



**University of
Leicester**

DEPARTMENT OF ECONOMICS

Band-Limited Stochastic Processes in Discrete and Continuous Time

Stephen Pollock, University of Leicester, UK

**Working Paper No. 11/11
January 2011**

BAND-LIMITED STOCHASTIC PROCESSES IN DISCRETE AND CONTINUOUS TIME

By D.S.G. POLLOCK

University of Leicester

Email: stephen_pollock@sigmapl.u-net.com

A theory of band-limited linear stochastic processes is described and it is related to the familiar theory of ARMA models in discrete time. By ignoring the limitation on the frequencies of the forcing function, in the process of fitting a conventional ARMA model, one is liable to derive estimates that are severely biased. If the maximum frequency in the sampled data is less than the Nyquist value, then the underlying continuous function can be reconstituted by sinc function or Fourier interpolation. The estimation biases can be avoided by re-sampling the continuous process at a rate corresponding to the maximum frequency of the forcing function. Then, there is a direct correspondence between the parameters of the band-limited ARMA model and those of an equivalent continuous-time process.

Keywords: Stochastic Differential Equations, Band-Limited Stochastic Processes, Aliasing and Interference

1. Introduction

It is common to assume that the differential equations that are used for modelling stochastic processes in continuous time are driven by a continuous stream of infinitesimal impulses. These impulses, which constitute the increments of a Wiener process, are composed of an infinity of sinusoidal elements of all frequencies in the interval $[0, \infty)$.

Whereas a Wiener process is a fruitful mathematical abstraction that has many important applications, it is not always appropriate to macroeconomic modelling. An inspection of the periodograms of macroeconomic data sequences reveals that their various components tend to reside in strictly limited frequency bands; and it seems improbable that they should have arisen from the filtering of Wiener processes.

It is more realistic to assume that such components originate in the filtering of continuous white-noise processes that are band limited in the same manner as the components themselves. However, in adopting this point of view, we are challenged to produce a model of such a process. To achieve this is one of the purposes of this paper. A further purpose is to examine the effects of using ordinary autoregressive moving average (ARMA) models in an inappropriate way to derive parametric representations of band-limited processes.

Conventional ARMA processes are driven by discrete-time white-noise forcing functions, which have spectral density functions that are uniform across a range of frequencies running from zero to the Nyquist frequency of π radians per sampling interval, which is the limiting frequency that is observable in sampled data. When

the driving process is band limited to a subset of the interval $[0, \pi]$, there are liable to be severe biases in the estimated parameters, unless some account is taken of this fact. We will describe the steps that must be taken to obtain appropriate estimates.

2. Evidence of Band-Limited Processes

We should begin by presenting some evidence to support the assertion that macroeconomic data sequences are commonly composed of components that fall within limited frequency bands.

Figure 1 displays a sequence of the logarithms of the quarterly series of U.K. consumption over the period from 1955 to 1994, which comprises a total of 160 observations. Interpolated through this sequence is a quadratic trend, which can be taken to represent the growth path of consumption.

The deviations from this growth path are a combination of a low-frequency motion, reflecting the business cycle, with some high-frequency fluctuations that are due to the seasonal nature of economic activity. These deviations are represented in Figure 2, which also shows an interpolated continuous function that is designed to represent the effects of the business cycle. Figure 3 shows that the effect of taking the first differences of the logarithmic data is to emphasise the seasonal fluctuations at the expense of the business cycle.

The periodograms of Figures 4–6 clearly reflect the features of the corresponding data sequences. The periodogram is a discrete periodic function, which is a function of the Fourier transform of the periodic extension of the data. A single cycle of the periodogram occurs in the Nyquist interval $[-\pi, \pi]$ or, equivalently, in the interval $[0, 2\pi]$. However, since the data are real-valued, the periodograms are symmetric about the zero frequency and, therefore, they are characterised completely by graphs over the interval $[0, \pi]$

Figure 4 is the periodogram of the saw tooth function that corresponds to the periodic extension of the trended data. It owes its basic profile, which is that of a rectangular hyperbola, to the radical disjunctions that occur at the points where the end of one replication of the sample is joined to the beginning of the next replication. The spike in the vicinity of the zero frequency is so dominant in this periodogram that the remaining features, which are on a much smaller scale, are almost invisible.

