be written as

$$w_{y.d}(\lambda_s)^* = w_{x.d}(\lambda_s)^*\beta(1) - w_{\Delta x.d}(\lambda_s)^*\tilde{\beta}(e^{-i\lambda_s}) + w_\varepsilon(\lambda_s)^* + o_p(1). \tag{111}$$

First, take $\omega = 0$. In this case, the regressors $w_{x.d}(\lambda_s)$ and $w_{\Delta x.d}(\lambda_s)$ in (111) are asymptotically orthogonal after appropriate scaling because

$$m^{-1/2}n^{-1}\sum_{\lambda_s\in\mathscr{B}_0} w_{\tilde{x}.d}(\lambda_s)w_v(\lambda_s)^* = O_p(m^{-1/2}), \tag{112}$$

in view of (99). It then follows from (98), Lemma C(k), (112), and (111) that

$$n(\tilde{\beta}_0 - \beta(1)) \xrightarrow{d} \left[\int_0^1 B_{x.u}B_{x.u}'\right]^{-1}\left[\int_0^1 B_{x.u}dB_\varepsilon\right],$$

as required.

Now take the case $\omega \neq 0$. From part (f) of Lemma C and (52) we have

$$w_{x.z}(\lambda_s) = w_{x.d}(\lambda_s) = \frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} - \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}} - (n^{-3/2}\widetilde{X}'D)(n^{-1}D'D)^{-1}(n^{1/2}w_d(\lambda_s))$$
$$= \frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} + w_{\frac{t}{n}}(\lambda_s)[\tilde{x}_n - \tilde{x}_0] - (n^{-3/2}\widetilde{X}'D)(n^{-1}D'D)^{-1}(n^{1/2}w_d(\lambda_s))$$
$$= \frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} + A_n(n^{1/2}w_d(\lambda_s)).$$

Here, $n^{1/2}w_d(\lambda_s) = O_p(1)$ from (42) and

$$A_n' = \left[0, \frac{\tilde{x}_n - \tilde{x}_0}{\sqrt{n}}\right] - (n^{-3/2}\widetilde{X}'D)(n^{-1}D'D)^{-1} = O_p(1).$$

Then (111) is

$$w_{y.d}(\lambda_s)^* = \left[\frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} + A_n(n^{1/2}w_d(\lambda_s))\right]^*\beta(1) - w_v(\lambda_s)^*\tilde{\beta}(e^{-i\lambda_s}) + w_\varepsilon(\lambda_s)^* + o_p(1)$$
$$= w_v(\lambda_s)^*\left[\frac{\beta(1)}{1-e^{-i\lambda_s}} - \tilde{\beta}(e^{-i\lambda_s})\right] + w_d(\lambda_s)^*(n^{1/2}A_n) + w_\varepsilon(\lambda_s)^* + o_p(1). \tag{113}$$

The augmented narrow band regression (32) around frequency ω can therefore be written as

$$w_{y.z}(\lambda_s) = \tilde{a}_{1\omega}'\left[\frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} + A_n(n^{1/2}w_d(\lambda_s))\right] + \tilde{a}_{2\omega}'[w_v(\lambda_s) + o_p(1)] + \text{residual}, \tag{114}$$

which is asymptotically equivalent to the regression

$$w_{y.z}(\lambda_s)^* = w_v(\lambda_s)^*\tilde{b}_{1,-\omega} + (n^{1/2}w_d(\lambda_s))^*\tilde{b}_{2,-\omega} + \text{residual} \tag{115}$$

with

$$\tilde{b}_{1,-\omega} = \frac{\tilde{a}_{1,-\omega}}{1-e^{-i\omega}} + \tilde{a}_{2,-\omega}, \quad \tilde{b}_{2\omega} = \tilde{a}_{1,-\omega}A_n.$$

In view of (113), (115), and the asymptotic orthogonality of the regressors in (115) (i.e., $m^{-1}\sum_{\lambda_s \in} w_v(\lambda_s)(n^{1/2}w_d(\lambda_s))^* \to_p 0$), we find that

$$\tilde{b}_{1,-\omega} \to_p \frac{\beta(1)}{1-e^{-i\omega}} - \tilde{\beta}(e^{-i\omega}) = \frac{\beta_\omega}{1-e^{-i\omega}}$$

and

$$\sqrt{m}\left[\tilde{b}_{1,-\omega} - \frac{\beta_\omega}{1-e^{-i\omega}}\right] \to_d N_c\left(0, \frac{2\pi f_{\varepsilon\varepsilon}(\omega)}{2\pi f_{vv}(\omega)}\right).$$

