COINTEGRATION AND SAMPLING FREQUENCY

Marcus J. CHAMBERS*

University of Essex

June 2001

Abstract

This paper analyses the effects of sampling frequency on the properties of spectral regression estimators of cointegrating parameters. Large sample asymptotic properties are derived under three scenarios concerning the span of data and sampling frequency, each scenario depending on whether span or frequency (or both) tends to infinity. The limiting distributions are shown to be different in each case. Furthermore, the asymptotic efficiency of the estimators obtained with a fixed sampling frequency is compared with that obtained with a continuous record of data, and it is shown that the only inefficiencies arise with respect to stock variables. Some simulation results and an empirical illustration are also provided.

Keywords: Sampling frequency; cointegration; spectral regression.

J.E.L. Number: C32

* I would like to thank Rex Bergstrom and Joanne McGarry for comments on this paper. The financial support provided by the ESRC under grant number R000222961 is gratefully acknowledged.

Address for Correspondence: Professor Marcus J. Chambers, Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England.
Tel: +44 1206 872756; fax: +44 1206 872724; e-mail: mchamb@essex.ac.uk.

1. INTRODUCTION

In some areas of economics, most notably finance, data frequency is increasingly becoming an element of the econometrician's decision set. In addition to making choices concerning, *inter alia*, functional form and the appropriate methods of estimation and inference to use, the frequency of data with which to conduct the analysis must also be chosen. In macroeconomics the choice is typically between annual and quarterly frequencies, although an increasing number of macroeconomic variables are now available on a monthly basis. In finance the choice is even greater, with near-continuous sampling being possible in some applications. The use of different data frequencies presumably has some effect on the properties of the estimation and inference procedures employed. A question of some interest is, therefore: precisely what are those effects?

Whilst it may be difficult (if not impossible) to answer this question in the generality in which it is posed, this paper attempts to address some more specific questions, the answers to which have a bearing on econometric research in certain applications. The focus is a particular class of time series models that are in widespread use, namely models of cointegration. The class of estimators of the cointegrating parameters that is considered is the class of spectral regression estimators. The large sample asymptotic properties of the estimators can be examined in a number of ways, depending on the way in which sample size tends to infinity. For a fixed sampling frequency, the number of observations tends to infinity if the span covered by the data tends to infinity. Conversely, for a fixed span of data, the number of observations grows if sampling becomes more frequent. Obviously, a combination of an increasing span and an increase in frequency also leads to an increasing sample size. All three modes of asymptotics are considered, and it is shown that the limiting properties of the estimators differ in each case. Hence the following more precise question is addressed: what are the effects of increasing data span and/or increasing sampling frequency on the asymptotic properties of spectral regression estimators of cointegrating parameters?

In view of cointegration being a feature of the long-run relationship between integrated time series, an answer to the previous question enables further issues to be explored. For example: is it possible to consistently estimate cointegrating parameters when the data span is fixed? Consistency in this case refers to the asymptotic analysis in which sampling frequency increases. Related research on testing for unit roots, for example the simulation results of Shiller and Perron (1985), suggests that increasing span is the important factor for test consistency, a finding confirmed by the theoretical results of Perron (1991). It is interesting to assess whether the same is true for the consistent estimation of cointegrating parameters. The characterisation of the limiting distributions also enables the investigation, in cases where span tends to infinity, of the question: is there an efficiency loss associated with sampling at a fixed frequency compared to the limiting case of continuous sampling? It turns out that the efficiency loss can be quantified (at least in theory) and that there is only a loss in efficiency where stock variables are concerned. Expressed in a slightly different way, the implication is that the estimators obtained with flow data at a fixed sampling frequency are as efficient as when based on a continuous sample. This remarkable result generalises related work in Chambers (2000).

The effects of sampling frequency on estimators and test statistics have been analysed in a variety of settings. Sargan (1974) derived the order of magnitude (in terms of the sampling frequency) of the asymptotic bias of various estimators of the parameters of stationary continuous time systems derived from approximate discrete models. Most recent research has been univariate in nature but has relaxed the stationarity requirement. Phillips (1987a,b) derived continuous record asymptotics for the ordinary least squares (OLS) estimator in a first-order autoregression with a unit root. Perron (1991) considered the consistency of tests of the random walk hypothesis and of randomness and, as mentioned above, shows that it is the increasing span of the data, rather than the frequency, that is important.

The recent research described in the preceding paragraph has been based on a univariate model with Brownian motion characterising the random disturbance process in continuous time. Such a model has the advantage of generating a discrete time process that satisfies a first-order autoregressive model with an independently and normally distributed disturbance term. Whilst this greatly facilitates the analysis and allows the precise effects of sampling frequency to be pinpointed, such an assumption, even extended to a multivariate Brownian motion process, would not be appropriate in the setting of this paper. Because, for reasons that will become apparent, the analysis here is based on the triangular error correction model (ECM) of Phillips (1991a,b), it is important to allow the disturbance process to characterise fully the dynamics of the cointegrated system. A process such as Brownian motion with independent increments is thus inappropriate for this task, and so a much more general forcing process is allowed that imposes much weaker conditions on the serial correlation and heterogeneity properties of the random disturbance. Whilst allowing for greater generality such an assumption requires a number of new results to be derived, in particular the invariance principles that describe the limiting properties of suitably scaled partial sum processes and on which many subsequent results depend. Hence the results in this paper represent a significant advance in the sophistication of model that can be analysed in this branch of the literature in three important dimensions. Simultaneously, the model is multivariate; the random variables are nonstationary; and the random forcing process is only required to satisfy much weaker conditions than the increment of Brownian motion.

The layout of the paper is as follows. Section 2 defines the underlying continuous time model and derives some results concerning its discrete time representation for any arbitrary sampling frequency, while Section 3 provides some preliminary notation and discussion of the spectral estimators under examination. The important asymptotic results appear in Section 4 which also contains the results related to asymptotic efficiency comparisons and discusses large sample inference. The results of a simulation experiment are reported in Section 5, along with an investigation of the performance of the spectral estimators in an empirical setting when sampling frequency is allowed to vary. Section 6 concludes the paper. The proofs of all lemmas and theorems, as well as additional details concerning the simulations, are contained in four appendices. This is done so as to aid the flow of the development of the results of interest in the main body of the paper.

Finally, the following notation shall be used throughout the paper. L denotes the lag operator such that, for a variable x_t, $L^j x_t = x_{t-j}$ for some integer j. D denotes the mean square differential operator such that, for a variable $x(t)$ defined in continuous time, $Dx(t)$ is defined by $\lim_{h\downarrow 0} E\{h^{-1}[x(t+h)-x(t)]-Dx(t)\}^2 = 0$. For a random $m\times 1$ vector process $x(t)$, $\|x(t)\|_\delta = [\sum_{j=1}^m E|x_j(t)|^\delta]^{1/\delta}$, while $\|x(t)\|$ shall denote the Euclidean norm $\|x(t)\| = [\sum_{j=1}^m x_j(t)^2]^{1/2}$. For an $m\times m$ matrix A, this norm is defined by $\|A\| = [\sum_{i=1}^m \sum_{j=1}^m a_{ij}^2]^{1/2}$. Finally, $\Rightarrow$ denotes weak convergence of the associated probability measures, and $[x]$ denotes the integer part of the real number x.

2. MODEL SPECIFICATION AND DISCRETE TIME REPRESENTATION

Consider the continuous time triangular ECM

$$dy(\tau) = -JAy(\tau)d\tau + w(\tau)d\tau, \quad \tau > 0, \tag{1}$$

where the cointegrated variables of interest are contained in the $m\times 1$ vector $y(\tau)$, τ denotes the continuous time parameter, and $w(\tau)$ satisfies the following assumption.

Assumption 1. $w(\tau)$ is a wide-sense stationary separable continuous time random process for which the function $Ew(\tau)w(s)'$ is measurable. Furthermore, $\|w_j(\tau)\|_2 < \infty$ for $j = 1,\ldots,m$.

This assumption ensures that $w(\tau)$ is integrable; see Rozanov (1967, Theorem 2.3). It is weaker than requiring $w(\tau)$ to be mean square continuous, and hence $w(\tau)$ could, in principle, incorporate jumps, which can be important for the modelling of financial time series. Defining $y = [y_1', y_2']'$, where y_1 is $m_1 \times 1$, y_2 is $m_2 \times 1$, and $m_1 + m_2 = m$, the ECM representation is consistent with an underlying cointegrating relationship between the sub-vectors y_1 and y_2 such that $y_1 - Cy_2$ is stationary, where C denotes the $m_1 \times m_2$ matrix of cointegrating parameters. The matrix C enters (1) via the matrix $A = [I_{m_1}, -C]$, while $J = [I_{m_1}, 0]'$. The first m_1 equations of (1) then give $dy_1(\tau) = -[y_1(\tau) - Cy_2(\tau)]d\tau + w_1(\tau)d\tau$, while the last m_2 equations in (1) depict the common stochastic trends $dy_2(\tau) = w_2(\tau)d\tau$, where w has been partitioned conformably with y. The solution to (1) is given by

$$y(\tau) = \int_0^\tau e^{-(\tau - r)JA} w(r)dr + e^{-\tau JA} y(0), \quad \tau > 0, \tag{2}$$

where the matrix exponential e^A is defined by the infinite series $e^A = \sum_{j=0}^\infty A^j/j!$ and $y(0)$ represents the initial state.

It will be assumed that the vectors y_1 and y_2 are each comprised of a mixture of stock variables and flow variables. Without loss of generality the variables in each vector will be arranged with the stocks first followed by the flows, and the cointegrating matrix C will be partitioned accordingly, so that

$$y_1(\tau) = \begin{bmatrix} y_1^S(\tau) \\ y_1^F(\tau) \end{bmatrix}, \quad y_2(\tau) = \begin{bmatrix} y_2^S(\tau) \\ y_2^F(\tau) \end{bmatrix}, \quad C = \begin{bmatrix} C_{SS} & C_{SF} \\ C_{FS} & C_{FF} \end{bmatrix}.$$

The vectors y_1^S and y_1^F are of dimensions $m_1^S \times 1$ and $m_1^F \times 1$ respectively, with $m_1^S + m_1^F = m_1$, while the subvectors of y_2 are of similarly-defined dimensions with $m_2^S + m_2^F = m_2$. The *sampling interval*, i.e. the period between observations, will be denoted by h, so that the *sampling frequency* is given by h^{-1}. Observations on the stock variables are made at points in time separated by a period of h while observations on flow variables are of the form of integrals of the underlying rate of flow over each successive interval of length h. That the necessary integrals exist is assured by Assumption 1. Introducing the variable t to index observations, the observations are of the form

$$y_{1,th} = \begin{bmatrix} y_{1,th}^S \\ y_{1,th}^F \end{bmatrix} = \begin{bmatrix} y_1^S(th) \\ \frac{1}{h}\int_0^h y_1^F(th - s)ds \end{bmatrix},$$

$$y_{2,th} = \begin{bmatrix} y_{2,th}^S \\ y_{2,th}^F \end{bmatrix} = \begin{bmatrix} y_2^S(th) \\ \frac{1}{h}\int_0^h y_2^F(th - s)ds \end{bmatrix},$$

where $t = 1, \ldots, T$ and T denotes the sample size. Denoting the span of the data by N, it follows that $T = N/h$. The observations are therefore made at the points th $(t = 1, \ldots, T)$, which divides (continuous) time (indexed by τ) into T intervals each of length h. Note that the flow variables are normalised by the factor $1/h$. The importance of this normalisation will become apparent below. The formulae defining the discrete time ECM are presented in Lemma 1.

Lemma 1. *Let $y(\tau)$ be generated by (1) and let $y_{th} = \left[y'_{1,th},\ y'_{2,th}\right]'$ $(t = 1, \ldots, T)$ denote the vector of observations on y_1 and y_2. Then, under Assumption 1, y_{th} satisfies the triangular ECM given by*

$$\Delta_h y_{th} = -\phi_h JA y_{th-h} + \xi_{th}, \quad t = 1, \ldots, T, \tag{3}$$

where $\Delta_h = 1 - L^h$, $\phi_h = 1 - e^{-h}$, and the subvectors of ξ_{th} are related to $w(\tau)$ as follows:

$$\begin{aligned}
\xi^S_{1,th} &= \int_0^h [1-\phi(r)]\, w^S_1(th-r)dr + C_{SS}\int_0^h \phi(r) w^S_2(th-r)dr \\
&\quad + C_{SF}\left[\int_0^h \phi(r) w^F_2(th-r)dr + \frac{\phi_h}{h}\int_0^h \psi_1(r) w^F_2(th-h-r)dr \right.\\
&\quad \left. + \frac{\phi_h}{h}\int_0^h \psi_2(r) w^F_2(th-2h-r)dr\right], \\
\xi^F_{1,th} &= \frac{1}{h}\int_0^h \phi(r) w^F_1(th-r)dr + \frac{1}{h}\int_0^h [\phi_h - \phi(r)]\, w^F_1(th-h-r)dr \\
&\quad + C_{FS}\left[\frac{1}{h}\int_0^h \psi_5(r) w^S_2(th-r)dr + \frac{1}{h}\int_0^h \psi_6(r) w^S_2(th-h-r)dr \right.\\
&\quad \left. - \frac{\phi_h}{h}\int_0^h \psi_1(r) w^S_2(th-h-r)dr - \frac{\phi_h}{h}\int_0^h \psi_2(r) w^S_2(th-2h-r)dr\right] \\
&\quad + C_{FF}\left[\frac{1}{h}\int_0^h \psi_5(r) w^F_2(th-r)dr + \frac{1}{h}\int_0^h \psi_6(r) w^F_2(th-h-r)dr\right], \\
\xi^S_{2,th} &= \int_0^h w^S_2(th-r)dr, \\
\xi^F_{2,th} &= \frac{1}{h}\int_0^h \psi_3(r) w^F_2(th-r)dr + \frac{1}{h}\int_0^h \psi_4(r) w^F_2(th-h-r)dr,
\end{aligned}$$

where $\phi(x) = 1 - e^{-x}$, $\xi_t = [\xi'_{1t}, \xi'_{2t}]' = [\xi^{S\prime}_{1t}, \xi^{F\prime}_{1t}, \xi^{S\prime}_{2t}, \xi^{F\prime}_{2t}]'$, and

$$\begin{aligned}
&\psi_1(x) = [h^2 - (x-h)^2]/2, \quad && \psi_2(x) = (x-h)^2/2, \quad && \psi_3(x) = x, \\
&\psi_4(x) = h - x, && \psi_5(x) = x - \phi(x), && \psi_6(x) = h - x - [\phi_h - \phi(x)].
\end{aligned}$$

The dynamics of the continuous time system, embodied in the stationary process $w(\tau)d\tau$ in (1), feed through into the discrete time ECM disturbance ξ_{th} via the sequence of formulae given in Lemma 1. Even in the simplest case in which $w(\tau)d\tau$ is an orthogonal increment process in continuous time, these formulae show that the dynamics of ξ_{th} will be rather more sophisticated than white noise. In particular, the presence of the lagged integrals imposes a higher-order moving average onto the discrete time dynamics. The discrete time triangular ECM representation in (3) exists provided that $w(\tau)$ is wide-sense integrable, which follows from Assumption 1.[1]

The normalisation of the flow variables by the factor $1/h$ puts them into the same units of measurement, regardless of the value of h. For example, suppose that $h = 1$ corresponds to one year and that y denotes the rate of flow of consumers' expenditure in dollars. Then $y_{t1} = \int_0^1 y(t-r)dr$ denotes annual consumers' expenditure measured in dollars *per annum*. If, however, the sampling frequency is quarterly, so that $h = 1/4$, then $y_{t\frac{1}{4}} = \int_0^{1/4} y(t\frac{1}{4} - r)dr$ measures quarterly consumption in dollars per quarter, while $[1/(1/4)]y_{t\frac{1}{4}}$ measures quarterly consumption in dollars *per annum*. Hence, in the latter case, the units of measurement remain constant, regardless of the sampling frquency h.

There are, however, even more important statistical reasons for normalising the flow variables by the factor $1/h$. One of these concerns the very validity of the discrete time ECM representation itself. Inspection of the derivation of the formulae defining $\xi_{1,th}$ in the proof of Lemma 1 in Appendix B reveals that terms of the form $y(th) - h^{-1}\int_0^h y(th-s)ds$ feature prominently. Lemma A1 in Appendix A provides a representation for this difference in terms of an integral of $w(\tau)$ in the form

$$y(th) - \frac{1}{h}\int_0^h y(th-s)ds = \frac{1}{h}\int_0^h (h-s)w(th-s)ds,$$

which is clearly stationary if $w(\tau)$ is stationary. However, if the flow variables were not normalised in this way, the resulting expression would be (using the proof of Lemma A1)

$$y(th) - \int_0^h y(th-s)ds = (1-h)y(th-h) + \int_0^h (1-s)w(th-s)ds.$$

The first term is clearly nonstationary, and hence its appearance as a component of $\xi_{1,th}$ would also render the discrete time ECM disturbance nonstationary as well. Thus the discrete time ECM would no longer be a valid representation of the cointegrated system.

The relationship between ξ_{th} and $w(\tau)$ can also be depicted in terms of a linear (matrix) filter. The form of this filter is presented in Lemma 2 below, although the precise definitions of its components are confined to Appendix B in order to avoid burdening the main text with unnecessary definitions.

[1]For a definition of wide-sense integrability, see Bergstrom (1984) or Rozanov (1967).

Lemma 2. *Let $y(\tau)$ be generated by (1). Then, under Assumption 1, the disturbance vector ξ_{th} in the discrete time ECM (3) is related to the disturbance vector $w(\tau)$ in the continuous time ECM (1) by the filtering equation $\xi_{th} = M_h(D)w(th)$, where the filter function $M_h(z)$ is defined by*

$$M_h(z) = \begin{bmatrix} m_1^S(z)I_{m_1^S} & 0 & m_{12}^{SS}(z)C_{SS} & m_{12}^{SF}(z)C_{SF} \\ 0 & m_1^F(z)I_{m_1^F} & m_{12}^{FS}(z)C_{FS} & m_{12}^{FF}(z)C_{FF} \\ 0 & 0 & m_2^S(z)I_{m_2^S} & 0 \\ 0 & 0 & 0 & m_2^F(z)I_{m_2^F} \end{bmatrix}$$

and its component filters are defined in Appendix B.

The filtering equation in Lemma 2 plays two important roles. First, it is particularly convenient for deriving the spectral density matrix of ξ_{th} from that of $w(\tau)$. If $f_{ww}^c(\lambda)$ $(-\infty < \lambda < \infty)$ denotes the spectral density matrix of the continuous time process $w(\tau)$, it follows, by noting that the frequency response function of the operator D is $i\lambda$, that the spectral density matrix of ξ_{th}, *regarded as a continuous time process*, is given by

$$f_{h,\xi\xi}^c(\lambda) = M_h(i\lambda)f_{ww}^c(\lambda)M_h(-i\lambda)', \quad -\infty < \lambda < \infty. \tag{4}$$

Note the dependence of this spectral density on h. The spectral density matrix of ξ_{th}, *regarded as a discrete time process*, is then obtained by applying the folding formula to (4) to yield[2]

$$f_{h,\xi\xi}(\lambda) = \sum_{k=-\infty}^{\infty} f_{h,\xi\xi}^c\left(\lambda + \frac{2k\pi}{h}\right), \quad -\frac{\pi}{h} < \lambda < \frac{\pi}{h}. \tag{5}$$

The spectral density function (5) plays a role in the asymptotics in section 4. Note that the range of λ in (5) is $(-\pi/h, \pi/h]$. In the limit, as $h \downarrow 0$, sampling becomes continuous and $f_{h,\xi\xi}(\lambda) \to f_{h,\xi\xi}^c(\lambda)$.

The second important role of the filtering equation in Lemma 2 is in the investigation of the order of magnitude of ξ_{th} in terms of h as $h \downarrow 0$. This latter property is important for the asymptotic analysis in section 4, and is presented in Lemma 3.

[2]See Priestley (1981, pp.504–507) for details.