The periodogram of the differenced data, which is in Figure 5, shows that, in eliminating the trend, the differencing strongly suppresses the low-frequency components of the data, which include the business cycle.

The periodogram of the deviations from the quadratic trend, which is in Figure 6, gives a clear representation of the spectral effects both of the business cycle and of the seasonal fluctuations, and it shows that they reside in separate frequency bands.

The spectral structure extending from zero frequency up to $\pi/8$ belongs to the business cycle. The prominent spikes located at the frequency $\pi/2$ and at the limiting Nyquist frequency of π are the property of the seasonal fluctuations. Elsewhere in the periodogram, there are wide dead spaces, which are punctuated by the spectral traces of minor elements of noise.

The slowly varying continuous function $z(t)$ interpolated through the deviations of Figure 2 has been created by combining a set of sine and cosine functions of

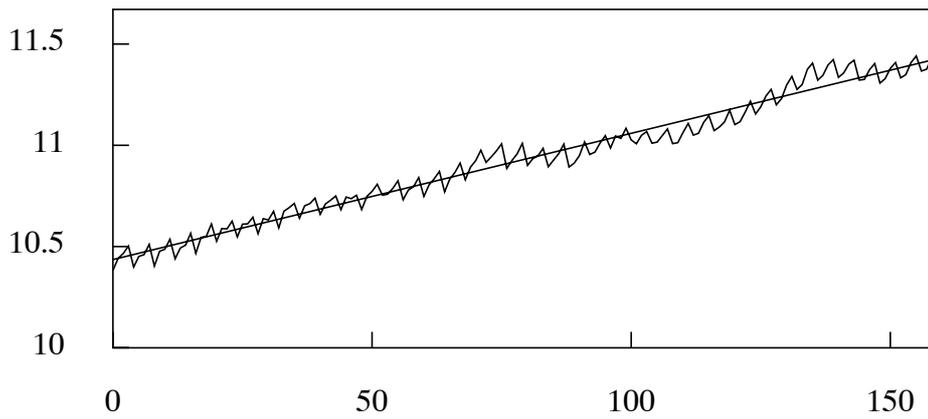


Figure 1. The quarterly series of the logarithms of consumption in the U.K., for the years 1955 to 1994, together with a quadratic trend interpolated by least-squares regression.

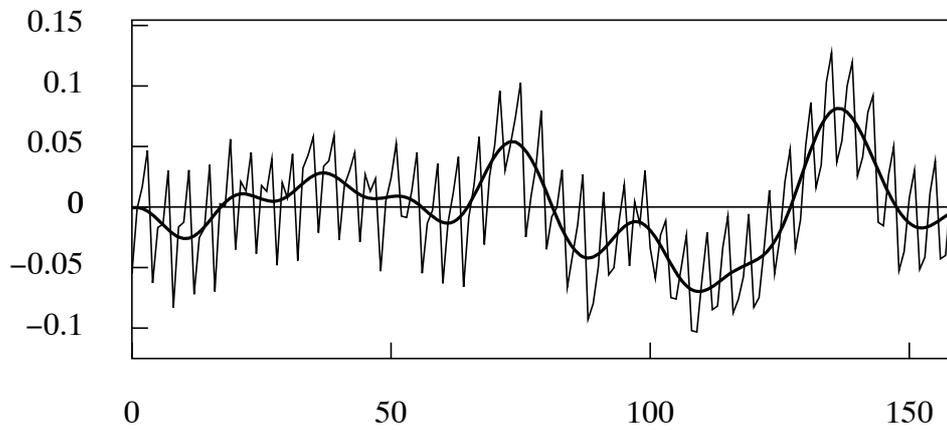


Figure 2. The residual sequence from fitting a quadratic trend to the logarithmic consumption data. The interpolated line, which represents the business cycle, has been synthesised from the Fourier ordinates in the frequency interval $[0, \pi/8]$.