We deduce that

$$\tilde{\beta}_\omega = \tilde{a}_{1,-\omega} + (1-e^{-i\omega})\tilde{a}_{2,-\omega} = \tilde{b}_{1,-\omega}(1-e^{-i\omega}) \to_p \beta_\omega$$

and

$$\sqrt{m}[\tilde{\beta}_\omega - \beta_\omega] \to_d N_c\left(0, \frac{f_{\varepsilon\varepsilon}(\omega)}{f_{xx}(\omega)}\right),$$

giving the stated result.

PROOF OF THEOREM 3: Note that

$$\begin{aligned} n(\hat{\beta}^f_A - \beta_A) &= (n^{-2}X'W^*AQ_{AWZ}AWX)^{-1}(n^{-1}X'W^*AQ_{AWZ}AW\varepsilon) \\ &= (n^{-2}X'Q_V\Psi Q_V X)^{-1}(n^{-1}X'Q_V\Psi Q_V\varepsilon) \\ &= (n^{-2}\widetilde{X}'Q_V\Psi Q_V\widetilde{X})^{-1}(n^{-1}\widetilde{X}'Q_V\Psi Q_V\varepsilon) \\ &= (n^{-2}\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\widetilde{X})^{-1}(n^{-1}\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\varepsilon). \end{aligned}$$

Next,

$$\begin{aligned} \widetilde{X}'Q_{\Psi Z}\Psi &= \widetilde{X}'\Psi - \widetilde{X}'P_{\Psi Z}\Psi = \widetilde{X}'\Psi - \widetilde{X}'\Psi Z(Z'\Psi Z)^{-1}Z'\Psi \\ &= \widetilde{X}'\Psi - \widetilde{X}'\Psi D(D'\Psi D)^{-1}D'\Psi, \end{aligned}$$

and so

$$\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\widetilde{X} = (\widetilde{X}'Q_{\Psi Z}\Psi)(\Psi Q_{\Psi Z}\widetilde{X}) = \widetilde{X}'\Psi\widetilde{X}' - \widetilde{X}'\Psi D(D'\Psi D)^{-1}D'\Psi\widetilde{X}',$$

which is

$$\begin{aligned} &\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* \\ &\quad - \Bigg(\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\Bigg)\Bigg(\sum_{\lambda_s\in\mathcal{B}_A} w_d(\lambda_s)w_d(\lambda_s)^*\Bigg)^{-1}\Bigg(\sum_{\lambda_s\in\mathcal{B}_A} w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\Bigg). \end{aligned}$$

Now, as in Lemma C(a), (b) and (i), we have

$$\begin{aligned} &n^{-2}\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_xB_x', \\ &n^{-3/2}\Bigg(\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\Bigg) \xrightarrow{d}_d \Bigg(\int_0^1 B_xu'\Bigg), \\ &n^{-1}\sum_{\lambda_s\in\mathcal{B}_A} w_d(\lambda_s)w_d(\lambda_s)^* \to \int_0^1 uu'. \end{aligned}$$

Thus,

$$n^{-2}\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\widetilde{X} \xrightarrow{d} \int_0^1 B_xB_x' - \Bigg(\int_0^1 B_xu'\Bigg)\Bigg(\int_0^1 uu'\Bigg)^{-1}\Bigg(\int_0^1 f_0B_x'\Bigg) = \int_0^1 B_{x.u}B_{x.u}'. \tag{116}$$

Next, consider the limit behavior of

$$\begin{aligned} n^{-1}\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\varepsilon &= n^{-1}[\widetilde{X}'\Psi\varepsilon - \widetilde{X}'\Psi D(D'\Psi D)^{-1}D'\Psi\varepsilon] \\ &= n^{-1}\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_\varepsilon(\lambda_s)^* - \Bigg(n^{-3/2}\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\Bigg) \\ &\quad\times\Bigg(n^{-1}\sum_{\lambda_s\in\mathcal{B}_A} w_d(\lambda_s)w_d(\lambda_s)^*\Bigg)^{-1}\Bigg(n^{-1/2}\sum_{\lambda_s\in\mathcal{B}_A} w_d(\lambda_s)w_\varepsilon(\lambda_s)^*\Bigg). \end{aligned}$$