Lemma 3. *Under Assumption 1, the discrete time disturbance vector* $\xi_{th} = O_p(h)$ *as* $h \downarrow 0$*, and satisfies the decomposition* $\xi_{th} = \zeta_{th} + \rho_{th}$*, where*

$$\zeta_{th} = g_h(D)w(th) = \int_0^h w(th-s)ds = O_p(h) \quad and \quad \rho_{th} = Q_h(D)w(th) = O_p(h^2)$$

as $h \downarrow 0$*, and where* $Q_h(z) = [M_h(z) - g_h(z)I_m]$ *and* $g_h(z) = (1-e^{-hz})/z$*. Furthermore,* $\rho^S_{2,th} = 0$.

Lemma 3 suggests that some care may need to be taken with respect to estimating the cointegrating parameters in view of ξ_{th} tending to zero in probability with h. This will manifest itself more precisely in the next section in which issues of estimation are treated more fully. The stated orders of magnitude are obtained by investigating the orders of magnitude of the integrating filter $g_h(z)$ and of the various filters that constitute $M_h(z)$ and by noting that $w(\tau) = O_p(1)$. The orders of magnitude of the filters are derived in a sequence of lemmas in Appendix A. The decomposition of ξ_{th} into the integral of $w(\tau)$ plus a remainder plays an important role in establishing the asymptotic properties of partial sums of ξ_{th} and related quantities.

As a by-product of the type of analysis leading to the results in Lemma 3, it is interesting to note that the normalisation of flow variables by the factor $1/h$ has the effect of normalising the discrete time flow variable to be $O_p(1)$. To see this, consider the unnormalised scalar flow variable $Y_{th} = \int_0^h y(th-s)ds = g_h(D)y(th)$. Since $y(th) = O_p(1)$ and Lemma A4 in Appendix A establishes that $g_h(z) = O(h)$ as $h \downarrow 0$, it follows that $Y_{th} = O_p(h)$. The normalised variable $y_{th} = h^{-1}Y_{th}$ is then $O_p(1)$.

3. SPECTRAL REGRESSION ESTIMATION: SOME PRELIMINARIES

The discrete time ECM (3) provides the basis for estimating the unknown elements of the matrix C of cointegrating parameters. In principle, a variety of methods could be considered for this task. If a parametric model was specified for the continuous time disturbance vector $w(\tau)$ in (1) then it would be possible to derive the precise dynamic properties (autocovariance structure) of ξ_{th} and to apply (quasi)-likelihood methods to jointly estimate C and the parameters determining the evolution of $w(\tau)$. If this parametric model was a system of stochastic differential equations then the problem would be one of estimating a system of (higher-order) differential equations of reduced rank (reflecting the cointegration properties). Discrete time models that enable this to be carried out have been derived by Chambers (1999). In this paper, by contrast, weaker conditions are imposed on the continuous time disturbances with the aim being to treat the system dynamics in a nonparametric way. The spectral regression estimators proposed by Phillips (1991a) for continuous time systems and by Phillips (1991c) for discrete time systems are ideally suited to this task.

It is convenient to rewrite the ECM (3) in a form more amenable to application of the spectral regression estimators. The first m_1 equations of (3) may be written

$$\Delta_h y_{1,th} + \phi_h y_{1,th-h} = C\phi_h y_{2,th-h} + \xi_{1,th}, \tag{6}$$

while the last m_2 equations of (3) are simply $\Delta_h y_{2,th} = \xi_{2,th}$. Combining these equations and normalising by h (in view of Lemma 3) gives

$$Y_{th} = JCX_{th} + w_{th}, \quad t = 1, \ldots, T = N/h, \tag{7}$$

where $Y_{th} = h^{-1}[(\Delta_h y_{1,th} + \phi_h y_{1,th-h})', \Delta_h y_{2,th}']'$, $X_{th} = h^{-1}\phi_h y_{2,th-h}$ and $w_{th} = h^{-1}\xi_{th}$. Note that $Ew_{th} = 0$ and $Ew_{th}w_{th}' = O(1)$ as $h \downarrow 0$.

Three main scenarios will be considered with regard to the sampling scheme, reflecting different joint behaviour of span N and frequency h^{-1}. The first is where h is fixed but $N \uparrow \infty$. This represents the usual situation in which sample size $T(= N/h) \uparrow \infty$ but emphasizes the dependence on a given sampling frequency, not necessarily equal to unity. The second is where $h \downarrow 0$ and $N \uparrow \infty$ jointly, so that the data are tending towards a continuous record limit at the same time as span increases. The third case keeps N fixed but allows $h \downarrow 0$ so that a continuous record is the result in the limit but one which covers a fixed span. Note that in all cases sample size $T \uparrow \infty$.

The analysis of the estimators in the sampling schemes of interest is aided by considering a triangular array of random variables $\{\{y_{nt}\}_{t=1}^{T_n}\}_{n=1}^{\infty}$ and by allowing the span and data frequency to be indexed by n, giving N_n and h_n. In this setup sample size $T_n = N_n/h_n$ always tends to infinity with n, while $N_n \uparrow \infty$ or $N_n = N$ and $h_n \downarrow 0$ or $h_n = h$. The system (7) then becomes

$$Y_{nt} = JCX_{nt} + w_{nt}, \quad t = 1, \ldots, T_n = N_n/h_n, \tag{8}$$

where $Y_{nt} = Y_{th_n}$, $X_{nt} = X_{th_n}$, and $w_{nt} = w_{th_n}$. The linearity of (8) in the unknown matrix C makes this an appealing equation as regards estimation.

The spectral regression estimators utilise estimates of certain spectral density matrices. For generic random variables x and y, let $\Gamma_{n,xy}(s) = Ex_{nt}y_{nt+s}'$ denote the autocovariance function which is estimated by $\widehat{\Gamma}_{n,xy}(s) = T_n^{-1}\sum_{t=1}^{T_n-s} x_{nt}y_{nt+s}'$. That this autocovariance matrix depends on n arises from the dependence of sampling frequency on n, so in terms of h_n this function is in fact $Ex_{th_n}y_{th_n+sh_n}'$. Although $\Gamma_{n,xy}(s)$ is the covariance between random variables separated by sh_n time units it is notationally convenient to suppress h_n in the argument of this function. The cross spectral density function of x and y, sampled at intervals of h_n, is given by[3]

$$f_{n,xy}(\lambda) = \frac{h_n}{2\pi} \sum_{s=-\infty}^{\infty} \Gamma_{n,xy}(s) e^{-ish_n\lambda}, \quad -\frac{\pi}{h_n} < \lambda \leq \frac{\pi}{h_n}$$

and can be estimated using

$$\widehat{f}_{n,xy}(\lambda) = \frac{h_n}{2\pi} \sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) \widehat{\Gamma}_{n,xy}(s) e^{-ish_n\lambda}, \tag{9}$$

where M_n is a bandwidth parameter and $k(z)$ is a kernel (or weighting) function. The precise properties that M_n and $k(z)$ are assumed to possess are defined in Assumption 3 later in this paper.

Two spectral regression estimators will be considered. The first utilises information contained in the full frequency range $(-\pi/h_n, \pi/h_n]$ and is defined by[4]

$$\begin{aligned} \text{vec}\left(\widehat{C}_n\right) &= \left[\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left(\widehat{f}_{n,XX}(\omega_j)' \otimes J' \widehat{f}_{n,\widehat{w}\widehat{w}}(\omega_j)^{-1} J\right)\right]^{-1} \\ &\quad \times \left[\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left(I_{m_2} \otimes J' \widehat{f}_{n,\widehat{w}\widehat{w}}(\omega_j)^{-1}\right) \text{vec}\left(\widehat{f}_{n,YX}(\omega_j)\right)\right], \end{aligned} \tag{10}$$

where $\omega_j = \pi j/(h_n M_n)$ $(j = -M_n+1, \ldots, M_n)$ and $\widehat{w}_{nt}$ denotes a consistent estimator of w_{nt} obtained, for example, by taking the residuals from an OLS regression applied to (8). Since, from (8),

$$\text{vec}\left(\widehat{f}_{n,YX}(\omega_j)\right) = \left(\widehat{f}_{n,XX}(\omega_j)' \otimes J\right) \text{vec}\left(C_0\right) + \text{vec}\left(\widehat{f}_{n,wX}(\omega_j)\right),$$

where C_0 denotes the true value of the matrix C, it follows that

$$\text{vec}\left(\widehat{C}_n - C_0\right) = \left[\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \Theta_{nj}\right]^{-1} \left[\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \theta_{nj}\right],$$

where

$$\Theta_{nj} = \widehat{f}_{n,XX}(\omega_j)' \otimes J' \widehat{f}_{n,\widehat{w}\widehat{w}}(\omega_j)^{-1} J, \quad j = -M_n+1, \ldots, M_n, \tag{11}$$

$$\theta_{nj} = \left(I_{m_2} \otimes J' \widehat{f}_{n,\widehat{w}\widehat{w}}(\omega_j)^{-1}\right) \text{vec}\left(\widehat{f}_{n,wX}(\omega_j)\right), \quad j = -M_n+1, \ldots, M_n. \tag{12}$$

[3]See equation (7.1.14) of Priestley (1981).

[4]The expression in (10) differs from the corresponding expression in Phillips (1991a) due to the use of column vectorisation here as opposed to row vectorisation in that article.

The second estimator uses information solely in the frequency band ω_0 and is defined by

$$\mathrm{vec}\left(\widehat{C}_{n0}\right) = \left[\widehat{f}_{n,XX}(0)' \otimes J'\widehat{f}_{n,\widehat{w}\widehat{w}}(0)^{-1}J\right]^{-1}\left[I_{m_2} \otimes J'\widehat{f}_{n,\widehat{w}\widehat{w}}(0)^{-1}\right]\mathrm{vec}\left(\widehat{f}_{n,YX}(0)\right). \quad (13)$$

Normalising in the same way as for the full-band estimator yields $\mathrm{vec}(\widehat{C}_{n0} - C_0) = \Theta_{n0}^{-1}\theta_{n0}$. The asymptotic behaviour of these two estimators is determined by the asymptotic properties of the expressions (11) and (12), to which attention is now turned.

4. SPECTRAL REGRESSION ESTIMATION: ASYMPTOTIC RESULTS

4.1. Asymptotic distributions

For the purposes of investigating the asymptotic properties of the spectral regression estimator in the different sampling scenarios, it is necessary to make further assumptions about the stochastic environment in which $y(\tau)$, and hence y_{th}, evolve. For this purpose, let $(\Omega_w, \mathcal{F}, P)$ denote the probability space on which $w(\tau)$ is defined, and let $\mathcal{F}_a^b$ $(a < b)$ denote a σ-subfield of $\mathcal{F}$. The strong mixing coefficients, α_j, are then defined by $\alpha_j = \sup_t \alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty)$, where

$$\alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty) = \sup_{G\in\mathcal{F}_{-\infty}^t, H\in\mathcal{F}_{t+j}^\infty} |P(G\cap H) - P(G)P(H)|.$$

The mixing coefficients are said to be of size $-p$ if $\alpha_j = O(j^{-p-\epsilon})$ for some $\epsilon > 0$ as $j \uparrow \infty$, which ensures that $\sum_{j=1}^\infty \alpha_j^{1/p} < \infty$. The following assumption is made with regard to the continuous time disturbance process $w(\tau)$ in (1).

Assumption 2. For some $\delta > \eta > 2$, $w(\tau)$ is a stationary strong mixing continuous time process with zero mean, $\|w(\tau)\|_\delta < \infty$, and with strong mixing coefficients of size $-\delta\eta/(\delta-\eta)$. Furthermore, the spectral density function, $f_{ww}^c(\lambda)$ $(-\infty < \lambda < \infty)$, of $w(\tau)$ is Hermitian positive definite with elements satisfying $0 < f_{ww,jj}^c(\lambda) < \infty$ $(j = 1, \ldots, m)$ and $|f_{ww,jk}^c(\lambda)| < \infty$ $(j \neq k,\ j,k = 1,\ldots,m)$ for all $-\infty < \lambda < \infty$.

The assumption that $w(\tau)$ is a strong mixing process is particularly convenient in the present circumstances in which the disturbances in the discrete time ECM involve integrals of $w(\tau)$ over finite intervals and are, therefore, strong mixing themselves.[5] Furthermore the mixing coefficients of such integrals are of the same size as those of the underlying process.

[5]See, for example, Theorem 14.1 of Davidson (1994).

The estimator asymptotics make use of the limiting properties of various sample moments concerning X_{nt} and w_{nt}. These, in turn, can be derived from the properties of the partial sum process

$$S_{n[T_n r]} = \sum_{j=1}^{[T_n r]} w_{nj}, \quad r \in [0,1], \tag{14}$$

and of the composite process

$$U_{nT_n} = \frac{1}{T_n} \sum_{t=1}^{T_n} S_{n,t-1} w_{nt}'. \tag{15}$$

The limiting properties of these random quantities are presented in Lemma 4 for each of the three sampling schemes of interest.

Lemma 4. *(a) Under Assumptions 1 and 2, if $h_n = h$ and $N_n \uparrow \infty$ as $n \uparrow \infty$,*

$$\frac{1}{T_n^{1/2}} S_{n[T_n r]} \Rightarrow B_h(r), \tag{16}$$

$$\frac{1}{T_n} U_{nT_n} \Rightarrow \int_0^1 B_h dB_h' + \Lambda_h(1), \tag{17}$$

where B_h denotes a Brownian motion process with variance matrix $\Omega_h = 2\pi h^{-3} f_{h,\xi\xi}(0)$ and $\Lambda_h(1) = \sum_{k=1}^{\infty} \Gamma_{h,k}$ where $\Gamma_{h,k} = E w_0 w_{kh}' = h^{-2} E \xi_0 \xi_{kh}'$.
(b) Under Assumptions 1 and 2, if $h_n \downarrow 0$, $N_n \uparrow \infty$ and $h_n N_n \downarrow 0$ as $n \uparrow \infty$,

$$\frac{h_n^{1/2}}{T_n^{1/2}} S_{n[T_n r]} \Rightarrow B(r), \tag{18}$$

$$\frac{h_n}{T_n} U_{nT_n} \Rightarrow \int_0^1 B dB' + \Lambda(1), \tag{19}$$

where B denotes a Brownian motion process with variance matrix $\Omega = 2\pi f_{ww}^c(0)$ and $\Lambda(1) = \lim_{n\uparrow\infty} h_n \sum_{k=1}^{T_n - 1} E w_{n0} w_{nk}'$.
(c) Under Assumption 1, if $h_n \downarrow 0$ and $N_n = N$ as $n \uparrow \infty$,

$$h_n S_{n[T_n r]} \Rightarrow Z(Nr), \tag{20}$$

$$h_n^2 U_{nT_n} \Rightarrow \int_0^N Z dZ', \tag{21}$$

where $Z(x) = \int_0^x w(s) ds$.

Part (a) of Lemma 4 extends the usual analysis of partial sums of discrete time processes to the case where the sampling interval h is not equal to one. The mixing decay rate in

Assumption 2 is slightly stronger than is strictly needed for (16) to hold, which only requires $\sum_{j=1}^{\infty} \alpha_j^{1-2/\eta} < \infty$. For $\eta > 2$, note that $0 < 1 - 2/\eta < 1$ while $0 < (\delta - \eta)/\delta\eta < 1/2$. The latter condition satisfies the former and is required for (17).

Part (b) of Lemma 4 extends the analysis further to allow $h_n \downarrow 0$. Here Ω is expressed in terms of $f_{ww}^c(0)$, since $w_{nj} \xrightarrow{p} w(\tau)$ as $n \uparrow \infty$. The matrix $\Lambda(1)$ is left in the form of a limit because, in the analysis of the asymptotic properties of the spectral regression estimator in this case, this limit will be taken in conjunction with another limit at the appropriate point. Note, too, the requirement that $h_n N_n \downarrow 0$ as $n \uparrow \infty$. This ensures that the contribution of the higher-order (in h_n) terms contained in ρ_{nt} are negligible in the asymptotics and, as a result, simplifies the analysis somewhat. The requirement for this to be valid in practice is that the observation interval h_n gets smaller at a faster rate than the span N_n gets larger, and is perhaps not an unreasonable requirement.

Part (c) of Lemma 4 treats the case where an infinitely large sample size is obtained by allowing $h_n \downarrow 0$ while holding the span fixed. Such continuous record asymptotics were also considered by Phillips (1987a,b) and by Perron (1991) although much stronger assumptions were made in those articles concerning the underlying continuous time random process than are being made here.

Lemma 4 provides a basis for developing the asymptotic properties of the estimator $\widehat{C}_n$. From (11) and (12) it can be seen that it is the covariance matrix estimators $\widehat{\Gamma}_{n,XX}(s)$ and $\widehat{\Gamma}_{n,wX}(s)$ that will determine the relevant asymptotics via their use in the construction of the spectral density estimators $\widehat{f}_{n,XX}(\lambda)$ and $\widehat{f}_{n,wX}(\lambda)$ respectively. In order to examine the properties of these covariance matrix estimators, consider the $m \times 1$ integrated process defined by $y_{nt} = y_{nt-1} + \xi_{nt}$ $(t = 1, \ldots, T_n = N_n/h_n)$, where $y_{n0} = y_0 = y(0)$ is an $O_p(1)$ random variable and ξ_{nt} is defined in Lemma 1. Further, define $x_{nt} = (\phi_{h_n}/h_n)y_{nt-1}$ and $w_{nt} = h_n^{-1}\xi_{nt}$, so that with this notation the variable of interest is simply $X_{nt} = x_{2,nt}$. Since $y_{nt} = y_0 + \sum_{j=1}^t \xi_{nj}$, it follows that $x_{nt} = (\phi_{h_n}/h_n)y_0 + \phi_{h_n} S_{nt-1}$, where $S_{nt} = \sum_{j=1}^t w_{nt}$. Consider the random quantities $\mu_n(s) = \sum_{t=1}^{T_n - s} x_{nt} w'_{nt+s}$ and $M_n(s) = \sum_{t=1}^{T_n - s} x_{nt} x'_{nt+s}$. By making the appropriate substitutions for x_{nt} in terms of S_{nt-1}, it can be shown that

$$\mu_n(s) = \phi_{h_n}\left(U_{nT_n} - U_{ns}\right) + \frac{\phi_{h_n}}{h_n} y_0 \left(S'_{nT_n} - S'_{ns}\right) - \phi_{h_n} T_n \sum_{k=1}^{s} \widehat{\Gamma}_{n,ww}(k), \tag{22}$$

where $\widehat{\Gamma}_{n,ww}(k) = T_n^{-1} \sum_{t=1}^{T_n - s} w_{nt+s-k} w'_{nt+s}$, and that

$$\begin{aligned} M_n(s) \;=\; & \phi_{h_n}^2 T_n \int_0^{1-s/T_n} S_{n[T_n r]} S'_{n[T_n r]} dr + \frac{\phi_{h_n}^2}{h_n} T_n y_0 \int_0^{1-s/T_n} S'_{n[T_n r]} dr \\ & + \frac{\phi_{h_n}^2}{h_n} T_n \int_0^{1-s/T_n} S_{n[T_n r]} dr y'_0 + \frac{\phi_{h_n}^2}{h_n^2} (T_n - s) y_0 y'_0 + \phi_{h_n}^2 \sum_{j=0}^{s-1} \mu_n(j). \end{aligned} \tag{23}$$

Note that, if $\mu_n(s) = [\mu_{n,1}(s), \mu_{n,2}(s)]$ then $\widehat{\Gamma}_{n,wX}(s) = T_n^{-1}\mu_{n,2}(s)$ while $\widehat{\Gamma}_{n,XX}(s) = T_n^{-1}M_{n,22}(s)$ where $M_{n,22}(s)$ is the lower right-hand block of $M_n(s)$. These expressions, combined with the results in Lemma 4, enable the results of interest to be derived.