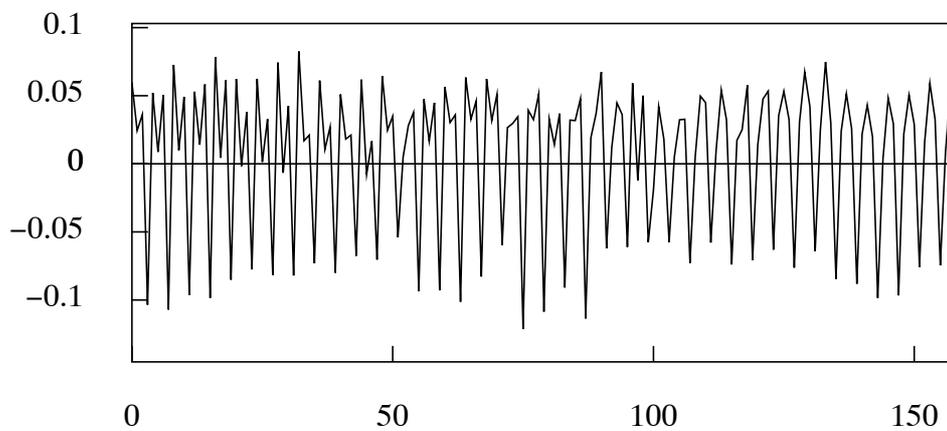


Figure 3. The differences of the logarithmic consumption data.

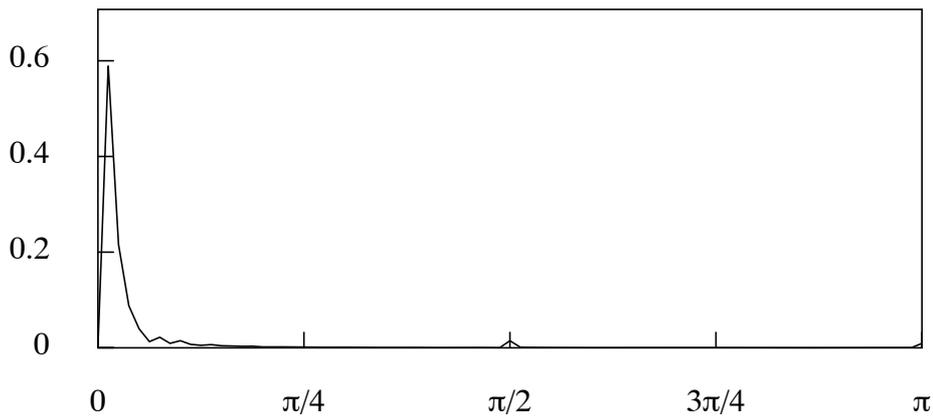


Figure 4. The periodogram of the logarithms of consumption in the U.K., for the years 1955 to 1994.

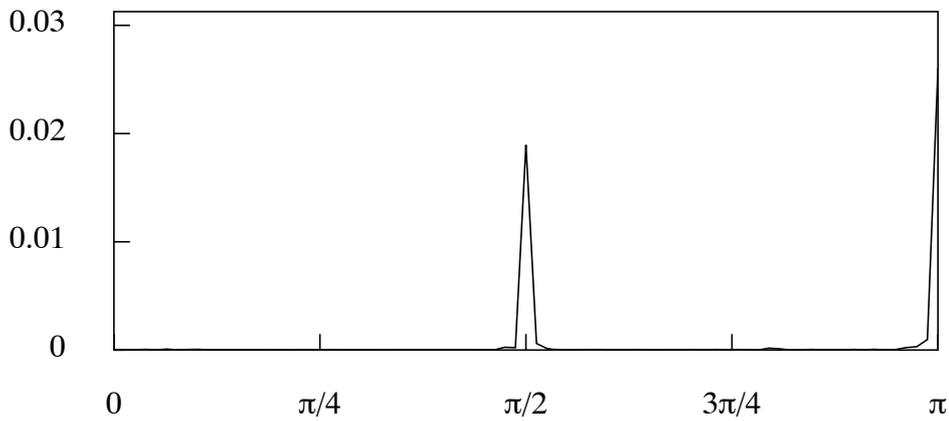


Figure 5. The periodogram of the first differences of the logarithmic consumption data.