As in the proof of (91),

$$n^{-1}\sum_{\lambda_s\in\mathcal{B}_A} w_{\tilde{x}}(\lambda_s)w_\varepsilon(\lambda_s)^* \xrightarrow{d} \int_0^1 B_x\,dB_\varepsilon,$$

and

$$n^{-1/2}\sum_{\lambda_s\in\mathscr{B}_A} w_d(\lambda_s)w_\varepsilon(\lambda_s)^* \xrightarrow{d} \int_0^1 u\,dB_\varepsilon.$$

Thus,

$$n^{-1}\widetilde{X}'Q_{\Psi Z}\Psi Q_{\Psi Z}\varepsilon \xrightarrow{d} \int_0^1 B_x\,dB_\varepsilon - \left(\int_0^1 B_x u'\right)\left(\int_0^1 uu'\right)^{-1}\int_0^1 u dB_\varepsilon = \int_0^1 B_{x.u}\,dB_\varepsilon.$$

It follows that

$$n(\hat{\beta}_A^f - \beta_A) \xrightarrow{d} \left(\int_0^1 B_{x.u}B_{x.u}'\right)^{-1}\left(\int_0^1 B_{x.u}\,dB_\varepsilon\right) \equiv MN\left(0, \left(\int_0^1 B_{x.u}B_{x.u}'\right)^{-1} 2\pi f_{\varepsilon\varepsilon}(0)\right),$$

giving the stated result for the band $\mathscr{B}_A$.

For the band $\mathscr{B}_A^c$, we have

$$\sqrt{n}(\hat{\beta}_{A^c}^f - \beta_{A^c}) = (n^{-1}X'W^*A^cQ_{A^cWZ}A^cWX)^{-1}(n^{-1/2}X'W^*A^cQ_{A^cWZ}A^cW\varepsilon)$$
$$= (n^{-1}\widetilde{X}'Q_{\Psi^cZ}\Psi^cQ_{\Psi^cZ}\widetilde{X})^{-1}(n^{-1/2}\widetilde{X}'Q_{\Psi^cZ}\Psi^cQ_{\Psi^cZ}\varepsilon). \tag{117}$$

As above,

$$\widetilde{X}'Q_{\Psi^cZ}\Psi^cQ_{\Psi^cZ}\widetilde{X} = (\widetilde{X}'Q_{\Psi^cZ}\Psi^c)(\Psi^cQ_{\Psi^cZ}\widetilde{X})$$
$$= \widetilde{X}'\Psi^c\widetilde{X}' - \widetilde{X}'\Psi^cD(D'\Psi^cD)^-D'\Psi^c\widetilde{X}'$$
$$= \sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)w_{\tilde{x}}(\lambda_s)^* - \left(\sum_{\lambda_s\in\mathscr{B}_A^c} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\right)$$
$$\times\left(\sum_{\lambda_s\in\mathscr{B}_A^c} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-1}\left(\sum_{\lambda_s\in\mathscr{B}_A^c} w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\right).$$

From the above expression and Lemma C(d), (e), and (h) we deduce that

$$n^{-1}\widetilde{X}'Q_{\Psi^cZ}\Psi^cQ_{\Psi^cZ}\widetilde{X}$$
$$\xrightarrow{d} \int_{\mathscr{B}_A^c}\left[f_{xx}(\omega) + \frac{1}{2\pi}\frac{1}{|1-e^{i\omega}|^2}B_x(1)B_x(1)'\right]d\omega - \left[\frac{1}{2\pi}B_x(1)\int_{\mathscr{B}_A^c}\frac{e^{i\omega}f_1(\omega)^*}{1-e^{i\omega}}\,d\omega\right]$$
$$\times\left[\frac{1}{2\pi}\int_{\mathscr{B}_A^c} f_1(\omega)f_1(\omega)^*\,d\omega\right]^-\left[\frac{1}{2\pi}\int_{\mathscr{B}_A^c} f_1(\omega)\frac{e^{-i\omega}}{1-e^{-i\omega}}\,d\omega B_x(1)'\right]$$
$$= \int_{\mathscr{B}_A^c}\left[f_{xx}(\omega) + \frac{1}{2\pi}\frac{1}{|1-e^{i\omega}|^2}B_x(1)B_x(1)'\right]d\omega - \frac{1}{2\pi}B_x(1)B_x(1)'\left[\int_{\mathscr{B}_A^c}\frac{e^{i\omega}f_1(\omega)^*}{1-e^{i\omega}}\,d\omega\right]$$
$$\times\left[\int_{\mathscr{B}_A^c} f_1(\omega)f_1(\omega)^*\,d\omega\right]^-\left[\int_{\mathscr{B}_A^c} f_1(\omega)\frac{e^{-i\omega}}{1-e^{-i\omega}}\,d\omega\right]. \tag{118}$$