Lemma 5. *(a) Under Assumptions 1 and 2, if $h_n = h$ and $N_n \uparrow \infty$ as $n \uparrow \infty$,*

$$\widehat{\Gamma}_{n,wX}(s) \Rightarrow \phi_h \left[\int_0^1 dB_h B_{h2}' + \Lambda_{h2}(s+1)'\right], \tag{24}$$

$$\frac{1}{T_n}\widehat{\Gamma}_{n,XX}(s) \Rightarrow \phi_h^2 \int_0^1 B_{h2}B_{h2}', \tag{25}$$

where $\Lambda_h(s+1) = [\Lambda_{h1}(s+1)', \Lambda_{h2}(s+1)']' = \sum_{k=s+1}^{\infty}\Gamma_{h,k}$.
(b) Under Assumptions 1 and 2, if $h_n \downarrow 0$, $N_n \uparrow \infty$ and $h_nN_n \downarrow 0$ as $n \uparrow \infty$,

$$\frac{h_n}{\phi_{h_n}}\widehat{\Gamma}_{n,wX}(s) \Rightarrow \int_0^1 dBB_2' + \Lambda_2(s+1)', \tag{26}$$

$$\frac{h_n^2}{\phi_{h_n}^2 N_n}\widehat{\Gamma}_{n,XX}(s) \Rightarrow \int_0^1 B_2B_2', \tag{27}$$

where $\Lambda(s+1) = [\Lambda_1(s+1)', \Lambda_2(s+1)']' = \lim_{n\uparrow\infty} h_n \sum_{k=s+1}^{T_n-(s+1)} Ew_{n0}w_{nk}'$.
(c) Under Assumption 1, if $h_n \downarrow 0$ and $N_n = N$ as $n \uparrow \infty$,

$$\frac{h_nN_n}{\phi_{h_n}}\widehat{\Gamma}_{n,wX}(s) \Rightarrow F_2(Z, y_{02}) \equiv \int_0^N dZZ_2' + Z(N)y_{02}', \tag{28}$$

$$\frac{h_n^2}{\phi_{h_n}^2}\widehat{\Gamma}_{n,XX}(s) \Rightarrow F_1(Z_2, y_{02}) \equiv \int_0^N Z_2Z_2' + y_{02}\int_0^N Z_2' + \int_0^N Z_2y_{02}' + y_{02}y_{02}'. \tag{29}$$

The convergence rates in Lemma 5 determine the rates of convergence of the terms Θ_{nj} and θ_{nj} that are used in constructing the spectral regression estimators. Once more, the results in part (a) generalise existing results in the literature to the case where the sampling interval h is not equal to one. Part (b) provides the extension where the sampling interval h_n tends to zero, and in part (c) this is achieved but with span held fixed. Note the dependency of the results in part (c) on the initial value y_{02} which enters the limiting expressions because the data span is held fixed.

In order to consider the limiting distributions of the terms Θ_{nj} and θ_{nj}, it is necessary to impose some conditions on the bandwidth parameter M_n and the kernel function $k(z)$.

Assumption 3. $M_n = o(T_n^{1/2})$ as $n \uparrow \infty$ and $k(z)$ is an even, bounded function for $z \in [-1,1]$ with $k(0) = 1$ and $k(z) = 0$ for $z \notin [-1,1]$.

These conditions are quite standard in the spectral regression literature; see, for example, Hannan (1963). It is also convenient to define the constant $\nu = (1/2\pi)\int_{-1}^{1} k(s)ds$ which appears in the results below.

Theorem 1. *(a) Under Assumptions 1–3, if $h_n = h$ and $N_n \uparrow \infty$ as $n \uparrow \infty$,*

$$\frac{1}{T_n}\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\Theta_{nj}\right] \Rightarrow \phi_h^2 h \int_0^1 B_{h2}B_{h2}' \otimes J'\Omega_h^{-1}J, \tag{30}$$

$$\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\theta_{nj}\right] \Rightarrow \phi_h h \left(I_{m_2} \otimes J'\Omega_h^{-1}\right) \mathrm{vec}\left(\int_0^1 dB_h B_{h2}'\right), \tag{31}$$

$$\frac{1}{T_n M_n}\Theta_{n0} \Rightarrow \nu\phi_h^2 h \int_0^1 B_{h2}B_{h2}' \otimes J'\Omega_h^{-1}J, \tag{32}$$

$$\frac{1}{M_n}\theta_{n0} \Rightarrow \nu\phi_h h \left(I_{m_2} \otimes J'\Omega_h^{-1}\right) \mathrm{vec}\left(\int_0^1 dB_h B_{h2}'\right). \tag{33}$$

(b) Under Assumptions 1–3, if $h_n \downarrow 0$, $N_n \uparrow \infty$ and $h_n N_n \downarrow 0$ as $n \uparrow \infty$,

$$\frac{h_n^2}{\phi_{h_n}^2 N_n}\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\Theta_{nj}\right] \Rightarrow \int_0^1 B_2 B_2' \otimes J'\Omega^{-1}J, \tag{34}$$

$$\frac{h_n}{\phi_{h_n}}\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\theta_{nj}\right] \Rightarrow \left(I_{m_2} \otimes J'\Omega^{-1}\right) \mathrm{vec}\left(\int_0^1 dB B_2'\right), \tag{35}$$

$$\frac{1}{T_n\phi_{h_n}^2 M_n}\Theta_{n0} \Rightarrow \nu \int_0^1 B_2 B_2' \otimes J'\Omega_h^{-1}J, \tag{36}$$

$$\frac{1}{\phi_{h_n} M_n}\theta_{n0} \Rightarrow \nu \left(I_{m_2} \otimes J'\Omega_h^{-1}\right) \mathrm{vec}\left(\int_0^1 dB B_2'\right). \tag{37}$$

(c) Under Assumptions 1 and 3, if $h_n \downarrow 0$ and $N_n = N$ as $n \uparrow \infty$,

$$\frac{h_n^2}{\phi_{h_n}^2}\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\Theta_{nj}\right] \Rightarrow F_1(Z_2, y_{02}) \otimes J'\Omega^{-1}J, \tag{38}$$

$$\frac{h_n N_n}{\phi_{h_n}}\left[\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n}\theta_{nj}\right] \Rightarrow \left(I_{m_2} \otimes J'\Omega^{-1}\right) \mathrm{vec}\left(F_2(Z, y_{02})\right), \tag{39}$$

$$\frac{h_n}{\phi_{h_n}^2 M_n}\Theta_{n0} \Rightarrow \nu F_1(Z_2, y_{02}) \otimes J'\Omega^{-1}J, \tag{40}$$

$$\frac{N}{\phi_{h_n} M_n}\theta_{n0} \Rightarrow \nu \left(I_{m_2} \otimes J'\Omega^{-1}\right) \mathrm{vec}\left(F_2(Z, y_{02})\right). \tag{41}$$

Using Theorem 1 it is a straightforward task to derive the limiting distributions of the appropriately normalised estimators, which are presented in Theorem 2.

Theorem 2. *(a) Under Assumptions 1–3, if* $h_n = h$ *and* $N_n \uparrow \infty$ *as* $n \uparrow \infty$*,*

$$T_n \text{vec}\left(\widehat{C}_n - C_0\right) \Rightarrow \phi_h^{-1}\left[\left(\int_0^1 B_{h2}B_{h2}'\right)^{-1} \otimes I\right]\left[\int_0^1 B_{h2} \otimes dB_{h,1.2}\right], \tag{42}$$

where $B_{h,1.2} = B_{h1} - \Omega_{h,12}\Omega_{h,22}^{-1}B_{h2}$ *is Brownian motion with variance matrix* $\Omega_{h,11.2} = \Omega_{h,11} - \Omega_{h,12}\Omega_{h,22}^{-1}\Omega_{h,21}$*. The same result holds for* $T_n\text{vec}\left(\widehat{C}_{n0} - C_0\right)$*.*
(b) Under Assumptions 1–3, if $h_n \downarrow 0$*,* $N_n \uparrow \infty$ *and* $h_nN_n \downarrow 0$ *as* $n \uparrow \infty$*,*

$$T_n\phi_{h_n}\text{vec}\left(\widehat{C}_n - C_0\right) \Rightarrow \left[\left(\int_0^1 B_2B_2'\right)^{-1} \otimes I\right]\left[\int_0^1 B_2 \otimes dB_{1.2}\right], \tag{43}$$

where $B_{1.2} = B_1 - \Omega_{12}\Omega_{22}^{-1}B_2$ *is Brownian motion with variance matrix* $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$*. The same result holds for* $T_n\phi_{h_n}\text{vec}\left(\widehat{C}_{n0} - C_0\right)$*.*
(c) Under Assumptions 1 and 3, if $h_n \downarrow 0$ *and* $N_n = N$ *as* $n \uparrow \infty$*,*

$$T_n\phi_{h_n}\text{vec}\left(\widehat{C}_n - C_0\right) \Rightarrow \left[F_1(Z_2, y_{02})^{-1} \otimes I\right]\left\{\left[\int_0^N Z_2 \otimes dZ_{1.2}\right] + \left[y_{02} \otimes Z_{1.2}(N)\right]\right\}, \tag{44}$$

where $Z_{1.2} = Z_1 - \Omega_{12}\Omega_{22}^{-1}Z_2$*. The same result holds for* $T_n\phi_{h_n}\text{vec}\left(\widehat{C}_{n0} - C_0\right)$*.*

In part (a) of Theorem 2, since h is fixed and $T_n = N_n/h$, it is clearly the increasing span of the data that is important in this case. The result can be written in terms of $N_n\text{vec}(\widehat{C}_n - C_0)$ with the limiting distribution being that given in (42) multiplied by h. Note that the distribution in (42) is the familiar mixed normal distribution from cointegration theory, as would be expected. If $P_h(G_h)$ denotes the probability measure associated with the random matrix $G_h = (\int_0^1 B_{h2}B_{h2}')^{-1}$, then the distribution has the representation

$$\int_{G_h>0} N\left(0, \phi_h^{-2}G_h \otimes \Omega_{h,11.2}\right) dP_h(G_h), \tag{45}$$

which, conditional on a given realisation of y_2, is normal. Similar comments apply to part (b) of Theorem 2. In this case, $T_n\phi_{h_n} = N_n\phi_{h_n}/h_n$. Since $\phi_{h_n}/h_n \to 1$ as $n \uparrow \infty$, it follows that $N_n\text{vec}(\widehat{C}_n - C_0)$ has the same distribution in the limit as that given in (43), which has the mixed normal representation

$$\int_{G>0} N\left(0, G \otimes \Omega_{11.2}\right) dP(G), \tag{46}$$

where $G = (\int_0^1 B_2 B_2')^{-1}$ and $P(G)$ is its associated probability measure.

In part (c) of Theorem 2, it is span that is fixed, and since $T_n \phi_{h_n} = N\phi_{h_n}/h_n \to N$ as $n \uparrow \infty$, it follows that $\text{vec}(\widehat{C}_n - C_0)$ has the same limiting distribution as N^{-1} times the distribution in (44). Furthermore, if $y_{02} = 0$, then the asymptotics are governed by

$$T_n \phi_{h_n} \text{vec}\left(\widehat{C}_n - C_0\right) \Rightarrow \left[\left(\int_0^N Z_2 Z_2'\right)^{-1} \otimes I\right]\left[\int_0^N Z_2 \otimes dZ_{1.2}\right], \tag{47}$$

since the final component in (44) is null. The distribution in (44) depends on the distribution of the underlying continuous time disturbance process $w(\tau)$ via the variable Z. In cases where $w(\tau)d\tau$ has independent increments[6] and variance $\Sigma d\tau$, the random variable $Z(x) = \int_0^x w(\tau)d\tau$ is Brownian motion with variance Σx. The limiting distribution in (47) is then the familiar mixed normal distribution, but when $y_{02} \neq 0$ the distribution in (44) contains an additional term involving $Z_{1.2}(N) \sim N(0, \Sigma_{11.2})$ with $\Sigma_{11.2}$ defined in terms of the sub-matrices of Σ in the same way that $\Omega_{11.2}$ is defined in terms of the sub-matrices of Ω.

4.2. Efficiency comparisons

Theorem 2 enables some interesting questions concerning the effects of observation frequency on the estimation of cointegrating parameters to be addressed. Although a number of comparisons could be explored, one in particular is addressed here. This concerns the potential inefficiency that might be conjectured to arise as a result of having a fixed sampling interval h as compared to a continuous sample (the limiting case when $h \downarrow 0$). Investigations of the asymptotic bias (as a function of sampling interval) of estimators of the parameters of stationary continuous time systems, derived from approximate discrete time models, are well established; see Bergstrom (1984) for a summary. Rather less attention has been paid to the efficiency of estimators, as measured by the variance of the asymptotic distribution, although Chambers (2000) provides some results for cointegration estimators that focus on the effects of the way in which data are recorded (i.e. stocks versus flows).

The analysis will be based on the limiting distribution of $N_n \text{vec}(\widehat{C}_n - C_0)$ which (when h is fixed) is given by (45) with the covariance matrix multiplied by h^2. A more convenient

[6]In this case it would be common to write $w(\tau)d\tau = \zeta(d\tau)$, where $\zeta(d\tau)$ is a vector random measure.

representation of this distribution, for the purposes of making comparisons, is given by[7]

$$\int_{\gamma>0} N\left(0, \gamma V(h)\right) dP_{\gamma}(\gamma), \tag{48}$$

where $V(h) = h^2\phi_h^{-2}\Omega_{h,22}^{-1} \otimes \Omega_{h,11.2}$, $\gamma = e_2'(\int_0^1 W_2 W_2')^{-1}e_2$, e_2 is any unit $m_2 \times 1$ vector, $P_\gamma(\gamma)$ is the probability measure associated with γ, and W_2 is an $m_2 \times 1$ vector of standard Brownian motions or Wiener processes (i.e. W_2 is Brownian motion with covariance matrix I_{m_2}). When h is allowed to tend to the limit of zero, the relevant distribution is given by (46), which may be written in the more covenient form

$$\int_{\gamma>0} N\left(0, \gamma V_0\right) dP_{\gamma}(\gamma), \tag{49}$$

where $V_0 = \Omega_{22}^{-1} \otimes \Omega_{11.2}$. Notice that the mixing variate, γ, is the same as in the fixed-h case in (48), because it is purely the sampling frequency that is different, not the underlying random process ($w_2(\tau)$ in continuous time) that generates the data. The precise form of the matrix difference $V(h) - V_0$ is given in Theorem 3.

Theorem 3. *Under Assumptions 1–3, the difference* $V(h) - V_0 = f_{ww,22}^c(0)^{-1} \otimes \widetilde{V}_h$ *is positive semi-definite for any fixed* $h > 0$, *where*

$$\widetilde{V}_h = \begin{bmatrix} \widetilde{V}_{h,11} & 0 \\ 0 & 0 \end{bmatrix},$$

$\widetilde{V}_{h,11} = J_C \sum_{k\neq 0} \left[1 + (4\pi^2k^2/h^2)\right]^{-1} f_{ww}^c(2\pi k/h) J_C'$ *and* $J_C = [I_{m_1^S},\ 0,\ -C_{SS},\ -C_{SF}]$.

Theorem 3 shows that there is an inefficiency associated with discrete time sampling relative to continuous sampling in view of $V(h)-V_0$ being positive semi-definite. Inspection of this matrix difference shows that, in fact, this inefficiency can be more accurately pinpointed. The qualitative implication of Theorem 3 is presented in Proposition 1 below.

Proposition 1. *The estimator inefficiencies caused by sampling at a discrete interval* h *only affect the estimation of the matrices* C_{SS} *and* C_{SF}. *The estimation of the matrices* C_{FS} *and* C_{FF} *is as efficient when based on data sampled at intervals of length* h *as when based on a continuous record of data.*

[7]Details of the equivalence of the representations (45) and (48) can be found in Phillips (1989).

Proposition 1 establishes that it is in the estimation of the cointegrating relationships in which the normalised (left-hand side) variables are stocks where the inefficiencies will arise. It is quite remarkable that, as far as the estimation of the matrices C_{FS} and C_{FF} is concerned i.e. the parameters of the cointegrating relationships in which the left-hand side (normalised) variables are flows, that there is no efficiency gain to be made from a continuous sample as compared to a fixed sampling interval of length h. This is presumably a result of flows being observed as integrals over the interval $(th-h, th]$ and hence contain information about the evolution of the variable over that interval. With stocks, however, such information is not contained in the observations which are made at points in time. Such results are in accordance with the findings of Chambers (2000) who demonstrated, for a fixed sampling interval $h = 1$, that the discrete time sampling of stock variables results in a loss of estimator efficiency as compared to flow variables, which in turn are as efficient as continuous sampling. Theorem 3 and Proposition 1 generalise these results in two directions. First, they allow for an arbitrary sampling interval h, and secondly, they are based on a more general system containing stocks and flows simultaneously.

4.3. Large sample inference

One of the principal advantages of the mixed normal limiting distribution of the estimators is that inference concerning the cointegrating parameters can be conducted using traditional methods, as emphasised by Phillips (1991b). For example, the usual t-ratios have limiting normal distributions, confidence intervals can be constructed using normal critical values, and Wald tests of possibly non-linear restrictions on the cointegrating parameters can be based on asymptotic chi-square criteria. Such comments obviously apply to situations in which the span of the data tends to infinity, regardless of whether the sampling frequency is fixed or not. This is not the case, however, when span is fixed and the sampling interval tends to zero, because the limiting distribution in this case is typically not mixed normal; see Theorem 2(c). This suggests that small-h asymptotic inference may be difficult in these circumstances and hence the focus here is on situations in which span tends to infinity (covered by Theorem 2, parts (a) and (b)).

A question that arises concerns the estimation of the asymptotic variances to use in t-tests or Wald tests. The case in which $h_n \downarrow 0$ and $N_n \uparrow \infty$ will be considered here, although the same arguments apply if $h_n = h$ is fixed, with minor modifications. Conditional on the realisation $\{y_{2,th}\}$, it follows from Theorem 2(b) and the representation (46) that $T_n \phi_{h_n} \text{vec}(\widehat{C}_n - C_0)$ is asymptotically $N(0, G \otimes \Omega_{11.2})$, where $G = (\int_0^1 B_2 B_2')^{-1}$ denotes the limit of the matrix $[(h_n/(\phi_{h_n}^2 T_n^2)) \sum_{t=1}^{T_n} X_{nt} X_{nt}']^{-1}$; see Lemma 5(b). Large sample (conditional) inference for $\text{vec}(\widehat{C}_n - C_0)$ can therefore be based on the distribution $N(0, \widehat{V}_1)$, where

$$\widehat{V}_1 = \frac{1}{T_n^2 \phi_{h_n}^2} \left[\left(\frac{h_n}{\phi_{h_n}^2 T_n^2} \sum_{t=1}^{T_n} X_{nt} X_{nt}' \right)^{-1} \otimes \widehat{\Omega}_{11.2} \right] \tag{50}$$

and the estimator $\widehat{\Omega}_{11.2}$ is derived from the sub-matrices of the matrix $\widehat{\Omega} = (2\pi/h_n)\widehat{f}_{n,\widehat{w}\widehat{w}}(0)$. The variance matrix $\widehat{V}_1$ is not, however, the usual covariance matrix estimator associated with spectral regression. The estimator suggested by Phillips (1991a,c) for constructing the Wald statistic to test hypotheses concerning the cointegrating parameters is based on the theory of spectral regression for stationary time series[8], suitably adapted for the faster rate of convergence of the estimator in the case of cointegration. This estimator is given by

$$\widehat{V}_2 = \frac{1}{T_n} \left[\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \Theta_{nj} \right]^{-1}, \tag{51}$$

where Θ_{nj} is defined in (11). It is this estimator that is usually computed in spectral regression software packages.[9]

Analagous expressions for the covariance matrix estimators can be derived for the band-limited spectral regression estimator, $\widehat{C}_{n0}$. Since $T_n \phi_{h_n} \text{vec}(\widehat{C}_{n0} - C_0)$ is also asymptotically $N(0, G \otimes \Omega_{11.2})$ conditional on the realisation $\{y_{2,th}\}$, large sample inference for $\text{vec}(\widehat{C}_{n0} - C_0)$ can be based on the distribution $N(0, \widehat{V}_{10})$, where $\widehat{V}_{10} = \widehat{V}_1$. Although the same expression is used for the covariance matrix estimator, the two will only coincide numerically if the same kernel function and bandwidth value are used in the construction of the spectral density estimates. The analogue of the second estimator is, however, different, and is given by[10]

$$\widehat{V}_{20} = \frac{M_n \nu}{T_n} \Theta_{n0}^{-1}, \tag{52}$$

with Θ_{n0} defined in (11) and ν defined prior to Theorem 1. The relative effects of the two types of covariance matrix estimator on conducting inference in finite samples is explored in the simulation experiments that follow.

[8] See, for example, Hannan (1970, p.442).

[9] For example, the COINT module in GAUSS uses the expression (51).

[10] See Phillips (1991b, pp.423–424) for details.