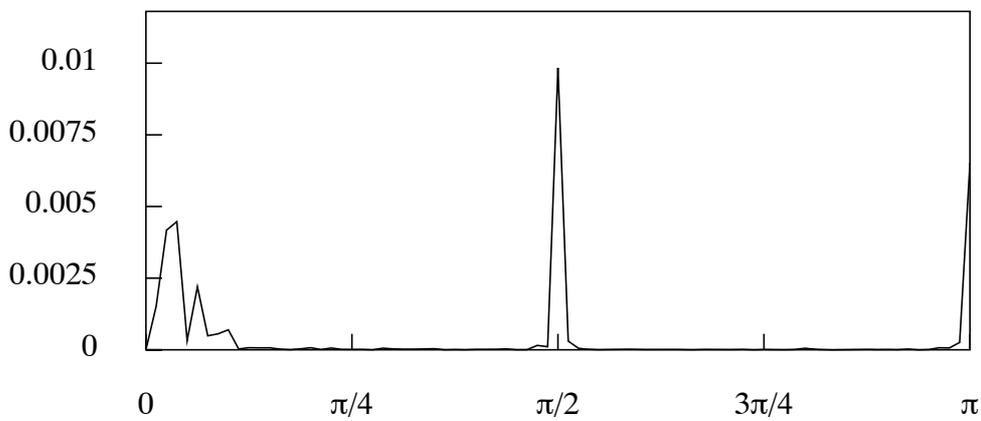


Figure 6. The periodogram of the residual sequence obtained from the linear detrending of the logarithmic consumption data.

increasing frequencies, which are regularly spaced and which extend no further than the limiting frequency of the business cycle, which is $\omega_c = \pi/8$. In the case of a sample of size T , the Fourier frequencies are the values $\omega_j = 2\pi j/T; j = 0, 1, \dots, [T/2]$ where $[T/2]$ is the integral part of $T/2$. Thus,

$$\begin{aligned} z(t) &= \sum_{j=0}^c \{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\} \\ &= \sum_{j=-c}^c \xi_j e^{i\omega_j t} \quad \text{with } \omega_c = \pi/8, \end{aligned} \tag{1}$$

where $\xi_j = (\alpha_j - i\beta_j)/2$ and $\xi_{-j} = \xi_j^* = (\alpha_j + i\beta_j)/2$. With $\omega_c = 2\pi c/T = \pi/8$ and $T = 160$, there is $c = 10$.

Some justification ought to be given for characterising the spectral structure of a trended sequence in terms of the periodogram of its deviations from an interpolated polynomial trend. If y is the vector of the data, then the vector of the deviations is given by the formula

$$e = Q(Q'Q)^{-1}Q'y, \tag{2}$$

wherein Q' is a submatrix of the p th-order matrix difference operator $\nabla_T^p = (I_T - L_T)^p$, which is obtained by deleting the first p rows. The difference operator $\nabla_T = I_T - L_T$ is formed from the identity matrix $I_T = [e_0, e_1, \dots, e_{T-1}]$ and from the matrix $L_T = [e_1, \dots, e_{T-1}, 0]$, obtained by deleting the leading vector from I_T and appending column of zeros to the end of the array.

Thus, for example, in the case of a second-order difference operator, which is appropriate to linear detrending, there is

$$\nabla_6^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} Q' \\ Q' \end{bmatrix}. \tag{3}$$

More generally, if $\nabla(z) = 1 - z$ is the difference operator, wherein z is an indeterminate algebraic symbol, then $\nabla_T^p = \nabla(L_T)^p$ and $Q' = [e_p, \dots, e_{T-1}]'\nabla_T^p$. It is apparent from equation (2) that the residual vector contains exactly the same information as the vector $Q'y$ of differences.