Next, observe that

$$\left[\int_{\mathscr{B}_A^c}\frac{e^{i\omega}f_1(\omega)^*}{1-e^{i\omega}}\,d\omega\right]\left[\int_{\mathscr{B}_A^c} f_1(\omega)f_1(\omega)^*\,d\omega\right]^- f_1(\omega)$$

is the $L_2(\mathscr{B}_A^c)$ projection of the function $e^{i\omega}(1-e^{i\omega})^{-1}$ onto the space spanned by $f_1(\omega)$. When the deterministic variable z_t includes a linear time trend, we know from (39) that the vector $f_1(\omega)$ includes the function $e^{i\omega}(1-e^{i\omega})^{-1}$ as one of its components. Hence, in this case we have

$$\left[\int_{\mathscr{B}_A^c} e^{i\omega}(1-e^{i\omega})^{-1}f_1(\omega)^*\,d\omega\right]\left[\int_{\mathscr{B}_A^c} f_1(\omega)f_1(\omega)^*\,d\omega\right]^- f_1(\omega) = e^{i\omega}(1-e^{i\omega})^{-1} \tag{119}$$

for $\omega \in \mathscr{B}_A^c$. It follows that (118) is simply $\int_{\mathscr{B}_A^c} f_{xx}(\omega)\,d\omega$.

Proceeding with the proof, the second factor of (117) decomposes as

$$
\begin{aligned}
& n^{-1/2}\widetilde{X}' Q_{\Psi^c Z}\Psi^c Q_{\Psi^c Z}\varepsilon \\
&\quad = n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A} w_{\tilde{x}}(\lambda_s)w_\varepsilon(\lambda_s)^* \\
&\qquad -\left(n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A} w_{\tilde{x}}(\lambda_s)w_d(\lambda_s)^*\right)\left(\sum_{\lambda_s\in\mathscr{B}^c_A} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-}\left(\sum_{\lambda_s\in\mathscr{B}^c_A} w_d(\lambda_s)w_\varepsilon(\lambda_s)^*\right).
\end{aligned}
$$

Using (119) and the independence of $\tilde{x}_t$ and ε_t we find

$$
\begin{aligned}
n^{-1/2}\widetilde{X}' Q_{\Psi^c Z}\Psi^c Q_{\Psi^c Z}\varepsilon &\overset{d}{\sim} n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A} w_{\tilde{x}}(\lambda_s)w_\varepsilon(\lambda_s)^* \\
&\quad -\left(\frac{1}{2\pi}B_x(1)\int_{\mathscr{B}^c_A}\frac{e^{i\omega}f_1(\omega)^*}{1-e^{i\omega}}\,d\omega\right)\left(\frac{1}{2\pi}\int_{\mathscr{B}^c_A} f_1(\omega)f_1(\omega)^*\,d\omega\right)^{-} \\
&\quad \times n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A} f_1(\lambda_s)w_\varepsilon(\lambda_s)^* \\
&\overset{d}{\sim} n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A}\left[w_{\tilde{x}}(\lambda_s)-\left(\frac{1}{2\pi}B_x(1)\int_{\mathscr{B}^c_A}\frac{e^{i\omega}f_1(\omega)^*}{1-e^{i\omega}}\,d\omega\right)\right. \\
&\qquad \left.\times\left(\frac{1}{2\pi}\int_{\mathscr{B}^c_A} f_1(\omega)f_1(\omega)^*\,d\omega\right)^{-} f_1(\lambda_s)\right]w_\varepsilon(\lambda_s)^* \\
&= n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A}\left[w_{\tilde{x}}(\lambda_s)-B_x(1)\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\right]w_\varepsilon(\lambda_s)^* \\
&= n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A}\left[\frac{1}{1-e^{i\lambda_s}}\,w_\nu(\lambda_s)-\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{[\tilde{x}_n-\tilde{x}_0]}{n^{1/2}}-B_x(1)\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\right] \\
&\quad \times w_\varepsilon(\lambda_s)^* \\
&= n^{-1/2}\sum_{\lambda_s\in\mathscr{B}^c_A}\left[\frac{1}{1-e^{i\lambda_s}}\,w_\nu(\lambda_s)\right]w_\varepsilon(\lambda_s)^*+o_p(1) \\
&\overset{d}{\sim} N\left(0,\frac{2\pi}{n}\sum_{\lambda_s\in\mathscr{B}^c_A}\left[\frac{1}{|1-e^{i\lambda_s}|^2}\,w_\nu(\lambda_s)w_\nu(\lambda_s)^*\right]f_{\varepsilon\varepsilon}(\lambda_s)\right) \\
&\overset{d}{\sim} N\left(0,2\pi\int_{\mathscr{B}^c_A} f_{xx}(\omega)f_{\varepsilon\varepsilon}(\omega)\,d\omega\right).
\end{aligned}
$$