5. SOME SIMULATION RESULTS AND AN EMPIRICAL EXAMPLE

5.1. *Simulation results*

A small simulation experiment was conducted in order to assess the finite sample properties of the spectral regression estimators of the cointegrating parameters when sampling frequency and/or span varies. The simulation model consists of a bivariate system of stock variables in which the continuous time disturbance process follows a first-order stochastic differential equation system. The variables therefore evolve in continuous time according to

$$dy_1(\tau) = -[y_1(\tau) - Cy_2(\tau)]d\tau + w_1(\tau)d\tau, \quad (53)$$

$$dy_2(\tau) = w_2(\tau)d\tau. \quad (54)$$

There is a single cointegrating parameter which is set to unity in the simulations i.e. $C = 1$ in (53), while the initial condition for the system is taken to be $y_1(0) = y_2(0) = 0$. The disturbance vector $w(\tau) = [w_1(\tau),\ w_2(\tau)]'$ satisfies

$$dw(\tau) = Gw(\tau)d\tau + \zeta(d\tau), \quad (55)$$

where $\zeta(d\tau)$ is a Gaussian random measure satisfying $E\zeta(d\tau) = 0$, $E\zeta(d\tau)\zeta(d\tau)' = I_2 d\tau$, and $E\zeta(\tau_2 - \tau_1)\zeta(\tau_4 - \tau_3)' = 0$ whenever the intervals $\tau_2 - \tau_1$ and $\tau_4 - \tau_3$ do not intersect. The matrix G in (55) is assumed to take the form

$$G = \begin{bmatrix} \gamma_1 & 0 \\ \gamma_2 & -1.5 \end{bmatrix},$$

where the parameter γ_1 is required to be negative for $w(\tau)$ to be stationary and the parameter γ_2 represents the strength of feedback from w_1 to w_2. The assumed values are $\gamma_1 \in \{-0.5, -5\}$ and $\gamma_2 \in \{-1, +1\}$, so that both positive and negative feedback are considered. This yields four combinations of parameters, referred to as Experiments 1 to 4:

Experiment 1: $\gamma_1 = -0.5, \quad \gamma_2 = -1;$ Experiment 2: $\gamma_1 = -5, \quad \gamma_2 = -1;$
Experiment 3: $\gamma_1 = -0.5, \quad \gamma_2 = 1;$ Experiment 4: $\gamma_1 = -5, \quad \gamma_2 = 1.$

In view of the system being comprised solely of stock variables, a modification of Lemma 1 reveals that the discretely observed vector y_{th} satisfies the discrete time triangular ECM of (3) but with the disturbance vector defined as

$$\xi_{th} = \int_0^h e^{-sJA} w(th - s)ds = \int_0^h [I_2 - \phi(s)JA]w(th - s)ds.$$

The discrete time data can therefore be generated using (3) once a set of discrete time disturbances $\{\xi_{th}\}_{t=1}^{T}$ satisfying the appropriate properties have been generated. Details of how this is achieved are provided in Appendix D but suffice it to note here that ξ_{th} is an ARMA(1,1) process that satisfies quite complicated restrictions that arise as a result of the temporal aggregation.

In order to assess the effects of sampling frequency and span on the estimators, three frequencies and three spans were considered in the simulations. More precisely, the values for the sampling interval are $h \in \{1/12, 1/4, 1\}$, corresponding to monthly, quarterly and annual frequencies, while $N \in \{25, 50, 100\}$, corresponding to spans of 25, 50 and 100 years. It is only necessary to generate 100 years of monthly data for each replication of the experiment, because the shorter spans, as well as the less frequently observed series, are simply obtained from this underlying series of 1200 observations. A total of 10,000 replications of each of the four experiments (or parameter combinations) were conducted.

The spectral regression estimators require a choice of kernel function and bandwidth value in order to become operational. The Parzen kernel was chosen in view of its relatively superior performance against an averaged periodogram estimator of the spectral density function in the study of Chambers (2001). The Parzen kernel is defined by

$$k(z) = \begin{cases} 1 - 6z^2 + 6|z|^3, & |z| \leq 1/2, \\ 2(1 - |z|)^3, & 1/2 < |z| \leq 1, \\ 0, & |z| > 1. \end{cases}$$

Concerning bandwidth choice, a pilot simulation study for each experiment, consisting of 1000 replications, was conducted using bandwidths of the form $M = [T^{\alpha}]$, for values of $\alpha \in \{1/10, 1/5, 1/3, 2/5\}$. Such choices of M are clearly $o(T^{1/2})$ as required. As there is typically a trade-off between bias and variance in choosing the bandwidth parameter, the bandwidth that resulted in the smallest mean square error (MSE) of the estimator (as sample size increases) was chosen. The resulting values were $\alpha = 1/10$ for the spectral estimator and $\alpha = 2/5$ for the band-limited spectral estimator. These MSE-minimising values were the same for each of the four experiments, and were employed to compute the values of the estimators in the simulations.

The MSEs of the two spectral estimators are reported in Table 1, which also includes the MSEs of the OLS estimator of the parameter C for purposes of comparison. The spectral estimator $\widehat{C}_n$ is denoted SPEC in Table 1, while the band limited estimator $\widehat{C}_{n0}$ is denoted BAND. Note that the MSEs are of the same order of magnitude for each h although they decrease, as is to be expected, with increasing N. They also have a tendency to be smaller when the parameter $\gamma_1 = -5$ (Experiments 2 and 4) than when $\gamma_1 = -0.5$ (Experiments 1 and 3), for given γ_2. Since γ_1 represents a root of the system, this suggests that the estimation is more precise when such roots are larger in absolute value. There is also a tendency for the band-limited spectral estimator to have smaller MSEs on the whole than the spectral estimator that is constructed using all frequency bands. It is also interesting to note that the MSEs of the spectral regression estimators in Experiment 1 are comparable to those of the OLS estimator, so there would appear to be relatively little gained in this case by using spectral regression, at least judged by this particular criterion. In Experiments 2, 3 and 4, however, the MSEs of both spectral estimators are uniformly smaller than those of the OLS estimator, except for the estimator using all frequency bands when $h = 1/12$ in Experiments 2 and 4.

[Table 1 about here.]

The inspection of MSEs, whilst providing useful information about the estimators, represents only a partial assessment of their performance. Estimated coefficients are typically used to make inferences concerning the true (but unknown) value of the parameter. This is commonly achieved by conducting a t-test, which requires an estimate of the variance of the estimated coefficient. This variance estimate can also be employed in the construction of confidence intervals, and it is confidence intervals that shall be considered here. Tables 2 and 3 contain the percentage coverage rates of 90% and 95% confidence intervals for the OLS estimator and the two spectral estimators. These coverage rates represent the proportion of the replications in which the true value of the coefficient ($C_0 = 1$) fell within the calculated confidence interval. For each of the spectral estimators two coverage rates are reported, each one based on the different estimators of the variance considered in section 4.3. In the simulations here, the Parzen kernel is used for both estimators but the bandwidths are different, so that the computed values of $\widehat{V}_1$ and $\widehat{V}_{10}$ will be different. Also, for the Parzen kernel employed here, it is straightforward to show that the constant $\nu = 0.75$.

[Tables 2 and 3 about here.]

Inspection of Tables 2 and 3 reveals a number of interesting features. It is immediately obvious that the coverage rates of the OLS confidence intervals are very poor and actually have a tendency to decline as span increases. In contrast the coverage rates of confidence intervals based on the spectral estimators are much better and are closer to the nominal values although discrepancies do occur. There are also differences between the different variance estimators used in constructing confidence intervals for the spectral estimators. Those based on the asymptotic distribution (46) tend to increase with N and to decrease with h, while those based on the usual expressions from spectral regression ($\widehat{V}_2$ and $\widehat{V}_{02}$)

tend to increase with h. The former expressions appear to provide more accurate confidence intervals than the latter.

5.2. An empirical example

Provided that a set of variables can be observed sufficiently frequently over a long enough span, it is possible to assess the behaviour of spectral regression estimators in an empirical setting when span and/or frequency are allowed to vary. This example focuses on the long run purchasing power parity (PPP) relationship between the UK and the US, assuming that the relevant variables cointegrate. Defining $P(\tau)$ to be the UK price level, $\Pi(\tau)$ the US price level, and $X(\tau)$ the exchange rate (expressed as dollars per pound sterling), PPP implies that $\Pi(\tau) = X(\tau)P(\tau)$. In circumstances in which the logarithms of the variables are individually integrated processes, the PPP relationship can be recast as the cointegrated system

$$d\ln\Pi(\tau) = -[\ln\Pi(\tau) - \beta_1 \ln X(\tau) - \beta_2 \ln P(\tau)]d\tau + w_1(\tau)d\tau, \quad \tau > 0, \tag{56}$$

$$d\ln X(\tau) = w_{21}(\tau)d\tau, \quad d\ln P(\tau) = w_{22}(\tau)d\tau, \quad \tau > 0, \tag{57}$$

where $\beta_1 = \beta_2 = 1$ is required if PPP holds and the vector $w(\tau) \equiv [w_1(\tau),\ w_{21}(\tau),\ w_{22}(\tau)]'$ is a stationary random disturbance vector. This model is in the form of the cointegrated system (1) and hence the preceeding results apply to the estimation of the parameters β_1 and β_2.

The underlying data used in this example constitute daily exchange rates and monthly producer prices for the UK and the US over the period January 1972 to December 1998. Each of the variables in the model is, in principle, observed as a stock variable, although there may be some averaging involved during the reporting of the price indices. The exchange rate series is, quite clearly, observed as a stock variable, but there remains the question of how to relate it to the monthly series. It is, for example, possible to use a monthly exhange rate series based on the end-month value or on the average daily value throughout the month. The results obtained by Chambers (2000) suggest that the latter, averaged, form of the stock variable yields more efficient estimators (asymptotically) than the end-month value, so it is of interest to compare estimators obtained using both series. In the regressions the span is kept fixed at 27 years but the frequency is allowed to vary between monthly, quarterly and annual, with corresponding sample sizes of 324, 108 and 27 respectively.[11]

The results of the estimations are contained in Tables 4 and 5. Table 4 contains the results using the end-month exchange rate data, while Table 5 contains the results with the

[11] The quarterly and annual data are obtained from the monthly data by skip-sampling in view of the variables being stocks.

monthly-averaged exchange rate data. Nine estimates of β_1 and β_2 are reported based on four types of estimator, as follows. The first estimates reported in each table are based on OLS. The next three are spectral estimates using the estimator $\widehat{C}_n$ constructed with the Parzen kernel and bandwidths equal to the integer parts of $T^{1/10}$, $T^{1/3}$ and $T^{2/5}$. These estimates are denoted SPEC(1/10), SPEC(1/3) and SPEC(2/5), respectively. The next three estimates are the band-limited versions ($\widehat{C}_{n0}$) of the spectral estimators, and are denoted BAND(1/10), BAND(1/3) and BAND(2/5). The final two estimators are the fully modified OLS estimators of Phillips and Hansen (1990), the first constructed using the Parzen kernel, the second using the Bartlett kernel. Both use the automatic bandwidth selection method of Andrews (1991), and are denoted FM-OLS(P) and FM-OLS(B), respectively.

[Tables 4 and 5 about here.]

Inspection of the estimates reported in Table 4 reveals some striking differences between the different estimators. Using the monthly data as an example, the estimates of β_1 range from 0.5346 using the SPEC(2/5) estimator, to 1.0332 using the FM-OLS(B) estimator. The estimates of β_2 show less dispersion, however, ranging from 0.8877 using the FM-OLS(B) estimator, to 0.9458 using the SPEC(1/3) estimator. It is also interesting to note how the spectral estimators vary from the OLS estimator, even though the spectral density estimates that they employ are derived from the OLS residuals. These differences arise because the spectral estimators employ nonparametric corrections to account for serial correlation in the stationary disturbance process that drives the system. For a given estimator, Table 4 reveals that the estimates are, on the whole, remarkably stable across sampling frequencies. Since the same long run parameters are being estimated in each case, this is a reassuring feature.

The estimates reported in Table 5 are obtained with the monthly-averaged exchange rate data. The main differences to emerge, as compared to Table 4, concern the SPEC estimators of β_1, all of which increase in Table 5. The main reason for using the averaged data is to improve (asymptotic) efficiency, and it is interesting to note that the standard errors have dropped in most cases. This provides some finite sample support for the theoretical results concerning asymptotic efficiency obtained in Chambers (2000). As a final point, the estimates in both Tables 4 and 5 suggest that β_1 and β_2 are not equal to unity, as judged by simple t-tests applied to the coefficients separately.[12] The rejections of the implications of PPP are most clear for β_2.

[12]Although not reported here, suffice it to note that Wald tests of the joint hypothesis $\beta_1 = \beta_2 = 1$ have marginal probability values of zero when compared with the chi-square distribution with two degrees of freedom.

6. CONCLUDING COMMENTS

This paper has investigated the effects of sampling frequency, and of data span, on the large-sample asymptotic properties of spectral regression estimators of cointegrating parameters. In cases where span goes to infinity, the limiting distributions are mixed normal, thus enabling conditional normal and chi-square inference to be carried out. When span is fixed but sampling frequency becomes infinite i.e. a continuous record of data is available, the limiting distribution depends on initial conditions and is not necessarily mixed normal. The limiting distributions in the large-span cases reveal that inefficiencies associated with sampling at a fixed interval only affect the parameters associated with stock variables. Put another way, the estimators of parameters associated with flow variables are as efficient when based on a fixed sampling frequency as when based on a continuous record. Simulations reveal that the spectral estimators, in particular the band-limited version, are successful in eradicating the second-order biases inherent in the distribution of the OLS estimator. A limited empirical example is also provided which assesses the performance of the estimators when sampling frequency varies. The estimators are found to be remarkably stable across frequencies although there are significant differences in estimates between different estimators.

On a technical level, the theoretical results derived in this paper have extended those available in the corresponding literature in a number of directions. First, a multivariate system of cointegrated variables has been considered, rather than the typical univariate processes. Secondly, the random forcing process has been allowed to be a fairly general stationary mixing process, thus considerably relaxing the usual assumption of Brownian motion. The resulting invariance principles established here therefore extend those that are currently used in studies of sampling frequency and continuous time processes. Thirdly, to capture the more complicated dynamics that arise because of the previous point, spectral regression estimators have been considered, the analysis of which is more complicated than the OLS estimators that have been considered in the literature so far.

There are a number of ways in which the results in this paper may be extended. It would be possible to consider other estimators that fall within the class of optimal estimators as defined by Phillips (1991b), although many of them require taking a stand on the precise law of motion of the underlying continuous time process $w(\tau)$. The qualitative results to be derived from such exercises are likely to be the same as those obtained here, however. An interesting area of investigation would be to derive the theoretical properties of tests for cointegration when sampling frequency varies. Such research will be helpful in explaining the simulation findings of Hooker (1993), Lahiri and Mamingi (1995) and Otero and Smith (2000). Also of interest would be more extensive empirical applications to assess the effects of sampling frequency more generally. These, and other topics, are ripe for further research.

APPENDIX A

This appendix states (without proof) a number of lemmas that are utilised in the proofs of the results in the main text. A document containing full proofs is available from the author by request or from his website at `http://privatewww.essex.ac.uk/~mchamb`.

Lemma A1. *Let $y(\tau)$ satisfy $dy(\tau) = w(\tau)d\tau$ $(\tau > 0)$ where $w(\tau)$ is a stationary continuous time random process. Then*

$$y(th) - \frac{1}{h}\int_0^h y(th-s)ds = \frac{1}{h}\int_0^h (h-s)w(th-s)ds.$$

Lemma A2. *Let $w(\tau)$ denote a stationary continuous time random process, and let $a(th) = \int_0^h w(th-s)ds$ and $b(th) = \int_0^h \phi(th-s)w(s)ds$, where $\phi(x) = 1 - e^{-x}$. Then*

$$\int_0^h (h-r)a(th-r)dr = \int_0^h \psi_1(r)w(th-r)dr + \int_0^h \psi_2(r)w(th-h-r)dr,$$

$$\int_0^h a(th-r)dr = \int_0^h \psi_3(r)w(th-r)dr + \int_0^h \psi_4(r)w(th-h-r)dr,$$

$$\int_0^h b(th-r)dr = \int_0^h \psi_5(r)w(th-r)dr + \int_0^h \psi_6(r)w(th-h-r)dr,$$

where

$$\psi_1(x) = [h^2 - (x-h)^2]/2, \quad \psi_2(x) = (x-h)^2/2, \quad \psi_3(x) = x,$$
$$\psi_4(x) = h - x, \quad \psi_5(x) = x - \phi(x), \quad \psi_6(x) = h - x - [\phi_h - \phi(x)].$$

Lemma A3. *Let $w(\tau)$ be a stationary continuous time random process. Then*

$$\int_0^h w(th-s)ds = g_h(D)w(th),$$

$$\int_0^h \phi(s)w(th-s)ds = k_h(z)w(th),$$

$$\int_0^h \psi_j(s)w(th-s)ds = \gamma_j(D)w(th), \quad j = 1, \ldots, 6,$$

where $g_h(z) = (1-e^{-hz})/z$, $k_h(z) = g_h(z) - g_h(1+z)$, and

$$\gamma_1(z) = \frac{h^2}{2}g_h(z) - \gamma_2(z), \qquad \gamma_2(z) = \frac{1}{2}\left[\frac{h^2}{z} - \frac{2h}{z^2} + \frac{2(1-e^{-hz})}{z^3}\right],$$
$$\gamma_3(z) = \frac{1}{z}\left[g_h(z) - he^{-hz}\right], \quad \gamma_4(z) = hg_h(z) - \gamma_3(z),$$
$$\gamma_5(z) = \gamma_3(z) - k_h(z), \qquad \gamma_6(z) = \gamma_4(z) - \phi_h g_h(z) + k_h(z).$$

Lemma A4. *Let the functions $g_h(z)$, $k_h(z)$ and $\gamma_j(z)$ be defined as in Lemma A3. Then for each fixed z and as $h \downarrow 0$,*

$$
\begin{aligned}
g_h(z) &= h - \frac{h^2 z}{2} + \frac{h^3 z^2}{6} + O(h^4), \\
k_h(z) &= \frac{h^2}{2} - \frac{h^3}{6}(1+2z) + O(h^4), \\
\gamma_1(z) &= \frac{h^3}{3} + O(h^4), \\
\gamma_2(z) &= \frac{h^3}{6} + O(h^4), \\
\gamma_3(z) &= \frac{h^2 z}{4} - \frac{5h^3 z^2}{12} + O(h^4), \\
\gamma_4(z) &= h^2\left(1 - \frac{z}{4}\right) + h^3\left(\frac{5z^2}{12} - \frac{z}{2}\right) + O(h^4), \\
\gamma_5(z) &= h^2\left(\frac{z}{4} - \frac{1}{2}\right) + h^3\left(\frac{1}{6} + \frac{z}{3} - \frac{5z^2}{12}\right) + O(h^4), \\
\gamma_6(z) &= h^2\left(\frac{1}{2} - \frac{z}{4}\right) + h^3\left(\frac{5z^2}{12} - \frac{z}{3} + \frac{1}{3}\right) + O(h^4).
\end{aligned}
$$

Lemma A5. *The component filters of the matrix filter function $M_h(z)$ defined in Lemma 2 satisfy, as $h \downarrow 0$ for fixed z,*

$$
\begin{aligned}
m_1^S(z) &= h + O(h^2), & m_1^F(z) &= h + O(h^2), \\
m_{12}^{SS}(z) &= \frac{h^2}{2} + O(h^3), & m_{12}^{SF}(z) &= \frac{h^2}{2} + O(h^3), \\
m_{12}^{FS}(z) &= \frac{h^2 z}{4} + O(h^3), & m_{12}^{FF}(z) &= \frac{h^2 z}{4} + O(h^3), \\
m_2^S(z) &= h + O(h^2), & m_2^F(z) &= h + O(h^2).
\end{aligned}
$$

Lemma A6. *The component filters of the matrix filter function $M_h(z)$ defined in Lemma 2 have the following values at $z = 0$ and $z = i\lambda_k$ for $\lambda_k \equiv 2\pi k/h$ and k an integer:*

$$
\begin{array}{llll}
m_1^S(0) = \phi_h, & m_1^S(i\lambda_k) = \mu_{k,h}, & m_1^F(0) = \phi_h, & m_1^F(i\lambda_k) = 0, \\
m_{12}^{SS}(0) = h - \phi_h, & m_{12}^{SS}(i\lambda_k) = -\mu_{k,h}, & & \\
m_{12}^{SF}(0) = h - \phi_h + \nu_h, & m_{12}^{SF}(i\lambda_k) = -\mu_{k,h}, & & \\
m_{12}^{FS}(0) = h - \phi_h - \nu_h, & m_{12}^{FS}(i\lambda_k) = 0, & m_{12}^{FF}(0) = h - \phi_h, & m_{12}^{FF}(i\lambda_k) = 0, \\
m_2^S(0) = h, & m_1^S(i\lambda_k) = 0, & m_2^F(0) = h, & m_2^F(i\lambda_k) = 0,
\end{array}
$$

where $\mu_{k,h} = h\phi_h/(h + 2\pi i k)$ and $\nu_h = h^2\phi_h/2$.