The periodogram of the polynomial residuals allows us to discern the spectral structure across the entire frequency range. The effect of increasing the degree of the polynomial, which generates a more flexible trend, is to attenuate the low-frequency elements relative to the high-frequency elements, but the bandwidths of the components in question remain the same. Thus, we have a device for accurately determining the domains of the various spectral structures that correspond to the components of the data sequence.

Our objective is to characterise the dynamics of the business cycle via the parameters of a fitted ARMA model. Such a model is liable to be applied to a

seasonally-adjusted version of the data, of which the periodogram will lack the spectral spikes at the seasonal frequency of $\pi/2$ and at the harmonic frequency of π . A second-order autoregressive AR(2) model with complex roots is the simplest of the models that might be appropriate to the purpose. The modulus of its roots should reveal the damping characteristics of the cycles, and their argument should indicate the angular velocity or, equivalently, the length of the cycles.

The parametric spectrum of an AR(2) model is supported on the entire frequency range $[0, \pi]$, which is to say that it is everywhere nonzero within this interval. However, the observed business cycle appears to belong to a band-limited process that has a zero-valued spectral density everywhere in the interval $(\pi/8, \pi]$.

The consequence of this disparity is that an AR(2) model that is fitted directly to the data is liable to deliver highly misleading estimates. Thus, it has been widely reported that, when it is fitted to seasonally-adjusted quarterly data, the model will deliver estimates that almost invariably imply real-valued roots, which fail adequately to represent the dynamics of the business cycle. (See Pagan 1997, for example.)

In order to estimate the parameters successfully, it is necessary to map the low-frequency spectral structure of the business cycle onto the full interval $[0, \pi]$. This involves creating a new data sequence at a lower sampling rate. It is also necessary to ensure that nothing is carried into the interval $[0, \pi]$ that does not belong to the low-frequency structure. This is achieved by applying an anti-aliasing or cleansing filter prior to resampling the data at the lesser rate.

In the process of describing the simple technique of resampling, we shall provide a model for the continuous-time band-limited process that generates the business cycle component. The parameters of this process are also the parameters of a discrete-time ARMA process that describes the resampled data.

3. The Sampling Process

To understand the methods that we propose to use in estimating the parameters of a band-limited processes, it is necessary to consider the nature of the sampling processes that deliver the discrete data on which econometric estimates are based.

Consider, therefore, the Fourier representation of a continuous real-valued square integrable function $x(t)$ defined over the real line. The following are the related expressions for the function and its Fourier transform $\xi(\omega)$:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \xi(\omega) d\omega \longleftrightarrow \xi(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} x(t) dt. \quad (4)$$

By sampling $x(t)$ at the integer time points, a sequence $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$ is generated of which the transform $\xi_S(\omega)$ is a 2π -periodic function. In that case,

$$x_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega t} \xi_S(\omega) d\omega \longleftrightarrow \xi_S(\omega) = \sum_{k=-\infty}^{\infty} x_k e^{-ik\omega}, \quad (5)$$

and $\xi_S(\omega)$ is described as the discrete-time Fourier transform of the data sequence.

At the sampled point x_t , to which the expressions under (4) and (5) both relate, there is,

$$x_t = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \xi(\omega) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega t} \xi_S(\omega) d\omega. \quad (6)$$