We deduce that

$$
\sqrt{n}(\hat{\beta}^f_{A^c}-\beta_{A^c})\overset{d}{\to} N\left(0,\left[\int_{\mathscr{B}^c_A} f_{xx}(\omega)\,d\omega\right]^{-1}\left[2\pi\int_{\mathscr{B}^c_A} f_{xx}(\omega)f_{\varepsilon\varepsilon}(\omega)\,d\omega\right]\left[\int_{\mathscr{B}^c_A} f_{xx}(\omega)\,d\omega\right]^{-1}\right),
$$

giving the stated result.

PROOF OF THEOREM 3′: *Part (a)*: First note that

$$
w_x(\lambda_s)=\Pi_2' w_z(\lambda_s)+w_{\tilde{x}}(\lambda_s)=\Pi_2'\delta_n w_d(\lambda_s)+w_{\tilde{x}}(\lambda_s),
$$

with a similar expression for $w_y(\lambda_s)$. The regression (45) is equivalent to the following regression with frequency domain detrended data:

$$
w^f_{\tilde{y}.d}(\lambda_s)=\tilde{c}_{1\omega}' w^f_{\tilde{x}.d}(\lambda_s)+\tilde{c}_{2\omega}' w^f_{\Delta\tilde{x}.d}(\lambda_s)+\text{residual}, \tag{120}
$$

where

$$w^f_{x.d}(\lambda_s)^* = w_x(\lambda_s)^* - w_d(\lambda_s)^*\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-1}\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_x(\lambda_s)^*\right)$$

$$= w_{\tilde{x}}(\lambda_s)^* - w_d(\lambda_s)^*\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-1}\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\right)$$

$$= w^f_{\tilde{x}.d}(\lambda_s)^*, \quad \text{say,}$$

with a similar expression for $w^f_{y.d}(\lambda_s) = w^f_{\tilde{y}.d}(\lambda_s)$, and

$$w^f_{\Delta\tilde{x}.d}(\lambda_s)^* = w_{\Delta\tilde{x}}(\lambda_s)^* - w_d(\lambda_s)^*\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-1}\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_{\Delta\tilde{x}}(\lambda_s)^*\right)$$

$$= w_v(\lambda_s)^* + o_p(1).$$

For $\omega = 0$, as in the proof of Theorem 2′, the regressors in (120) are asymptotically orthogonal because

$$n^{-3/2}\sum_{\lambda_s\in\mathscr{B}_0} w^f_{\tilde{x}.d}(\lambda_s)w_v(\lambda_s)^* = O_p(n^{-1/2}).$$

Next, using Lemma C(j)–(m), we find as in (116) above that

$$n^{-2}\sum_{\lambda_s\in\mathscr{B}_0} w^f_{\tilde{x}.d}(\lambda_s)w^f_{\tilde{x}.d}(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u}B'_{x.u},$$

and as in (89) we get

$$n^{-1}\sum_{\lambda_s\in\mathscr{B}_0} w^f_{\tilde{x}.d}(\lambda_s)w_\varepsilon(\lambda_s)^* \xrightarrow{d} \int_0^1 B_{x.u}dB_\varepsilon.$$