APPENDIX B

Proof of Lemma 1.

Upon replacing τ by th in (2) it can be shown that $y(th)$ satisfies the difference equation $y(th) = e^{-hJA}y(th-h) + \epsilon(th)$, where $\epsilon(th) = \int_0^h e^{-sJA}w(th-s)ds$. Noting that $e^{-hJA} = I_m - JA + e^{-h}JA = I_m - \phi_h JA$, where $\phi_h = 1 - e^{-h}$, this equation may be written

$$\Delta_h y(th) = -\phi_h JAy(th-h) + \epsilon(th). \tag{B1}$$

The decomposition of e^{-hJA} allows $\epsilon(th)$ to be written $\epsilon(th) = \int_0^h [I - \phi(r)JA]w(th-r)dr$, where $\phi(r) = 1 - e^{-r}$, the subvectors of which are

$$\begin{aligned}
\epsilon_1^S(th) &= \int_0^h [1-\phi(r)]w_1^S(th-r)dr + C_{SS}\int_0^h \phi(r)w_2^S(th-r)dr && \text{(B2)}\\
&\quad + C_{SF}\int_0^h \phi(r)w_2^F(th-r)dr, \\
\epsilon_1^F(th) &= \int_0^h [1-\phi(r)]w_1^F(th-r)dr + C_{FS}\int_0^h \phi(r)w_2^S(th-r)dr && \text{(B3)}\\
&\quad + C_{FF}\int_0^h \phi(r)w_2^F(th-r)dr, \\
\epsilon_2^S(th) &= \int_0^h w_2^S(th-r)dr, && \text{(B4)}\\
\epsilon_2^F(th) &= \int_0^h w_2^F(th-r)dr. && \text{(B5)}
\end{aligned}$$

It is convenient to pick out the equations determining the stocks and flows separately from (B1) to give

$$\begin{aligned}
\Delta_h y_1^S(th) &= -\phi_h\left[y_1^S(th-h) - C_{SS}y_2^S(th-h) - C_{SF}y_2^F(th-h)\right] + \epsilon_1^S(th), && \text{(B6)}\\
\Delta_h y_1^F(th) &= -\phi_h\left[y_1^F(th-h) - C_{FS}y_2^S(th-h) - C_{FF}y_2^F(th-h)\right] + \epsilon_1^F(th), && \text{(B7)}\\
\Delta_h y_2^S(th) &= \epsilon_2^S(th), && \text{(B8)}\\
\Delta_h y_2^F(th) &= \epsilon_2^F(th). && \text{(B9)}
\end{aligned}$$

In (B6), note that the variable $y_2^F(th-h)$ on the right-hand side is unobservable, and so adding and subtracting $\phi_h C_{SF} y_{2,th-h}^F$ yields

$$\Delta_h y_{1,th}^S = -\phi_h\left[y_{1,th-h}^S - C_{SS}y_{2,th-h}^S - C_{SF}y_{2,th-h}^F\right] + \xi_{1,th}^S,$$

where the disturbance $\xi_{1,th}^S$ absorbs the transformation involving y_2^F and is given by

$$\xi_{1,th}^S = \epsilon_1^S(th) + \phi_h C_{SF}\left[y_2^F(th-h) - y_{2,th-h}^F\right]$$

$$= \epsilon_1^S(th) + C_{SF}\frac{\phi_h}{h}\int_0^h (h-r)\epsilon_2^F(th-h-r)dr,$$

the second line utilising Lemma A1. The expression for $\xi_{1,th}^S$ in Lemma 1 is then obtained by substituting for $\epsilon_1^S(th)$ using (B2) and for the second term using Lemma A2. In order to transform (B7) into observable variables, it is necessary to first integrate over the interval $[0,h]$ and to divide by h to yield

$$\begin{aligned}\Delta_h y_{1,th}^F &= -\phi_h\left[y_{1,th-h}^F - C_{FS}\frac{1}{h}\int_0^h y_2^S(th-h-s)ds - C_{FF}y_{2,th-h}^F\right] \\ &\quad +\frac{1}{h}\int_0^h \epsilon_1^F(th-s)ds.\end{aligned}$$

The term involving the integral of y_2^S on the right-hand side is unobservable and so adding and subtracting $\phi_h C_{FS} y_{2,th-h}^S$ yields

$$\Delta_h y_{1,th}^F = -\phi_h\left[y_{1,th-h}^F - C_{FS}y_{2,th-h}^S - C_{FF}y_{2,th-h}^F\right] + \xi_{1,th}^F,$$

where

$$\begin{aligned}\xi_{1,th}^F &= \frac{1}{h}\int_0^h \epsilon_1^F(th-s)ds + \phi_h C_{FS}\left[\frac{1}{h}\int_0^h y_2^S(th-h-s)ds - y_2^S(th-h)\right] \\ &= \frac{1}{h}\int_0^h \epsilon_1^F(th-s)ds - C_{FS}\frac{\phi_h}{h}\int_0^h (h-s)\epsilon_2^S(th-h-s)ds,\end{aligned}$$

the second line following from Lemma A1. Substituting for ϵ_1^F using (B3) and then using Lemma A2 on the resulting terms yields the required expression for $\xi_{1,th}^F$. The expressions determining the evolution of y_2^S and y_2^F are easily obtained from (B8) and (B9) giving $\Delta_h y_{2,th}^S = \xi_{2,th}^S$ and $\Delta_h y_{2,th}^F = \xi_{2,th}^F$ with $\xi_{2,th}^S = \epsilon_2^S(th)$ and $\xi_{2,th}^F = h^{-1}\int_0^h \epsilon_2^F(th-s)ds$. These are expressible in terms of w by using (B4) and (B5) while the equation for $\xi_{2,th}^F$ in Lemma 1 also requires the results in Lemma A2. Finally, combining the equations for all the variables yields the discrete time ECM as required. ||

Proof of Lemma 2.

The filtering relationship is obtained from the expressions for the components of ξ_{th} in Lemma 1 using Lemma A3 in Appendix A. The component filters are defined by

$$\begin{aligned}m_1^S(z) &= g_h(1+z), \\ m_{12}^{SS}(z) &= k_h(z), \\ m_{12}^{SF}(z) &= k_h(z) + h^{-1}\phi_h e^{-hz}\left[\gamma_1(z) + e^{-hz}\gamma_2(z)\right],\end{aligned}$$

$$
\begin{aligned}
m_1^F(z) &= h^{-1}[g_h(z) - g_h(1+z)] + h^{-1}e^{-hz}[\phi_h g_h(z) - (g_h(z) - g_h(1+z))], \\
m_{12}^{FS}(z) &= h^{-1}[\gamma_5(z) + e^{-hz}\gamma_6(z) - \phi_h e^{-hz}(\gamma_1(z) + e^{-hz}\gamma_2(z))], \\
m_{12}^{FF}(z) &= h^{-1}[\gamma_5(z) + e^{-hz}\gamma_6(z)], \\
m_2^S(z) &= g_h(z), \\
m_2^F(z) &= h^{-1}[\gamma_3(z) + e^{-hz}\gamma_4(z)],
\end{aligned}
$$

in which $g_h(z) = (1 - e^{-hz})/z$, $k_h(z) = g_h(z) - g_h(1+z)$, and

$$
\begin{aligned}
&\gamma_1(z) = \frac{h^2}{2} g_h(z) - \gamma_2(z), \qquad &&\gamma_2(z) = \frac{1}{2}\left[\frac{h^2}{z} - \frac{2h}{z^2} + \frac{2(1-e^{-hz})}{z^3}\right], \\
&\gamma_3(z) = \frac{1}{z}\left[g_h(z) - he^{-hz}\right], &&\gamma_4(z) = hg_h(z) - \gamma_3(z), \\
&\gamma_5(z) = \gamma_3(z) - k_h(z), &&\gamma_6(z) = \gamma_4(z) - \phi_h g_h(z) + k_h(z).
\end{aligned}
$$

$\|$

Proof of Lemma 3.

Since $g_h(z) = O(h)$ by Lemma A4 it follows that $\zeta_{th} = O_p(h)$ because $w(th)$ is an $O_p(1)$ random variable. Lemma A5 establishes that the upper right $m_1 \times m_2$ block of $M_h(z)$ (corresponding to the response of $\xi_{1,th}$ to $w_2(th)$) and hence of $Q_h(z)$ is $O(h^2)$, and so it remains to show that the diagonal elements of $Q_h(z)$ are also $O(h^2)$. The first m_1^S elements on the diagonal are $m_1^S(z) - g_h(z) = g_h(1+z) - g_h(z) = -k_h(z) = O(h^2)$ by Lemma A4. The next m_1^F elements are $m_1^F(z) - g_h(z) = h + O(h^2) - [h + O(h^2)] = O(h^2)$, using Lemmas A4 and A5. The next m_2^S elements are given by the filter $m_2^S(z) - g_h(z) = 0$ since $m_2^S(z) = g_h(z)$, and hence $\rho_{2,th}^S = 0$ as also stated in the Lemma. The final m_2^F elements on the diagonal equal $m_2^F(z) - g_h(z) = h + O(h^2) - [h + O(h^2)] = O(h^2)$, again using Lemmas A4 and A5. $\|$

Proof of Lemma 4.

(a) The proof follows from Hansen (1992) if the process $w_{nt} = h_n^{-1}\xi_{nt}$, or $w_{th} = h^{-1}\xi_{th}$ (since $h_n = h$), satisfies his Assumption 1, which requires: (i) $Ew_{th} = 0$, (ii) w_{th} is strong mixing with mixing coefficients of size $-\delta\eta/(\delta - \eta)$, (iii) $\|w_{th}\|_\delta < \infty$. Part (i) is trivially satisfied, while (ii) follows from Assumption 2 and Theorem 14.1 of Davidson (1994) because ξ_{th} is a measurable function of $w(\tau)$ over a finite time interval. It remains to verify (iii). Note that

$$
\|w_{th}\|_\delta = \left\|h^{-1}\xi_{th}\right\|_\delta \le h^{-1}\left\{\left\|\xi_{1,th}^S\right\|_\delta + \left\|\xi_{1,th}^F\right\|_\delta + \left\|\xi_{2,th}^S\right\|_\delta + \left\|\xi_{2,th}^F\right\|_\delta\right\}.
$$

Taking each term in turn, using the definitions for the components of ξ_{th} in Lemma 1, and noting that $\| \int_0^h f(r)w(th-r)dr\|_\delta \le |\int_0^h f(r)dr|\|w(\tau)\|_\delta$ for a scalar function $f(\cdot)$,

$$
\begin{aligned}
\left\|\xi^S_{1,th}\right\|_\delta &= \left\| \int_0^h [1-\phi(r)]\, w^S_1(th-r)dr + C_{SS}\int_0^h \phi(r)w^S_2(th-r)dr \right.\\
&\quad + C_{SF}\left[\int_0^h \phi(r)w^F_2(th-r)dr + \frac{\phi_h}{h}\int_0^h \psi_1(r)w^F_2(th-h-r)dr \right.\\
&\quad \left.\left. + \frac{\phi_h}{h}\int_0^h \psi_2(r)w^F_2(th-2h-r)dr\right]\right\|_\delta \\
&\le \left|\int_0^h [1-\phi(r)]dr\right| \left\|w^S_1(\tau)\right\|_\delta + \|C_{SS}\|_\delta \left|\int_0^h \phi(r)dr\right| \left\|w^S_2(\tau)\right\|_\delta \\
&\quad + \|C_{SF}\|_\delta \left\{\left|\int_0^h \phi(r)dr\right| + \frac{|\phi_h|}{h}\left[\left|\int_0^h \psi_1(r)dr\right| + \left|\int_0^h \psi_2(r)dr\right|\right]\right\} \left\|w^F_2(\tau)\right\|_\delta
\end{aligned}
$$

and hence $\|\xi^S_{1,th}\|_\delta < \infty$ in view of the moment condition in Assumption 2 and the finiteness of the integrals of the functions. By a similar procedure,

$$
\begin{aligned}
\left\|\xi^F_{1,th}\right\|_\delta &\le h^{-1}\left\{\left|\int_0^h \phi(r)dr\right| + \left|\int_0^h [\phi_h-\phi(r)]\,dr\right|\right\}\left\|w^F_1(\tau)\right\|_\delta + h^{-1}\|C_{FS}\|_\delta \\
&\quad \times\left\{\left|\int_0^h \psi_5(r)dr\right| + \left|\int_0^h \psi_6(r)dr\right| + |\phi_h|\left[\left|\int_0^h \psi_1(r)dr\right| + \left|\int_0^h \psi_2(r)dr\right|\right]\right\} \\
&\quad \times\left\|w^S_2(\tau)\right\|_\delta + h^{-1}\|C_{FF}\|_\delta\left\{\left|\int_0^h \psi_5(r)dr\right| + \left|\int_0^h \psi_6(r)dr\right|\right\}\left\|w^F_2(\tau)\right\|_\delta
\end{aligned}
$$

which is also finite under Assumption 2 and the finiteness of the integrals. In the same way,

$$
\left\|\xi^S_{2,th}\right\|_\delta = \left\|\int_0^h w^S_2(th-s)ds\right\|_\delta \le h\left\|w^S_2(\tau)\right\|_\delta < \infty,
$$

$$
\left\|\xi^F_{2,th}\right\|_\delta \le h^{-1}\left\{\left|\int_0^h \psi_3(r)dr\right| + \left|\int_0^h \psi_4(r)dr\right|\right\}\left\|w^F_2(\tau)\right\|_\delta < \infty.
$$

Hence condition (iii) is satisfied and part (a) of the Lemma follows.

(b) Note that, from Lemma 3,

$$
\frac{h_n^{1/2}}{T_n^{1/2}} S_{n[T_n r]} = \frac{1}{N_n^{1/2}}\sum_{j=1}^{[T_n r]}\xi_{nj} = \frac{1}{N_n^{1/2}}\sum_{j=1}^{[T_n r]}\zeta_{nj} + \frac{1}{N_n^{1/2}}\sum_{j=1}^{[T_n r]}\rho_{nj}. \qquad \text{(B10)}
$$

The proof proceeds by first showing that $N_n^{-1/2}\sum_{j=1}^{[T_n r]}\rho_{nj}$ converges to zero in probability uniformly in r i.e. that $\sup_{r\in[0,1]}\|N_n^{-1/2}\sum_{j=1}^{[T_n r]}\rho_{nj}\| \xrightarrow{p} 0$ as $n\uparrow\infty$. Now

$$
\begin{aligned}
\sup_{r\in[0,1]}\left\|N_n^{-1/2}\sum_{j=1}^{[T_n r]}\rho_{nj}\right\| &\le \max_{1\le k\le T_n}\left\|N_n^{-1/2}\sum_{j=1}^{k}\rho_{nj}\right\| \\
&\le N_n^{-1/2}\max_{1\le k\le T_n}\sum_{j=1}^{k}\|\rho_{nj}\| \le N_n^{-1/2}T_n\max_{1\le j\le T_n}\|\rho_{nj}\|. \qquad \text{(B11)}
\end{aligned}
$$

It follows that, for some $\epsilon > 0$,

$$\begin{aligned}
\Pr\left(\sup_{r\in[0,1]}\left\|N_n^{-1/2}\sum_{j=1}^{[T_nr]}\rho_{nj}\right\| > \epsilon\right) &\leq \Pr\left(N_n^{-1/2}T_n\max_{1\leq j\leq T_n}\|\rho_{nj}\| > \epsilon\right) \\
&= \Pr\left(\|\rho_{n1}\| > \epsilon N_n^{1/2}T_n^{-1}\right) \quad \text{by stationarity} \\
&\leq \frac{E\|\rho_{n1}\|^2}{\epsilon^2 N_n T_n^{-2}}, \qquad \text{(B12)}
\end{aligned}$$

the last line using the Markov inequality. Now $E\|\rho_{n1}\|^2 = \sum_{j=1}^m E\rho_{n1,j}^2 = O(h_n^4)$ by Lemma 3. Furthermore, $N_nT_n^{-2} = h_n^2N_n^{-1}$ and so the right-hand-side of (B12) is $O(h_n^2N_n) = o(1)$ as $n\uparrow\infty$ since $h_nN_n\downarrow 0$. Hence it is the term involving the partial sum of the ζ_{nj} in (B11) that determines the asymptotic distribution of interest.

Let $x_j = \int_{j-1}^j w(s)ds$. Clearly $Ex_j = 0$ and $\|x_j\|_\delta \leq \|w(1)\|_\delta < \infty$ under Assumption 2. In fact, x_j also has the same mixing properties as $w(\tau)$, and so, from Hansen (1992),

$$N_n^{-1/2}\sum_{j=1}^{[N_nr]} x_j \Rightarrow B(r)$$

as $n\uparrow\infty$, where $B(r)$ is Brownian motion with variance matrix

$$\begin{aligned}
\Omega = \lim_{n\uparrow\infty} N_n^{-1}E\sum_{j=1}^{N_n}x_j\sum_{k=1}^{N_n}x_k' &= \lim_{n\uparrow\infty}\int_{-N_n}^{N_n}\left(1-\frac{|k|}{N_n}\right)Ew(0)w(k)'dk \\
&= \int_{-\infty}^{\infty}Ew(0)w(k)'dk = 2\pi f_{ww}^c(0).
\end{aligned}$$

If it can be shown that $N_n^{-1/2}\sum_{j=1}^{T_nr}\zeta_{nj}$ converges in probability uniformly in r to the partial sum $N_n^{-1/2}\sum_{j=1}^{[N_nr]}x_j$, then the claim in part (b) concerning $S_{n[T_nr]}$ is established. Now,

$$\begin{aligned}
&\sup_{r\in[0,1]}\left\|N_n^{-1/2}\sum_{j=1}^{T_nr}\zeta_{nj} - N_n^{-1/2}\sum_{j=1}^{[N_nr]}x_j\right\| \\
&= \sup_{r\in[0,1]}\left\|N_n^{-1/2}\left(\int_0^{[T_nr]h_n}w(s)ds - \int_0^{[N_nr]}w(s)ds\right)\right\| \\
&= \sup_{r\in[0,1]}\left\|N_n^{-1/2}\int_{[N_nr]}^{[T_nr]h_n}w(s)ds\right\| \quad \text{since } [T_nr]h_n \geq [N_nr] \\
&\leq \sup_{r\in[0,1]} N_n^{-1/2}\int_{[N_nr]}^{[T_nr]h_n}\|w(s)\|ds \\
&\leq N_n^{-1/2}\sup_{r\in[0,1]}\int_{[N_nr]}^{N_nr}\|w(s)\|ds \quad \text{since } [T_nr]h_n \leq N_nr \\
&\leq N_n^{-1/2}\max_{1\leq j\leq N_n}\int_{j-1}^{j}\|w(s)\|ds. \qquad \text{(B13)}
\end{aligned}$$

Thus

$$\begin{aligned}
&\Pr\left(\sup_{r\in[0,1]}\left\|N_n^{-1/2}\sum_{j=1}^{T_n r}\zeta_{nj}-N_n^{-1/2}\sum_{j=1}^{[N_n r]}x_j\right\|>\epsilon\right)\\
&\le \Pr\left(N_n^{-1/2}\max_{1\le j\le N_n}\int_{j-1}^{j}\|w(s)\|ds>\epsilon\right)\\
&= \Pr\left(\int_0^1\|w(s)\|ds>\epsilon N_n^{1/2}\right)\quad\text{by stationarity}\\
&\le \Pr\left(\sup_{s\in[0,1]}\|w(s)\|>\epsilon N_n^{1/2}\right)\\
&= \Pr\left(\|w(1)\|>\epsilon N_n^{1/2}\right)\quad\text{by stationarity}\\
&\le \frac{E\|w(1)\|^2}{\epsilon^2 N_n}\downarrow 0\quad\text{as}\quad n\uparrow\infty \qquad\qquad (B14)
\end{aligned}$$

since $w(1)$ is $O_p(1)$ and $N_n\uparrow\infty$, thus establishing the result.