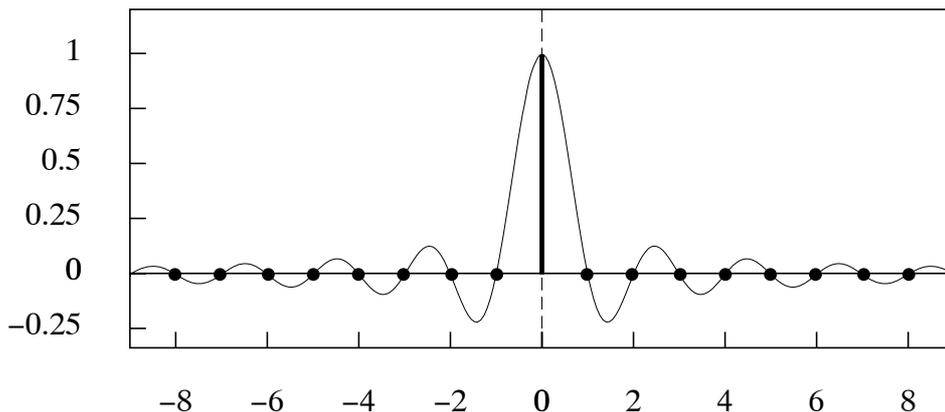


Figure 7. The sinc function wave-packet $\phi(t) = \sin(\pi t)/\pi t$ comprising frequencies in the interval $[0, \pi]$.

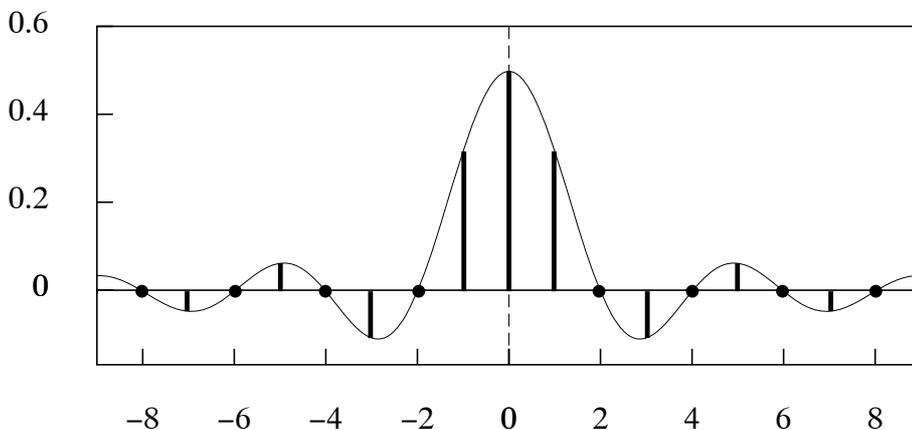


Figure 8. The sinc function wave-packet $\phi_1(t) = \sin(\pi t/2)/\pi t$ comprising frequencies in the interval $[0, \pi/2]$.

The equality of the two integrals implies that

$$\xi_S(\omega) = \sum_{j=-\infty}^{\infty} \xi(\omega + 2j\pi). \tag{7}$$

Thus, the function $\xi_S(\omega)$ is obtained by wrapping $\xi(\omega)$ around a circle of circumference of 2π and adding the coincident ordinates. The two functions will coincide at all frequencies in the interval $[-\pi, \pi]$ if $\xi(\omega) = 0$ for all $|\omega| \geq \pi$. Otherwise, $\xi_S(\omega)$ will be subject to a process of aliasing, whereby elements of the continuous function that are at frequencies in excess of π are confounded with elements at frequencies less than π . Thus, the so-called Nyquist frequency of π radians per period of observation represents the limit of what is directly observable in sampled data.

If the condition is fulfilled that $\xi(\omega) = 0$ for all $|\omega| \geq \pi$, then it should be possible to reconstitute the continuous function $x(t)$ from its sampled ordinates. This is the burden of the famous Nyquist–Shannon sampling theorem—see Shannon (1949, 1998)—which was foreshadowed in the work of Whittaker (1935).