It follows using (15) and these two results that

$$n(\tilde{\beta}^f_\omega - \beta_0) \xrightarrow{d} \left[\int_0^1 B_{x.u}B'_{x.u}\right]^{-1}\left[\int_0^1 B_{x.u}dB_\varepsilon\right].$$

Next consider the $\omega \neq 0$ case. First we simplify the regression equation (120):

(121) $$w^f_{\tilde{y}.d}(\lambda_s) = \tilde{c}'_{1\omega}w^f_{\tilde{x}.d}(\lambda_s) + \tilde{c}'_{2\omega}w^f_{\Delta\tilde{x}.d}(\lambda_s) + \text{residual}$$
$$= \tilde{c}'_{1\omega}w^f_{\tilde{x}.d}(\lambda_s) + \tilde{c}'_{2\omega}[w_v(\lambda_s) + o_p(1)] + \text{residual}.$$

Using (73) and Lemma C(n)–(o) we obtain

(122) $$w^f_{\tilde{x}.d}(\lambda_s)^* = w_{\tilde{x}}(\lambda_s)^* - w_d(\lambda_s)^*\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_d(\lambda_s)^*\right)^{-1}\left(\sum_{\lambda_s\in\mathscr{B}_\omega} w_d(\lambda_s)w_{\tilde{x}}(\lambda_s)^*\right)$$
$$= \frac{w_v(\lambda_s)^*}{1-e^{-i\lambda_s}} - \frac{e^{-i\lambda_s}}{1-e^{-i\lambda_s}}\frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}}$$
$$- \frac{1}{\sqrt{n}}f_1(\omega)^*\left[\frac{m}{n}f_1(\omega)f_1(\omega)^*\right]^{-}\left[\frac{m}{\sqrt{n}}f_1(\omega)\frac{e^{-i\omega}}{1-e^{-i\omega}}B(1)\right]$$

$$= \frac{w_v(\lambda_s)^*}{1-e^{-i\lambda_s}} - \frac{e^{-i\lambda_s}}{1-e^{-i\lambda_s}} \frac{[\tilde{x}_n - \tilde{x}_0]}{n^{1/2}} - \frac{e^{-i\omega}}{1-e^{-i\omega}} B(1)$$
$$= \frac{w_v(\lambda_s)^*}{1-e^{-i\lambda_s}} + o_p(1).$$

So the regression (121) is asymptotically equivalent to

$$w^f_{\tilde{y}.d}(\lambda_s)^* = \left[\frac{w_v(\lambda_s)}{1-e^{i\lambda_s}} + o_p(1)\right]^* \tilde{c}_{1\omega} + [w_v(\lambda_s) + o_p(1)]^* \tilde{c}_{2\omega} + \text{residual}. \tag{123}$$

From (122) and (15) we deduce, as in Theorem 2′, that

$$\tilde{b}^f_{1,-\omega} := \frac{\tilde{c}_{1,-\omega}}{1-e^{-i\omega}} + \tilde{c}_{2,-\omega p} \to_p \frac{\beta(1)}{1-e^{-i\omega}} - \tilde{\beta}(e^{-i\omega}) = \frac{\beta_\omega}{1-e^{-i\omega}}$$

and

$$\sqrt{m}\left[\tilde{b}_{1,-\omega} - \frac{\beta_\omega}{1-e^{-i\omega}}\right] \to_d N_c\left(0, \frac{f_{\varepsilon\varepsilon}(\omega)}{f_{vv}(\omega)}\right).$$

It follows that

$$\tilde{\beta}^f_\omega = \tilde{c}_{1,-\omega} + (1-e^{-i\omega})\tilde{c}_{2,-\omega} = \tilde{b}_{1,-\omega}(1-e^{-i\omega}) \to_p \beta_\omega$$

and

$$\sqrt{m}\,[\tilde{\beta}^f_\omega - \beta_\omega] \to_d N_c\left(0, \frac{f_{\varepsilon\varepsilon}(\omega)}{f_{xx}(\omega)}\right).$$

REFERENCES

CHOI, I., AND P. C. B. PHILLIPS (1993): “Testing for a Unit Root by Frequency Domain Regression,” *Journal of Econometrics*, 59, 263–286.

CORBAE, D., S. OULIARIS, AND P. C. B. PHILLIPS (1997): “Band Spectral Regression with Trending Data,” Cowles Foundation Discussion Paper, Yale University, 1163.

ENGLE, R. (1974): “Band Spectrum Regression,” *International Economic Review*, 15, 1–11.