Turning to U_{nT_n}, consider (noting that $h_n^{1/2}T_n^{-1/2}=h_nN_n^{-1/2}$)

$$\begin{aligned}
\frac{h_n^2}{N_n}U_{nT_n}=N_n^{-1}\sum_{t=1}^{T_n}\left(\sum_{j=1}^{t-1}\xi_{nj}\right)\xi_{nt}' &= N_n^{-1}\sum_{t=1}^{T_n}\left[\sum_{j=1}^{t-1}(\zeta_{nj}+\rho_{nj})\right](\zeta_{nt}+\rho_{nt})'\\
&= U_{n,\zeta\zeta}+U_{n,\zeta\rho}+U_{n,\rho\zeta}+U_{n,\rho\rho}, \qquad (B15)
\end{aligned}$$

where, for example, $U_{n,\zeta\rho}=N_n^{-1}\sum_{t=1}^{T_n}(\sum_{j=1}^{t-1}\zeta_{nj})\rho_{nt}'$. Note, first, that

$$\begin{aligned}
U_{n,\zeta\zeta} &= N_n^{-1}\sum_{t=1}^{T_n}\left(\sum_{j=1}^{t-1}\int_{jh_n-h_n}^{jh_n}w(s)ds\right)\int_{th_n-h_n}^{th_n}w(r)'dr\\
&= N_n^{-1}\sum_{t=1}^{T_n}\left(\int_0^{th_n-h_n}w(s)ds\right)\int_{th_n-h_n}^{th_n}w(r)'dr\\
&= N_n^{-1}\int_0^{N_n}\left(\int_0^r w(s)ds\right)w(r)'dr. \qquad (B16)
\end{aligned}$$

Now consider

$$\begin{aligned}
U_{n,xx}=N_n^{-1}\sum_{t=1}^{N_n}\left(\sum_{j=1}^{t-1}x_{nj}\right)x_{nt}' &= N_n^{-1}\sum_{t=1}^{N_n}\left(\int_0^{t-1}w(s)ds\right)\int_{t-1}^{t}w(r)'dr\\
&= N_n^{-1}\int_0^{N_n}\left(\int_0^r w(s)ds\right)w(r)'dr. \qquad (B17)
\end{aligned}$$

Since x_{nt} satisfies the assumptions of Hansen (1992), it follows that $U_{n,xx}\Rightarrow\int_0^1 BdB'+\Lambda_1$. That $U_{n,\zeta\zeta}$ also has the same asymptotic distribution follows from (B16). For the remaining terms in (B15), consider, first,

$$\|U_{n,\zeta\rho}\|=\left\|N_n^{-1}\sum_{t=1}^{T_n}\left(\sum_{j=1}^{t-1}\zeta_{nj}\right)\rho_{nt}'\right\| \le N_n^{-1}\sum_{t=1}^{T_n}\left\|\sum_{j=1}^{t-1}\zeta_{nj}\right\|\|\rho_{nt}\|$$

$$\leq N_n^{-1} \sum_{t=1}^{T_n} \|\rho_{nt}\| \sum_{j=1}^{T_n} \|\zeta_{nj}\|$$
$$\leq N_n^{-1} T_n^2 \max_{1\leq t\leq T_n} \|\rho_{nt}\| \max_{1\leq j\leq T_n} \|\zeta_{nj}\| .$$

Hence, for some $\epsilon > 0$,

$$\begin{aligned}
\Pr(\|U_{n,\zeta,\rho}\| > \epsilon) &\leq \Pr\left(N_n^{-1} T_n^2 \max_{1\leq t\leq T_n} \|\rho_{nt}\| \max_{1\leq j\leq T_n} \|\zeta_{nj}\| > \epsilon\right) \\
&\leq \Pr\left(\|\rho_{n1}\|\|\zeta_{n1}\| > \epsilon N_n T_n^{-2}\right) \quad \text{by stationarity} \\
&\leq \frac{E\left(\|\rho_{n1}\|\|\zeta_{n1}\|\right)^2}{\epsilon^2 N_n^2 T_n^{-4}} \quad \text{by Markov's inequality} \\
&\leq \frac{E\|\rho_{n1}\|^2 E\|\zeta_{n1}\|^2}{\epsilon^2 N_n^2 T_n^{-4}} \quad \text{by the Cauchy-Schwarz inequality.} \qquad \text{(B18)}
\end{aligned}$$

From Lemma 3, $E\|\rho_{n1}\|^2 = O(h_n^4)$ and $E\|\zeta_{n1}\|^2 = O(h_n^2)$, while $N_n^2 T_n^{-4} = N_n^2 (N_n/h_n)^{-4} = h_n^4 N_n^{-2}$, so that the right-hand side of (B18) is $O(h_n^2 N_n^2) = o(1)$ as $n \uparrow \infty$, since $h_n N_n \downarrow 0$. Hence $\|U_{n,\zeta\rho}\| = o_p(1)$ as $n \uparrow \infty$. Similar arguments can be used to show that

$$\Pr(\|U_{n,\rho\zeta}\| > \epsilon) \leq \frac{E\|\rho_{n1}\|^2 E\|\zeta_{n1}\|^2}{\epsilon^2 N_n^2 T_n^{-4}} = O(h_n^2 N_n^2),$$

$$\Pr(\|U_{n,\rho\rho}\| > \epsilon) \leq \frac{\left(E\|\rho_{n1}\|^2\right)^2}{\epsilon^2 N_n^2 T_n^{-4}} = O(h_n^4 N_n^2),$$

and hence $\|U_{n,\rho\zeta}\| = o_p(1)$ and $\|U_{n,\rho\rho}\| = o_p(1)$ as $n \uparrow \infty$. Thus part (b) is proved.

(c) In this case, $h_n S_{n[T_n r]} = \sum_{t=1}^{[T_n r]} \xi_{nt} = \sum_{t=1}^{[T_n r]} \zeta_{nt} + \sum_{t=1}^{[T_n r]} \rho_{nt}$. From (B11) and (B12), and noting that $N_n = N$, it follows that $\sup_{r\in[0,1]} \| \sum_{t=1}^{[T_n r]} \rho_{nt}\| = o_p(1)$. Now

$$\begin{aligned}
\sup_{r\in[0,1]} \left\| \int_0^{Nr} w(s)ds - \sum_{t=1}^{[T_n r]} \zeta_{nt} \right\| &= \sup_{r\in[0,1]} \left\| \int_0^{Nr} w(s)ds - \int_0^{[T_n r]h_n} w(s)ds \right\| \\
&= \sup_{r\in[0,1]} \left\| \int_{[T_n r]h_n}^{Nr} w(s)ds \right\| \\
&\leq \sup_{r\in[0,1]} \int_{[T_n r]h_n}^{Nr} \|w(s)\|ds \xrightarrow{p} 0
\end{aligned}$$

since

$$Nr - [T_n r]h_n = h_n\left(\frac{Nr}{h_n} - [T_n r]\right) = h_n\left(\frac{Nr}{h_n} - \left[\frac{Nr}{h_n}\right]\right) \leq h_n \downarrow 0$$

as $n \uparrow \infty$. Hence $\sup_{r\in[0,1]} \|h_n S_{n[T_n r]} - \int_0^{Nr} w(s)ds\| \xrightarrow{p} 0$ as $n \uparrow \infty$ and so $h_n S_{n[T_n r]} \Rightarrow Z(Nr)$ as required. Finally, consider

$$h_n^2 U_{nT_n} = N(U_{n,\zeta\zeta} + U_{n,\zeta\rho} + U_{n,\rho\zeta} + U_{n,\rho\rho}).$$

From the analysis in part (b), each of the last three terms is $o_p(1)$, and so

$$\begin{aligned} h_n^2 U_{nT_n} &= \sum_{t=1}^{T_n} \left(\sum_{j=1}^{t-1} \zeta_{nj} \right) \zeta_{nt}' + o_p(1) \\ &= \sum_{t=1}^{T_n} \left(\int_0^{th_n - h_n} w(s) ds \right) \int_{th_n - h_n}^{th_n} w(r)' dr + o_p(1) \\ &= \sum_{t=1}^{T_n} Z(th_n - h_n) \left[Z(th_n) - Z(th_n - h_n) \right]' + o_p(1) \\ &\Rightarrow \int_0^N Z(s) dZ(s)' \end{aligned}$$

as $n \uparrow \infty$, as required. ‖

Proof of Lemma 5.

The proof for each part is based on the expressions for $\mu_n(s)$ and $M_n(s)$ in (22) and (23) respectively, combined with the appropriate convergence rates for $S_{n[T_n r]}$ and U_{nT_n} given in Lemma 4. ‖

APPENDIX C

Proof of Theorem 1.

(a) This is a straightforward extension of the results of Phillips (1991a,c) to the case where the sampling frequency $h \neq 1$, and so the details are omitted. The proof can also be based on the results presented in (b) below provided the appropriate modifications are made.

(b) For convenience of notation let $a_n = h_n^2 \phi_{h_n}^{-2} N_n^{-1}$ and $b_n = h_n/\phi_{h_n}$ so that $a_n \widehat{\Gamma}_{n,XX}(s)$ and $b_n \widehat{\Gamma}_{n,wX}(s)$ converge to the limits given in Lemma 5. Define $\widehat{Q}_n = (1/2M_n) \sum_{j=-M_n+1}^{M_n} \Theta_{nj}$ and $\widehat{q}_n = (1/2M_n) \sum_{j=-M_n+1}^{M_n} \theta_{nj}$ so that $\mathrm{vec}(\widehat{C}_n - C_0) = \widehat{Q}_n^{-1} \widehat{q}_n$. Taking the component $\widehat{Q}_n$ first, let $\widehat{Q}_n = Q_n + R_n$, where

$$Q_n = \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left(\widehat{f}'_{n,XX} \otimes J' f_{ww}^{-1} J \right), \tag{C1}$$

$$R_n = \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left(\widehat{f}'_{n,XX} \otimes J' \left[\widehat{f}_{n,\widehat{w}\widehat{w}}^{-1} - f_{ww}^{-1} \right] J \right), \tag{C2}$$

and where the dependence of the spectral density matrices on frequency ω_j has been suppressed for ease of notation. The first step is to show that $\|(a_n/h_n) R_n\| = o_p(1)$. Consider

$$\begin{aligned}
\left\| \frac{a_n}{h_n} R_n \right\| &= \left\| \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left(\frac{a_n}{h_n} \widehat{f}'_{n,XX} \otimes J' \widehat{f}_{n,\widehat{w}\widehat{w}}^{-1} \left[f_{ww} - \widehat{f}_{n,\widehat{w}\widehat{w}} \right] f_{ww}^{-1} J \right) \right\| \\
&\leq \frac{\|J\|^2}{2M_n} \sum_{j=-M_n+1}^{M_n} \left\| \frac{a_n}{h_n} \widehat{f}_{n,XX} \right\| \left\| \widehat{f}_{n,\widehat{w}\widehat{w}}^{-1} \right\| \left\| f_{ww} - \widehat{f}_{n,\widehat{w}\widehat{w}} \right\| \left\| f_{ww}^{-1} \right\| \\
&\leq \max_{\lambda \in \Pi_{h_n}} \left\| \widehat{f}_{n,\widehat{w}\widehat{w}}^{-1} \right\| \max_{\lambda \in \Pi_{h_n}} \left\| f_{ww} - \widehat{f}_{n,\widehat{w}\widehat{w}} \right\| \max_{\lambda \in \Pi_{h_n}} \left\| f_{ww}^{-1} \right\| \frac{\|J\|^2}{2M_n} \sum_{j=-M_n+1}^{M_n} \left\| \frac{a_n}{h_n} \widehat{f}_{n,XX} \right\|
\end{aligned} \tag{C3}$$

where $\Pi_{h_n} = \{\lambda : -\pi/h_n < \lambda \leq \pi/h_n\}$. Now, as $n \uparrow \infty$,

$$\max_{\lambda \in \Pi_{h_n}} \left\| f_{ww}^{-1} \right\| \to \max_{\lambda \in \Pi_0} \left\| (f_{ww}^c)^{-1} \right\| \leq K,$$

by the assumed properties of $f_{ww}^c(\lambda)$. Meanwhile, outside a set $\widetilde{\Pi}_n$ whose probability measure tends to zero as $n \uparrow \infty$, $\left\| \widehat{f}_{n,\widehat{w}\widehat{w}} \right\| \geq a > 0$ and so

$$\max_{\lambda \in \Pi_{h_n}} \left\| \widehat{f}_{n,\widehat{w}\widehat{w}}^{-1} \right\| \to \max_{\lambda \in \Pi_0} \left\| (f_{ww}^c)^{-1} \right\| \leq K$$

as $n \uparrow \infty$. Furthermore,

$$\max_{\lambda \in \Pi_{h_n}} \left\| \widehat{f}_{n,\widehat{w}\widehat{w}} - f_{ww} \right\| \leq \max_{\lambda \in \Pi_{h_n}} \left\| \widehat{f}_{n,\widehat{w}\widehat{w}} - f_{\widehat{w}\widehat{w}} \right\| + \max_{\lambda \in \Pi_{h_n}} \left\| f_{\widehat{w}\widehat{w}} - f_{ww} \right\|. \tag{C4}$$

Taking the first component,

$$\begin{aligned}
&\max_\lambda \left\| \widehat{f}_{n,\widehat{w}\widehat{w}} - f_{\widehat{w}\widehat{w}} \right\| \\
&= \max_\lambda \left\| \frac{h_n}{2\pi} \sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) \widehat{\Gamma}_{n,\widehat{w}\widehat{w}}(s) e^{-ish_n\lambda} - \frac{h_n}{2\pi} \sum_{s=-\infty}^{\infty} \Gamma_{\widehat{w}\widehat{w}}(s) e^{-ish_n\lambda} \right\| \\
&= \max_\lambda \left\| \frac{h_n}{2\pi} \sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) \left[\widehat{\Gamma}_{n,\widehat{w}\widehat{w}}(s) - \Gamma_{\widehat{w}\widehat{w}}(s)\right] e^{-ish_n\lambda} \right. \\
&\qquad \left. + \frac{h_n}{2\pi} \sum_{s=-\infty}^{\infty} \left[k\left(\frac{s}{M_n}\right) - 1\right] \Gamma_{\widehat{w}\widehat{w}}(s) e^{-ish_n\lambda} \right\| \\
&\le \frac{h_n}{2\pi} \max_{s,\lambda} \left| e^{-ish_n\lambda} \right| \sum_{s=-M_n}^{M_n} \left| k\left(\frac{s}{M_n}\right) \right| \left\| \widehat{\Gamma}_{n,\widehat{w}\widehat{w}}(s) - \Gamma_{\widehat{w}\widehat{w}}(s) \right\| \\
&\quad + \frac{h_n}{2\pi} \max_{s,\lambda} \left| e^{-ish_n\lambda} \right| \sum_{s=-\infty}^{\infty} \left| k\left(\frac{s}{M_n}\right) - 1 \right| \left\| \Gamma_{\widehat{w}\widehat{w}}(s) \right\|. \qquad \text{(C5)}
\end{aligned}$$

Now $\max_{s,\lambda} \left| e^{-ish_n\lambda} \right| = 1$, $\left\| \widehat{\Gamma}_{n,\widehat{w}\widehat{w}}(s) - \Gamma_{\widehat{w}\widehat{w}}(s) \right\| \xrightarrow{p} 0$, and $|k(s/M_n) - 1| \to 0$ for each fixed s as $M_n \uparrow \infty$, and hence (C5)$\xrightarrow{p} 0$ as $n \uparrow \infty$. For the second component of (C4),

$$\begin{aligned}
\max_\lambda \left\| f_{\widehat{w}\widehat{w}} - f_{ww} \right\| &= \max_\lambda \left\| \frac{h_n}{2\pi} \sum_{s=-\infty}^{\infty} \left[\Gamma_{\widehat{w}\widehat{w}}(s) - \Gamma_{ww}(s)\right] e^{-ish_n\lambda} \right\| \\
&\le \frac{h_n}{2\pi} \sum_{s=-\infty}^{\infty} \left\| \Gamma_{\widehat{w}\widehat{w}}(s) - \Gamma_{ww}(s) \right\| \xrightarrow{p} 0
\end{aligned}$$

as $n \uparrow \infty$, since $h_n \downarrow 0$ and $\widehat{w}_n \xrightarrow{p} w_n$. Hence (C4)$\xrightarrow{p} 0$. Now consider

$$\begin{aligned}
\frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left\| \frac{a_n}{h_n} \widehat{f}_{n,XX} \right\| &= \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left\| \frac{h_n}{2\pi} \sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) \frac{a_n}{h_n} \widehat{\Gamma}_{n,XX}(s) e^{-ish_n\omega_j} \right\| \\
&\le \frac{1}{2M_n} \frac{1}{2\pi} \sum_{j=-M_n+1}^{M_n} \sum_{s=-M_n}^{M_n} \left| k\left(\frac{s}{M_n}\right) \right| \left\| a_n \widehat{\Gamma}_{n,XX}(s) \right\| \left| e^{-is\pi j/M_n} \right| \\
&= \frac{1}{2\pi} \sum_{s=-M_n}^{M_n} \left| k\left(\frac{s}{M_n}\right) \right| \left\| a_n \widehat{\Gamma}_{n,XX}(s) \right\| \quad \text{since} \quad \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} \left| e^{-is\pi j/M_n} \right| = 1 \quad \forall s \\
&\Rightarrow \frac{1}{2\pi} \int_{-1}^{1} |k(r)|\, dr \times O_p(1)
\end{aligned}$$

using Lemma 5(b), where the $O_p(1)$ limit is independent of s. Combining these results in (C3) establishes that $\|(a_n/h_n)R_n\| = o_p(1)$ as required, and hence $(a_n/h_n)\widehat{Q}_n = (a_n/h_n)Q_n + o_p(1)$. Now, from (C1),

$$\begin{aligned}
\frac{a_n}{h_n} Q_n &= \frac{1}{2M_n} \sum_j \left[\frac{h_n}{2\pi} \sum_s k\left(\frac{s}{M_n}\right) \frac{a_n}{h_n} \widehat{\Gamma}_{n,XX}(s)' e^{-ish_n\omega_j} \otimes J' \frac{h_n}{2\pi} \sum_g D_{g,n} e^{igh_n\omega_j} J \right] \\
&= \frac{h_n}{(2\pi)^2} \sum_{g=-\infty}^{\infty} \sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) a_n \widehat{\Gamma}_{n,XX}(s)' \otimes J' D_{g,n} J \frac{1}{2M_n} \sum_{j=-M_n+1}^{M_n} e^{i(g-s)h_n\omega_j},
\end{aligned}$$

which utilises the Fourier series $f_{ww}(\lambda)^{-1} = (h_n/2\pi)\sum_{g=-\infty}^{\infty} D_{g,n}e^{igh_n\lambda}$. But

$$\frac{1}{2M_n}\sum_{j=-M_n+1}^{M_n} e^{i(g-s)h_n\omega_j} = \frac{1}{2M_n}\sum_j e^{i(g-s)\pi j/M_n} = \begin{cases} 1, & (g-s) = 2lM_n, \\ 0, & \text{otherwise}, \end{cases}$$

where l denotes a (positive or negative) integer. Substituting $s = g + 2lM_n$ $(l = 0, \pm 1, \ldots)$ in the above yields

$$\frac{a_n}{h_n}Q_n = \frac{h_n}{(2\pi)^2}\sum_g\sum_l k\left(\frac{g+2lM_n}{M_n}\right) a_n\widehat{\Gamma}_{n,XX}(g+2lM_n)' \otimes J'D_{g,n}J = Q_{0n} + Q_{1n},$$

the first term corresponding to $l = 0$, the second to the sum over $l \neq 0$. Now, since $k(g/M_n) \to 1$ for all g as $M_n \uparrow \infty$,

$$Q_{0n} = \frac{h_n}{(2\pi)^2}\sum_g k\left(\frac{g}{M_n}\right) a_n\widehat{\Gamma}_{n,XX}(g)' \otimes J'D_{g,n}J \Rightarrow \int_0^1 B_2B_2' \otimes J'\Omega^{-1}J, \tag{C6}$$

where

$$\Omega^{-1} = \lim_{n\uparrow\infty}\frac{h_n}{(2\pi)^2}\sum_g D_{g,n} = \frac{1}{2\pi}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} D(v)dv\right] = \frac{1}{2\pi}f_{ww}^c(0)^{-1}.$$

The second term to consider is

$$Q_{1n} = \frac{h_n}{(2\pi)^2}\sum_g\sum_{l\neq 0} k\left(\frac{g+2lM_n}{M_n}\right) a_n\widehat{\Gamma}_{n,XX}(g+2lM_n)' \otimes J'D_{g,n}J.$$

For each g, $k((g+2lM_n)/M_n) \to k(2l)$ as $M_n \uparrow \infty$, but $k(2l) = 0$ for $l \neq 0$. Hence $Q_{1n} = o_p(1)$, and the limit in (C6) is the limit of $(a_n/h_n)\widehat{Q}_n$.