FRISCH, R., AND F. V. WAUGH (1993): “Partial Time Series Regression as Compared with Individual Trends,” *Econometrica*, 1, 211–223.

GEWEKE, J. (1982): “Measurement of Linear Dependence and Feedback Between Multiple Time Series,” *Journal of the American Statistical Association*, 77, 304–313.

GRADSHTEYN, I. S., AND I. M. RYZHIK (1965): *Tables of Integrals, Series and Products*. New York: Academic Press.

GRANGER, C. W. J., AND J.-L LIN (1995): “Causality in the Long Run,” *Econometric Theory*, 11, 530–536.

GRENANDER, U., AND M. ROSENBLATT (1957): *Statistical Analysis of Stationary Time Series*. New York: Wiley.

HANNAN, E. J. (1963a): “Regression for Time Series,” in *Time Series Analysis*, ed. by M. Rosenblatt. New York: Wiley.

——— (1963b): “Regression for Time Series with Errors of Measurement,” *Biometrika*, 50, 293–302.

——— (1970): *Multiple Time Series*. New York: Wiley.

——— (1973): “Central Limit Theorems for Time Series Regression,” *Zeitschrift für Wahrscheinlichkeitsheorie und verwandte Gebiete*, 26, 157–170.

HANNAN, E. J., AND P. J. THOMSON (1971): “Spectral Inference over Narrow Bands,” *Journal of Applied Probability*, 8, 157–169.

HANNAN, E. J., AND P. M. ROBINSON (1973): “Lagged Regression with Unknown Lags,” *Journal of the Royal Statistical Society*, Series B, 35, 252–267.

HOSOYA, Y. (1991): "The Decomposition and Measurement of the Interdependence between Second-order Stationary Processes," *Probability Theory and Related Fields*, 88, 429–444.

KIM, C., AND P. C. B. PHILLIPS (1999): "Fully Modified Estimation of Fractional Cointegration Models," Yale University, Mimeo.

MARINUCCI, D. (1998): "Band Spectrum Regression for Cointegrated Time Series with Long Memory Innovations," LSE Econometrics Discussion Paper #353.

PHILLIPS, P. C. B. (1991a): "Spectral Regression for Co-integrated Time Series," in *Nonparametric and Semiparametric Methods in Economics and Statistics*, ed. by W. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press.

——— (1991b): "Optimal Inference in Cointegrated Systems," *Econometrica*, 59, 283–306.

——— (1999): "Discrete Fourier Transforms of Fractional Processes," Cowles Foundation Discussion Paper #1243, Yale University.

——— (2000): "Unit Root Log Periodogram Regression," Cowles Foundation Discussion Paper #1244, Yale University.

PHILLIPS, P. C. B., AND B. HANSEN (1990): "Statistical Inference in Instrumental Variable Regression with *I(1)* processes," *Review of Economic Studies*, 57, 99–125.

PHILLIPS, P. C. B., AND C. C. LEE (1996): "Efficiency Gains from Quasi-differencing under Nonstationarity," in *Athens Conference on Applied Probability and Time Series: Essays in Memory of E. J. Hannan*, ed. by P. M. Robinson and M. Rosenblatt. New York: Springer-Verlag.

PHILLIPS, P. C. B., AND M. LORETAN (1991): "Estimating Long-run Economic Equilibria," *Review of Economic Studies*, 59, 407–436.

PHILLIPS, P. C. B., AND V. SOLO (1992): "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971–1001.

ROBINSON, P. M. (1991): "Automatic Frequency Domain Inference," *Econometrica*, 59, 1329–1363.

ROBINSON P. M., AND D. MARINUCCI (1998): "Semiparametric Frequency Domain Analysis of Fractional Cointegration," LSE Econometrics Discussion Paper #348.

SAIKKONEN, P. (1991): "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, 1–21.

SHORACK, G. R., AND J. A. WELLNER (1986): *Empirical Processes with Applications to Statistics*. New York: Wiley.

STOCK, J. H., AND M. W. WATSON (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.

XIAO, ZHIJIE, AND PETER C. B. PHILLIPS (1998): "Higher Order Approximations for Frequency Domain Time Series Regression," *Journal of Econometrics*, 86, 297–336.

——— (1999): "Higher Order Approximations for Wald Statistics in Time Series Regressions with Integrated Processes," Yale University, Mimeo.