A similar procedure can be applied to the component $\widehat{q}_n$, yielding the decomposition $(b_n/h_n)\widehat{q}_n = (b_n/h_n)q_n + (b_n/h_n)r_n = (b_n/h_n)q_n + o_p(1)$, while $(b_n/h_n)q_n = q_{0n} + q_{1n} = q_{0n} + o_p(1)$, where the important term is

$$q_{0n} = \frac{h_n}{(2\pi)^2}\sum_g (I \otimes J'D_{g,n})\,\mathrm{vec}\left[k\left(\frac{g}{M_n}\right) b_n\widehat{\Gamma}_{n,wX}(g)\right]. \tag{C7}$$

From the convergence of $b_n\widehat{\Gamma}_{n,wX}(g)$ given in Lemma 5(b), it follows that

$$q_{0n} \Rightarrow \left(I \otimes J'\Omega^{-1}\right)\mathrm{vec}\left(\int_0^1 dBB_2'\right) + \bar{\theta},$$

where

$$\begin{aligned}\bar{\theta} &= \lim_{n\uparrow\infty}\frac{h_n}{(2\pi)^2}\sum_g (I \otimes J'D_{g,n})\,\mathrm{vec}\left(h_n\sum_{k=g+1}^{T_n} Ew_{nk}w_{2,n0}'\right) \\ &= \left(\frac{1}{2\pi}\right)^2\int_{-\infty}^{\infty}(I \otimes J'D(v))\,\mathrm{vec}\left(\int_0^{\infty} Ew(s+v)w_2(0)'ds\right)dv, \end{aligned} \tag{C8}$$

the second expression transforming gh_n into the continuous variable v and kh_n into the continuous variable s. Let $\Lambda_2(v) = \int_0^\infty Ew_2(0)w(s+v)'ds$, and note that

$$\int_{-\infty}^{\infty} (I \otimes J'D(v)) \operatorname{vec}(\Lambda(v)')\, dv = \operatorname{vec}\left(J' \int_{-\infty}^{\infty} D(v)\Lambda_2(v)'dv\right). \tag{C9}$$

Defining $z(s) = \int_{-\infty}^{\infty} D(v)w(s+v)dv$ it is possible to write

$$\int_{-\infty}^{\infty} D(v)\Lambda_2(v)'dv = \int_{-\infty}^{\infty} D(v)\left[\int_0^\infty Ew(s+v)w_2(0)'ds\right]dv = \int_0^\infty Ez(s)w_2(0)'ds. \tag{C10}$$

Let $f_{zw_2}^c(\lambda)$ $(-\infty < \lambda < \infty)$ denote the spectral density function between z and w_2. Then $Ez(s)w_2(0)' = \int_{-\infty}^{\infty} e^{is\lambda} f_{zw_2}^c(\lambda)d\lambda$ so that

$$J' \int_{-\infty}^{\infty} D(v)\Lambda_2(v)'dv = J' \int_0^\infty Ez(s)w_2(0)'ds = \int_0^\infty \left[\int_{-\infty}^{\infty} e^{is\lambda} J' f_{zw_2}^c(\lambda)d\lambda\right] ds. \tag{C11}$$

Combining these results,

$$\begin{aligned}
f_{zw_2}^c(\lambda) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-is\lambda} Ez(s)w_2(0)'ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-is\lambda} E\left[\int_{-\infty}^{\infty} D(v)w(s+v)dv\right] w_2(0)'ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} D(v)e^{iv\lambda}dv \int_{-\infty}^{\infty} Ew(m)w_2(0)'e^{-im\lambda}dm \quad (m = s+v) \\
&= f_{ww}^c(\lambda)^{-1} 2\pi f_{ww_2}^c(\lambda) \\
&= 2\pi f_{ww}^c(\lambda)^{-1} f_{ww}^c(\lambda)\begin{bmatrix} 0 \\ I \end{bmatrix} = 2\pi \begin{bmatrix} 0 \\ I \end{bmatrix},
\end{aligned} \tag{C12}$$

which implies that

$$J' f_{zw_2}^c(\lambda) = [I\ 0] 2\pi \begin{bmatrix} 0 \\ I \end{bmatrix} = 0.$$

Substituting into (C11) shows that (C11) is null which implies that $\bar{\theta}$ is also null by (C9). Hence q_{0n} converges to the first term in the displayed expression following (C7), and hence $(b_n/h_n)\widehat{q}_n$ has the same limit by the arguments provided earlier.

Turning to Θ_{n0} and θ_{n0}, it is legitimate, for the reasons advanced earlier (i.e. that $\max_\lambda \|\widehat{f}_{n,\widehat{w}\widehat{w}} - f_{ww}\| \xrightarrow{p} 0$), to replace $\widehat{f}_{n,\widehat{w}\widehat{w}}(0)^{-1}$ with $f_{ww}(0)^{-1}$ in their definitions. Using Lemma 5, it then follows that

$$\frac{a_n}{h_n M_n}\widehat{f}_{n,XX}(0) = \frac{1}{2\pi M_n}\sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) a_n \widehat{\Gamma}_{n,XX}(s) \Rightarrow \nu \int_0^1 B_2 B_2',$$

where $\nu = \lim_{n\uparrow\infty} M_n^{-1}\sum_{s=-M_n}^{M_n} k(s/M_n) = (2\pi)^{-1}\int_{-1}^1 k(s)ds$. The expression (36) then follows from the definition of Θ_{n0}. Proceeding in a similar fashion,

$$\frac{b_n}{h_n M_n}\widehat{f}_{n,wX}(0) = \frac{1}{2\pi M_n}\sum_{s=-M_n}^{M_n} k\left(\frac{s}{M_n}\right) b_n \widehat{\Gamma}_{n,wX}(s) \Rightarrow \nu\int_0^1 dBB_2' + \frac{1}{2\pi}\Delta_2',$$

where $\Delta_2 = \lim_{n\uparrow\infty} h_n \sum_{k=-\infty}^{\infty} Ew_{2,n0}w_{nk}'$. The result for θ_{n0} stated in Theorem 1 follows because it can be shown that $(I\otimes J' f_{ww}(0)^{-1})\text{vec}((1/2\pi)\Delta_2') = 0$, which follows along the lines of the proof on p.433 of Phillips (1991c).
(c) Part (c) follows by the same arguments used to establish part (b), with the appropriate limits from Lemma 5(c) used where appropriate. ||

Proof of Theorem 2.

First, note that

$$J'\Omega^{-1} = [I,\ 0]\begin{bmatrix} \Omega_{11.2}^{-1} & -\Omega_{11.2}^{-1}\Omega_{12}\Omega_{22}^{-1} \\ \cdot & \cdot \end{bmatrix} = \Omega_{11.2}^{-1}\left[I,\ -\Omega_{12}\Omega_{22}^{-1}\right],$$

where $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. It follows that

$$J'\Omega^{-1}J = \Omega_{11.2}^{-1}\left[I,\ -\Omega_{12}\Omega_{22}^{-1}\right]\begin{bmatrix} I \\ 0 \end{bmatrix} = \Omega_{11.2}^{-1}.$$

Taking each part in turn:
(a) Using the above results applied to Ω_h,

$$\begin{aligned}\left(I\otimes J'\Omega_h^{-1}\right)\text{vec}\left(\int_0^1 dB_h B_{h2}'\right) &= \left(I\otimes J'\Omega_h^{-1}\right)\left(\int_0^1 B_{h2}\otimes dB_h\right) \\ &= \int_0^1 \left(B_{h2}\otimes J'\Omega_h^{-1}dB_h\right) \\ &= \int_0^1 \left(B_{h2}\otimes \Omega_{h,11.2}^{-1}dB_{h,1.2}\right).\end{aligned}$$

Combining (30) and (31) gives

$$\begin{aligned}T_n\text{vec}\left(\widehat{C}_n - C_0\right) &\Rightarrow \phi_h^{-1}\left[\left(\int_0^1 B_{h2}B_{h2}'\right)^{-1}\otimes\Omega_{h,11.2}\right]\int_0^1\left(B_{h2}\otimes\Omega_{h,11.2}^{-1}dB_{h,1.2}\right) \\ &= \phi_h^{-1}\left[\left(\int_0^1 B_{h2}B_{h2}'\right)^{-1}\otimes I\right]\int_0^1 \left(B_{h2}\otimes dB_{h,1.2}\right)\end{aligned}$$

as required.
(b) The same arguments apply as in part (a), simply replacing Ω_h by Ω and B_h by B.

(c) Again, the proof follows that in part (a), noting that

$$\text{vec}\left(F_2(Z, y_{02})\right) = \left(\int_0^N Z_2 \otimes dZ\right) + \left(y_{02} \otimes Z(N)\right).$$

The stated results then follow straightforwardly. ||

Proof of Theorem 3.

Recall that $\Omega_h = (2\pi/h^3) f_{h,\xi\xi}(0)$ and that $f_{h,\xi\xi}(0) = \sum_{k=-\infty}^{\infty} f_{h,\xi\xi}^c(2k\pi/h)$ where

$$f_{h,\xi\xi}^c\left(\frac{2k\pi}{h}\right) = M_h\left(\frac{2ki\pi}{h}\right) f_{ww}^c\left(\frac{2k\pi}{h}\right) M_h\left(\frac{-2ki\pi}{h}\right)'.$$

It is convenient to partition $f_{ww}^c(\lambda)$ and $M_h(i\lambda)$ as

$$f_{ww}^c(\lambda) = \begin{bmatrix} f_{11}(\lambda) & f_{12}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) \end{bmatrix}, \quad M_h(i\lambda) = \begin{bmatrix} M_{11}(i\lambda) & M_{12}(i\lambda) \\ 0 & M_{22}(i\lambda) \end{bmatrix}.$$

Defining $\lambda_k = 2\pi k/h$, the sub-matrices of Ω_h can be written

$$\begin{aligned} \Omega_{h,11} &= \frac{2\pi}{h^3}\sum_{k=-\infty}^{\infty}\left[M_{11}(i\lambda_k)f_{11}(\lambda_k)M_{11}(-i\lambda_k)' + M_{12}(i\lambda_k)f_{21}(\lambda_k)M_{11}(-i\lambda_k)'\right. \\ &\quad \left. + M_{11}(i\lambda_k)f_{12}(\lambda_k)M_{12}(-i\lambda_k)' + M_{12}(i\lambda_k)f_{22}(\lambda_k)M_{12}(-i\lambda_k)'\right], \qquad \text{(C13)} \\ \Omega_{h,12} &= \frac{2\pi}{h^3}\sum_{k=-\infty}^{\infty}\left[M_{11}(i\lambda_k)f_{12}(\lambda_k) + M_{12}(i\lambda_k)f_{22}(\lambda_k)\right]M_{22}(-i\lambda_k)', \qquad \text{(C14)} \\ \Omega_{h,22} &= \frac{2\pi}{h^3}\sum_{k=-\infty}^{\infty} M_{22}(i\lambda_k)f_{22}(\lambda_k)M_{22}(-i\lambda_k)', \qquad \text{(C15)} \end{aligned}$$

and where $\Omega_{h,21} = \Omega_{h,12}'$. Lemma A6 in Appendix A establishes that $M_{11}(0) = \phi_h I_{m_1}$, $M_{22}(0) = hI_{m_2}$ and $M_{22}(i\lambda_k) = 0$ for all $k \neq 0$, so that (C13), (C14) and (C15) simplify as

$$\begin{aligned} \Omega_{h,11} &= \frac{2\pi}{h^3}\left[\phi_h^2 f_{11}(0) + \phi_h\left(M_{12}(0)f_{21}(0) + f_{12}(0)M_{12}(0)'\right) + M_{12}(0)f_{22}(0)M_{12}(0)'\right] \\ &\quad + \widetilde{\Omega}_{h,11}, \qquad \text{(C16)} \\ \Omega_{h,12} &= \frac{2\pi}{h^2}\left[\phi_h f_{12}(0) + M_{12}(0)f_{22}(0)\right], \qquad \text{(C17)} \\ \Omega_{h,22} &= \frac{2\pi}{h}f_{22}(0), \qquad \text{(C18)} \end{aligned}$$

where

$$\begin{aligned} \widetilde{\Omega}_{h,11} &= \frac{2\pi}{h^3}\sum_{k\neq 0}\left[M_{11}(i\lambda_k)f_{11}(\lambda_k)M_{11}(-i\lambda_k)' + M_{12}(i\lambda_k)f_{21}(\lambda_k)M_{11}(-i\lambda_k)'\right. \\ &\quad \left. + M_{11}(i\lambda_k)f_{12}(\lambda_k)M_{12}(-i\lambda_k)' + M_{12}(i\lambda_k)f_{22}(\lambda_k)M_{12}(-i\lambda_k)'\right]. \qquad \text{(C19)} \end{aligned}$$

From (C17) and (C18),

$$\begin{aligned}\Omega_{h,12}\Omega_{h,22}^{-1}\Omega_{h,21} &= \frac{2\pi}{h^3}\left[\phi_h^2 f_{12}(0)f_{22}(0)^{-1}f_{21}(0) + \phi_h\left(M_{12}(0)f_{21}(0)' + f_{12}(0)M_{12}(0)'\right)\right.\\ &\quad \left. + M_{12}(0)f_{22}(0)M_{12}(0)'\right], \qquad \text{(C20)}\end{aligned}$$

so that, combining (C16) and (C20),

$$\Omega_{h,11.2} = \frac{2\pi}{h^3}\phi_h^2 f_{11.2}(0) + \widetilde{\Omega}_{h,11}, \qquad \text{(C21)}$$

where $f_{11.2}(0) = f_{11}(0) - f_{12}(0)f_{22}(0)^{-1}f_{21}(0)$. The variance matrix of interest in the conditional distribution therefore has the representation

$$\begin{aligned}V(h) &= h^2\phi_h^{-2}\left[\frac{h}{2\pi}f_{22}(0)^{-1}\otimes\left(\frac{2\pi}{h^3}\phi_h^2 f_{11.2}(0) + \widetilde{\Omega}_{h,11}\right)\right]\\ &= f_{22}(0)^{-1}\otimes\left(f_{11.2}(0) + \frac{h^3}{2\pi\phi_h^2}\widetilde{\Omega}_{h,11}\right). \qquad \text{(C22)}\end{aligned}$$

The most complicated term to investigate here is $\widetilde{\Omega}_{h,11}$. Bearing (C19) in mind, the definitions of the $M_{ij}(\lambda)$ in terms of the underlying scalar filter functions yields

$$M_{11}(i\lambda_k)f_{11}(\lambda_k)M_{11}(-i\lambda_k)' = \begin{bmatrix} |m_1^S(i\lambda_k)|^2 f_{11}^{SS}(\lambda_k) & 0 \\ 0 & 0 \end{bmatrix},$$

$$M_{12}(i\lambda_k)f_{21}(\lambda_k)M_{11}(-i\lambda_k)' = \begin{bmatrix} \begin{array}{c} C_{SS}m_{12}^{SS}(i\lambda_k)m_1^S(-i\lambda_k)f_{21}^{SS}(\lambda_k) \\ +C_{SF}m_{12}^{SF}(i\lambda_k)m_1^S(-i\lambda_k)f_{21}^{FS}(\lambda_k) \end{array} & 0 \\ 0 & 0 \end{bmatrix},$$

$$M_{22}(i\lambda_k)f_{22}(\lambda_k)M_{22}(-i\lambda_k)' = \begin{bmatrix} \begin{array}{c} C_{SS}|m_{12}^{SS}(i\lambda_k)|^2 f_{22}^{SS}(\lambda_k)C_{SS}' \\ +C_{SF}m_{12}^{SF}(i\lambda_k)m_{12}^{SS}(-i\lambda_k)f_{22}^{FS}(\lambda_k)C_{SS}' \\ +C_{SS}m_{12}^{SS}(i\lambda_k)m_{12}^{SF}(-i\lambda_k)f_{22}^{SF}(\lambda_k)C_{SF}' \\ +C_{SF}|m_{12}^{SF}(i\lambda_k)|^2 f_{22}^{FF}(\lambda_k)C_{SF}' \end{array} & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence $\widetilde{\Omega}_{h,11}$ is of the form

$$\widetilde{\Omega}_{h,11} = \begin{bmatrix} \widetilde{\Omega}_{h,11}^{SS} & 0 \\ 0 & 0 \end{bmatrix}.$$

Now, from Lemma A6, note that for $k \neq 0$,

$$|m_1^S(i\lambda_k)|^2 = |m_{12}^{SS}(i\lambda_k)|^2 = |m_{12}^{SF}(i\lambda_k)|^2 = m_{12}^{SS}(i\lambda_k)m_{12}^{SF}(-i\lambda_k) = \frac{h^2\phi_h^2}{h^2+4k^2\pi^2},$$

$$m_{12}^{SS}(i\lambda_k)m_1^S(-i\lambda_k) = m_{12}^{SF}(i\lambda_k)m_1^S(-i\lambda_k) = -\frac{h^2\phi_h^2}{h^2+4k^2\pi^2}.$$

Hence $\widetilde{\Omega}_{h,11}^{SS} = (2\pi/h^3)\sum_{k\neq 0} h^2\phi_h^2(h^2+4k^2\pi^2)^{-1}P_k^{SS}(h)$, where

$$\begin{aligned}P_k^{SS}(h) &= f_{11}^{SS}(\lambda_k) - C_{SS}f_{21}^{SS}(\lambda_k) - C_{SF}f_{21}^{FS}(\lambda_k) - f_{12}^{SS}(\lambda_k)C_{SS}' - f_{12}^{SF}(\lambda_k)C_{SF}' \\ &\quad + C_{SS}f_{22}^{SS}(\lambda_k)C_{SS}' + C_{SF}f_{22}^{FS}(\lambda_k)C_{SS}' + C_{SS}f_{22}^{SF}(\lambda_k)C_{SF}' + C_{SF}f_{22}^{FF}(\lambda_k)C_{SF}'.\end{aligned}$$

Let $J_S = [I_{m_1^S},\ 0]$, $C_S = [C_{SS},\ C_{SF}]$ and $J_C = [J_S,\ -C_S]$. Then

$$P_k^{SS}(h) = J_S f_{11}(\lambda_k)J_S' - C_S f_{21}(\lambda_k)J_S' - J_S f_{12}(\lambda_k)C_S' + C_S f_{22}(\lambda_k)C_S' = J_C f_{ww}^c(\lambda_k)J_C',$$

so that

$$\widetilde{\Omega}_{h,11}^{SS} = \frac{2\pi\phi_h^2}{h^3}J_C\sum_{k\neq 0}\frac{1}{1+(4\pi^2k^2/h^2)}f_{ww}^c\left(\frac{2\pi k}{h}\right)J_C' \equiv \frac{2\pi\phi_h^2}{h^3}\widetilde{V}_{h,11}, \qquad \text{(C23)}$$

where $\widetilde{V}_{h,11}$ is implicitly defined and is positive semi-definite under the assumed properties of $f_{ww}^c(\lambda)$ in Assumption 1. Hence the matrix $V(h)$ has the representation

$$V(h) = f_{22}(0)^{-1} \otimes \left(f_{11.2}(0) + \widetilde{V}_h\right),$$

where

$$\widetilde{V}_h = \begin{bmatrix} \widetilde{V}_{h,11} & 0 \\ 0 & 0 \end{bmatrix}.$$

Turning to the matrix $V_0 = \Omega_{22}^{-1} \otimes \Omega_{11.2}$, recall that $\Omega = 2\pi f_{ww}^c(0)$. It is easy to show that $\Omega_{11.2} = 2\pi f_{11.2}(0)$ and that $\Omega_{22} = 2\pi f_{22}(0)$, thus yielding

$$V_0 = f_{22}(0)^{-1} \otimes f_{11.2}(0).$$

It immediately follows that the matrix difference $V(h) - V_0$ is given by

$$V(h) - V_0 = f_{22}(0)^{-1} \otimes \widetilde{V}_h,$$

which is clearly positive semi-definite under Assumption 2. ||

Proof of Proposition 1.

Follows straightforwardly from the form of the matrix $f_{ww,22}^c(0)^{-1} \otimes \widetilde{V}_h$. ||

APPENDIX D

This Appendix provides further details concerning the generation of the discrete time data for the simulation experiments. From the expression for ξ_{th} it is clear that this disturbance vector can be expressed as the sum of two components, so that $\xi_{th} = e_{th} - JAu_{th}$, where

$$e_{th} = \int_0^h w(th-s)ds, \quad u_{th} = \int_0^h \phi(s)w(th-s)ds.$$

Deriving the autocovariance properties of these two components enables each one to be generated from a single set of N(0,1) random variables.

First, note that $w(th)$ satisfies the difference equation

$$w(th) = e^{hG}w(th-h) + \int_{th-h}^{th} e^{(th-s)G}\zeta(ds).$$

Integrating over the interval $[0,h]$ yields a difference equation for e_{th}, given by $e_{th} = e^{hG}e_{th-h} + v_{th}$, where

$$v_{th} = \int_0^h \int_{th-h}^{th} e^{(th-s-r)G}\zeta(ds)dr.$$

It is convenient to express v_{th} as a pair of single stochastic integrals with respect to the random measure $\zeta(d\tau)$. The justification of the change in the order of integration has been rigorously demonstrated by McCrorie (2000), and the method yields

$$v_{th} = \int_{th-h}^{th} \Phi_1(th-s)\zeta(ds) + \int_{th-2h}^{th-h} \Phi_2(th-h-s)\zeta(ds),$$

where $\Phi_1(z) = G^{-1}[e^{zG} - I_2]$ and $\Phi_2(z) = G^{-1}[e^{hG} - e^{zG}]$. The autocovariance properties of v_{th} then follow straightforwardly, yielding (given the autocovariance properties of $\zeta(d\tau)$)

$$Ev_{th}v_{th}' = \int_0^h \Phi_1(r)\Phi_1(r)'dr + \int_0^h \Phi_2(r)\Phi_2(r)'dr,$$

$$Ev_{th}v_{th-h}' = \int_0^h \Phi_2(r)\Phi_1(r)'dr,$$

while $Ev_{th}v_{th-jh}' = 0$ for $|j| \geq 2$. These integrals are straightforwardly expressed in terms of integrals of the matrices e^{rG} and e^{hG}, although the derivations are somewhat tedious and are omitted for brevity.

Applying a similar procedure to u_{th} yields the difference equation $u_{th} = e^{hG}u_{th-h} + z_{th}$, where

$$z_{th} = \int_0^h \phi(s) \int_{th-h-r}^{th-r} e^{(th-s-r)G}\zeta(ds)dr.$$

This double integral can also be reduced to a pair of single stochastic integrals with respect to $\zeta(d\tau)$, yielding

$$z_{th} = \int_{th-h}^{th} \Phi_3(th-s)\zeta(ds) + \int_{th-2h}^{th-h} \Phi_4(th-h-s)\zeta(ds),$$

where $\Phi_3(z) = K(0)e^{zG} - K(z)$, $\Phi_4(z) = K(z)e^{hG} - K(-h)e^{zG}$, and $K(z) = G^{-1} - e^z(I + G)^{-1}$. It follows that z_{th} is also an MA(1) process, with

$$Ez_{th}z'_{th} = \int_0^h \Phi_3(r)\Phi_3(r)'dr + \int_0^h \Phi_4(r)\Phi_4(r)'dr,$$

$$Ez_{th}z'_{th-h} = \int_0^h \Phi_4(r)\Phi_3(r)'dr,$$

while $Ez_{th}z'_{th-jh} = 0$ for $|j| \geq 2$. Once more some tedious algebra enables these integrals to be expressed in terms of integrals with respect to e^{zG} and $e^{z(I+G)}$, for example, which can be evaluated for the given values of parameters that define the matrix G.[13]

The processes v_{th} and z_{th} are both MA(1). It remains to describe the method by which they were generated from a sequence of independent N(0,1) variates. Consider, first, v_{th}. The procedure for z_{th} is identical. Denote the variance matrix by V_0 and the first-order autocovariance by V_1. Then, if $v_{th} = \epsilon_{th} + \Lambda\epsilon_{th-h}$, where ϵ_{th} is an uncorrelated sequence with variance matrix P, it follows that P, Λ, V_0 and V_1 are related by the formulae

$$V_0 = P + \Lambda P \Lambda', \quad V_1 = \Lambda P.$$

The matrices P and Λ were derived (numerically) to satisfy these equations for each experiment. Then, given a sequence of independent N(0,1) variates μ_{th}, and denoting the Cholesky factorisation of P by $P = P_c P_c'$, the ϵ_{th} are determined by $\epsilon_{th} = P_c\mu_{th}$. The process v_{th} can then be generated according to the MA(1) representation, from which e_{th} can be generated using the AR(1) representation. The same procedure, using the same set of μ_{th}, then determines z_{th} (using the appropriate variance matrix and MA(1) coefficient matrix), and hence u_{th}. The same set of underlying random variates must be used, because e_{th} and u_{th} are functions of the same underlying (continuous time) process $w(\tau)$.

[13]The expressions used assume that the matrices G and $(I + G)$ are nonsingular, conditions which are certainly satisfied in the experiments conducted here.

REFERENCES

ANDREWS, D. W. K. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817–858.

BERGSTROM, A. R. (1984): "Continuous time stochastic models and issues of aggregation over time," in *Handbook of Econometrics, Volume 2*, edited by Z. Griliches and M. D. Intriligator, pp. 1145–1212, North-Holland, Amsterdam.

CHAMBERS, M. J. (1999): "Discrete time representation of stationary and nonstationary continuous time systems," *Journal of Economic Dynamics and Control*, 23, 619–639.

CHAMBERS, M. J. (2000): "The asymptotic efficiency of cointegration estimators under temporal aggregation," University of Essex, unpublished.

CHAMBERS, M. J. (2001): "Temporal aggregation and the finite sample performance of spectral regression estimators in cointegrated systems: a simulation study," *Econometric Theory* (forthcoming).

DAVIDSON, J. E. H. (1994): *Stochastic Limit Theory*, Oxford University Press, Oxford.

HANNAN, E. J. (1963): "Regression for time series," in *Time Series Analysis*, edited by M. Rosenblatt, chapter 2, pp. 17–37, Wiley, New York.

HANNAN, E. J. (1970): *Multiple Time Series*, Wiley, New York.

HANSEN, B. E. (1992): "Convergence to stochastic integrals for dependent heterogeneous processes," *Econometric Theory*, 8, 489–500.

HOOKER, M. A. (1993): "Testing for cointegration: power versus frequency of observation," *Economics Letters*, 41, 359–362.

LAHIRI, K. AND MAMINGI, N. (1995): "Testing for cointegration: power versus frequency of observation - another view," *Economics Letters*, 49, 121–124.

MCCRORIE, J. R. (2000): "Deriving the exact discrete analog of a continuous time system," *Econometric Theory*, 16, 998–1015.

OTERO, J. AND SMITH, J. (2000): "Testing for cointegration: power versus frequency of observation - further Monte Carlo results," *Economics Letters*, 67, 5–9.

PERRON, P. (1991): "Test consistency with varying sampling frequency," *Econometric Theory*, 7, 341–368.

PHILLIPS, P. C. B. (1987a): "Time series regression with a unit root," *Econometrica*, 55, 277–301.

Phillips, P. C. B. (1987b): "Towards a unified asymptotic theory for autoregression," *Biometrika*, 74, 535–547.

Phillips, P. C. B. (1989): "Partially identified econometric models," *Econometric Theory*, 5, 181–240.

Phillips, P. C. B. (1991a): "Error correction and long run equilibrium in continuous time," *Econometrica*, 59, 967–980.

Phillips, P. C. B. (1991b): "Optimal inference in cointegrated systems," *Econometrica*, 59, 283–306.

Phillips, P. C. B. (1991c): "Spectral regression for cointegrated time series," in *Nonparametric and Semiparametric Methods in Economics and Statistics*, edited by W. A. Barnett, J. Powell and G. Tauchen, chapter 16, pp. 413–435, Cambridge University Press, Cambridge.

Phillips, P. C. B. and Hansen, B. E. (1990): "Statistical inference in instrumental variables regression with I(1) processes," *The Review of Economic Studies*, 57, 99–125.

Priestley, M. B. (1981): *Spectral Analysis and Time Series*, Academic Press, London.

Rozanov, Y. A. (1967): *Stationary Random Processes*, Holden Day, New York.

Sargan, J. D. (1974): "Some discrete approximations to continuous time models," *Journal of the Royal Statistical Society Series B*, 36, 74–90.

Shiller, R. J. and Perron, P. (1985): "Testing the random walk hypothesis: power versus frequency of observation," *Economics Letters*, 18, 381–386.

TABLE 1

Mean square errors

Estimator	h	Values of N 25	50	100	25	50	100
		Experiment 1			Experiment 2		
OLS	1	158.14	31.20	6.89	61.85	15.17	3.68
	1/4	168.63	33.78	7.37	35.19	8.50	1.98
	1/12	175.67	35.21	7.66	37.37	8.96	2.07
SPEC	1	167.10	33.00	7.15	18.79	4.02	0.91
	1/4	175.37	34.84	7.46	32.18	7.71	1.79
	1/12	197.79	38.79	8.22	46.40	10.92	2.58
BAND	1	175.32	34.10	7.28	17.54	3.58	0.80
	1/4	166.15	32.85	7.04	21.84	4.64	1.00
	1/12	170.94	33.55	7.11	29.14	6.34	1.29
		Experiment 3			Experiment 4		
OLS	1	160.90	32.91	7.89	56.72	13.51	3.51
	1/4	107.87	21.76	5.22	27.06	6.18	1.54
	1/12	101.84	20.51	4.92	26.49	6.01	1.49
SPEC	1	87.03	15.45	3.48	12.33	2.46	0.57
	1/4	81.72	15.27	3.50	23.27	5.37	1.34
	1/12	84.73	15.76	3.62	31.37	7.16	1.81
BAND	1	91.53	15.75	3.48	10.86	2.07	0.46
	1/4	81.55	15.05	3.41	15.36	3.12	0.69
	1/12	80.50	15.02	3.43	21.25	4.46	0.96

Note: All entries have been multiplied by 10^4.

TABLE 2

Percentage coverage rates of 90% confidence intervals

Estimator	h	Values of N					
		25	50	100	25	50	100
		Experiment 1			Experiment 2		
OLS	1	28.69	20.06	13.31	33.70	24.66	17.42
	1/4	26.44	18.94	12.75	46.34	34.11	23.93
	1/12	35.45	25.19	17.48	66.79	50.53	36.77
SPEC ($\widehat{V}_1$)	1	69.74	69.66	69.04	87.11	88.25	88.55
	1/4	84.97	84.84	84.65	99.52	99.51	99.52
	1/12	95.38	95.23	96.02	100.00	100.00	100.00
SPEC ($\widehat{V}_2$)	1	69.74	69.66	69.04	87.11	88.25	88.55
	1/4	52.50	53.35	52.20	82.11	81.86	81.99
	1/12	44.12	44.69	44.58	69.39	68.79	68.26
BAND ($\widehat{V}_{10}$)	1	75.98	80.10	83.12	86.77	88.38	88.55
	1/4	94.44	96.50	97.92	99.60	99.74	99.82
	1/12	99.32	99.58	99.86	100.00	100.00	100.00
BAND ($\widehat{V}_{20}$)	1	76.97	80.86	83.72	87.84	89.16	89.30
	1/4	69.55	73.95	76.79	87.00	87.89	88.84
	1/12	60.67	65.44	68.97	84.48	85.58	87.14
		Experiment 3			Experiment 4		
OLS	1	24.95	18.55	13.71	34.55	25.08	17.40
	1/4	20.15	14.84	10.87	42.01	30.04	21.34
	1/12	23.55	16.74	12.39	56.78	41.98	30.20
SPEC ($\widehat{V}_1$)	1	66.27	69.13	70.54	86.87	88.43	87.65
	1/4	78.04	80.01	81.10	98.44	98.64	98.52
	1/12	88.14	89.80	92.34	99.98	99.93	99.98
SPEC ($\widehat{V}_2$)	1	66.27	69.13	70.54	86.87	88.43	87.65
	1/4	46.63	48.29	49.72	75.36	75.26	74.17
	1/12	36.31	37.14	39.10	57.52	56.49	56.65
BAND ($\widehat{V}_{10}$)	1	72.28	79.40	83.48	87.80	89.90	89.73
	1/4	92.92	97.10	98.05	99.54	99.63	99.77
	1/12	98.53	99.62	99.92	100.00	100.00	100.00
BAND ($\widehat{V}_{20}$)	1	73.42	80.36	84.07	88.68	90.64	90.39
	1/4	67.18	73.68	78.10	86.16	87.58	88.21
	1/12	57.36	64.23	70.91	81.48	83.44	85.40

TABLE 3

Percentage coverage rates of 95% confidence intervals

		Values of N					
Estimator	h	25	50	100	25	50	100
		Experiment 1			Experiment 2		
OLS	1	33.67	23.42	15.69	40.01	29.41	21.09
	1/4	31.43	22.33	15.25	54.18	39.98	28.29
	1/12	41.75	29.85	20.79	75.65	58.50	43.30
SPEC ($\widehat{V}_1$)	1	77.83	77.95	77.45	92.63	93.59	94.05
	1/4	90.82	91.07	90.89	99.93	99.93	99.96
	1/12	98.16	97.94	98.59	100.00	100.00	100.00
SPEC ($\widehat{V}_2$)	1	77.83	77.95	77.45	92.63	93.59	94.05
	1/4	61.06	61.49	60.54	89.31	89.03	89.57
	1/12	50.91	51.82	51.69	79.65	79.32	78.71
BAND ($\widehat{V}_{10}$)	1	83.40	86.81	89.25	92.19	93.23	93.94
	1/4	97.49	98.61	99.29	99.93	99.60	99.97
	1/12	99.81	99.91	100.00	100.00	100.00	100.00
BAND ($\widehat{V}_{20}$)	1	84.32	87.43	89.72	93.00	93.87	94.46
	1/4	77.84	81.53	84.07	92.69	93.48	93.91
	1/12	69.07	73.15	77.70	91.04	91.86	93.04
		Experiment 3			Experiment 4		
OLS	1	29.22	22.25	16.14	40.69	29.62	20.71
	1/4	24.14	17.54	12.97	49.23	35.69	25.52
	1/12	27.80	20.00	14.75	64.93	49.33	35.51
SPEC ($\widehat{V}_1$)	1	74.37	77.12	78.51	92.46	93.47	93.32
	1/4	84.91	87.18	87.94	99.72	99.67	99.66
	1/12	92.92	94.65	96.07	100.00	100.00	100.00
SPEC ($\widehat{V}_2$)	1	74.37	77.12	78.51	92.46	93.47	93.32
	1/4	54.35	55.99	57.29	83.68	83.43	83.08
	1/12	42.21	43.88	45.46	68.10	67.49	66.47
BAND ($\widehat{V}_{10}$)	1	79.82	86.51	89.74	92.53	94.47	94.54
	1/4	96.29	98.88	99.41	99.91	99.94	99.98
	1/12	99.44	99.86	100.00	100.00	100.00	100.00
BAND ($\widehat{V}_{20}$)	1	80.63	87.18	90.18	93.37	95.04	95.07
	1/4	75.00	81.46	85.40	91.99	93.39	93.78
	1/12	65.87	72.40	79.16	88.78	90.70	91.49

TABLE 4

Estimates of the long-run PPP relationship

using month-end exchange rate data

	Monthly		Quarterly		Annual	
Estimator	β_1	β_2	β_1	β_2	β_1	β_2
OLS	0.9438	0.9061	0.9528	0.9048	0.9187	0.9071
	(0.0315)	(0.0045)	(0.0559)	(0.0079)	(0.1267)	(0.0178)
SPEC(1/10)	0.6877	0.9366	0.6383	0.9437	0.6231	0.9442
	(0.0234)	(0.0036)	(0.0448)	(0.0068)	(0.0949)	(0.0141)
SPEC(1/3)	0.5452	0.9458	0.5903	0.9404	0.5774	0.9502
	(0.0237)	(0.0045)	(0.0443)	(0.0080)	(0.0876)	(0.0136)
SPEC(2/5)	0.5346	0.9447	0.6141	0.9397	0.5774	0.9502
	(0.0235)	(0.0046)	(0.0402)	(0.0071)	(0.0876)	(0.0136)
BAND(1/10)	0.8415	0.9203	0.8129	0.9243	0.7998	0.9240
	(0.0325)	(0.0046)	(0.0630)	(0.0088)	(0.1292)	(0.0180)
BAND(1/3)	0.8006	0.9262	0.8114	0.9247	0.9312	0.9078
	(0.0766)	(0.0108)	(0.1233)	(0.0172)	(0.0340)	(0.0045)
BAND(2/5)	0.8063	0.9257	0.7993	0.9264	0.9312	0.9078
	(0.1007)	(0.0141)	(0.1559)	(0.0216)	(0.0340)	(0.0045)
FM-OLS(P)	0.9062	0.8878	0.9192	0.8880	0.9317	0.8873
	(0.0953)	(0.0135)	(0.0997)	(0.0139)	(0.0989)	(0.0135)
FM-OLS(B)	1.0332	0.8877	1.0343	0.8877	1.0445	0.8841
	(0.1521)	(0.0215)	(0.1565)	(0.0218)	(0.1833)	(0.0250)

Note: Figures in parentheses are standard errors.

TABLE 5

Estimates of the long-run PPP relationship

using monthly-averaged exchange rate data

	Monthly		Quarterly		Annual	
Estimator	β_1	β_2	β_1	β_2	β_1	β_2
OLS	0.9436	0.9063	0.9443	0.9064	0.9440	0.9053
	(0.0313)	(0.0044)	(0.0550)	(0.0077)	(0.1278)	(0.0177)
SPEC(1/10)	0.7811	0.9227	0.6708	0.9395	0.6499	0.9412
	(0.0229)	(0.0035)	(0.0455)	(0.0069)	(0.0958)	(0.0141)
SPEC(1/3)	0.6968	0.9204	0.6277	0.9339	0.6011	0.9475
	(0.0226)	(0.0042)	(0.0433)	(0.0079)	(0.0881)	(0.0136)
SPEC(2/5)	0.6946	0.9185	0.6533	0.9333	0.6011	0.9475
	(0.0226)	(0.0043)	(0.0391)	(0.0069)	(0.0881)	(0.0136)
BAND(1/10)	0.8412	0.9205	0.8050	0.9257	0.8215	0.9225
	(0.0324)	(0.0046)	(0.0623)	(0.0087)	(0.1313)	(0.0181)
BAND(1/3)	0.7999	0.9264	0.8064	0.9258	0.9567	0.9059
	(0.0765)	(0.0108)	(0.1218)	(0.0170)	(0.0339)	(0.0044)
BAND(2/5)	0.8061	0.9258	0.7938	0.9275	0.9567	0.9059
	(0.1008)	(0.0141)	(0.1537)	(0.0213)	(0.0339)	(0.0044)
FM-OLS(P)	0.8978	0.8860	0.9070	0.8895	0.9529	0.8872
	(0.0586)	(0.0083)	(0.0859)	(0.0120)	(0.1012)	(0.0136)
FM-OLS(B)	1.0093	0.8916	1.0150	0.8910	1.0625	0.8835
	(0.1477)	(0.0208)	(0.1539)	(0.0214)	(0.1858)	(0.0249)

Note: Figures in parentheses are standard errors.