When $\xi(\omega) = \xi_S(\omega)$ is a continuous function defined on the interval $[-\pi, \pi]$, it may be regarded as a periodic function of a period of 2π . Then, putting the RHS of (5) into the LHS of (4), and taking the integral over $[-\pi, \pi]$, in consequence of the band-limited nature of the function $x(t)$, gives

$$x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \sum_{k=-\infty}^{\infty} x_k e^{-ik\omega} \right\} e^{i\omega t} d\omega = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} x_k \int_{-\pi}^{\pi} e^{i\omega(t-k)} d\omega. \quad (8)$$

The integral on the RHS is evaluated as

$$\int_{-\pi}^{\pi} e^{i\omega(t-k)} d\omega = 2 \frac{\sin\{\pi(t-k)\}}{t-k}. \quad (9)$$

Putting this into the RHS of (8) gives

$$x(t) = \sum_{k=-\infty}^{\infty} x_k \frac{\sin\{\pi(t-k)\}}{\pi(t-k)} = \sum_{k=-\infty}^{\infty} x_k \phi(t-k), \quad (10)$$

where the continuous function

$$\phi(t-k) = \frac{\sin\{\pi(t-k)\}}{\pi(t-k)} \quad (11)$$

is the so-called sinc function, which is the Fourier transform of the following frequency function:

$$\phi(\omega) = \begin{cases} 1, & \text{if } |\omega| \in (0, \pi); \\ 1/2, & \text{if } \omega = \pm\pi, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

Equation (10) shows how the continuous function $x(t)$ can be reconstituted from the sampled ordinates $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$

In the case of a stationary stochastic process of an infinite duration, the sampled sequence would not be square summable and, therefore, in a strict sense, this proof of the interpolation via the Nyquist–Shannon Theory would not apply. Nevertheless, the stationary process can be regarded as a limiting case.

However, it is also appropriate to consider a finite data sequence as the product of a circular stochastic process. As we shall show, this leads to a modified theory of interpolation. An infinite stationary sequence can be accommodated on this basis by increasing circumference of the circle indefinitely.

The sequence of sinc functions $\phi(t-k); k \in \mathcal{Z} = \{0, \pm 1, \pm 2, \dots\}$ constitutes an orthogonal basis for the set of all functions band-limited to the frequency interval $[0, \pi]$. To show this, let $\phi(\omega)$ be the transform of $\phi(t)$ and consider the following autoconvolution:

$$\begin{aligned} \int_t \phi(t)\phi(\tau-t)dt &= \int_t \phi(t) \left\{ \frac{1}{2\pi} \int_{\omega} \phi(\omega) e^{i\omega(\tau-t)} d\omega \right\} dt \\ &= \frac{1}{2\pi} \int_{\omega} \phi(\omega) \left\{ \int_t \phi(t) e^{-i\omega t} dt \right\} e^{i\omega\tau} d\omega \\ &= \frac{1}{2\pi} \int_{\omega} \phi(\omega)\phi(\omega) e^{i\omega\tau} d\omega. \end{aligned} \quad (13)$$

The symmetry of $\phi(t)$ allows us to write $\phi(\tau - t) = \phi(t - \tau)$, whereas the idempotency of $\phi(\omega)$ gives $\phi^2(\omega) = \phi(\omega)$. Together, these two conditions indicate that $\phi(t)$ is its own autocorrelation function. Therefore, the condition

$$\phi(t) = 0 \quad \text{for } t \in \{\pm 1, \pm 2, \dots\} \quad (14)$$

indicates that sinc functions separated by integer distances are mutually orthogonal.

The sinc function $\phi(t)$ is represented in Figure 7. It is clear from this that the values of the ordinates of the function at the nonzero integer points $t \in \{\pm 1, \pm 2, \dots\}$ are zeros. Also, when the set of sinc functions $\{\phi(t - k); k \in \mathcal{Z}\}$ at unit displacements are sampled at the integer values of t , the result is nothing but the set of unit impulses at the integer points. This constitutes a basis for the set of all sequences defined over the set of integers.

The reconstruction or interpolation of a function in the manner suggested by the sampling theorem is not possible in practice, because it requires summing an infinite number of sinc functions, each of which is supported on the entire real line. Nevertheless, a continuous band-limited periodic function, defined on a finite interval, can be reconstituted from a finite number of wrapped or periodic sinc functions, which are Dirichlet kernels by another name. The Dirichlet kernel is obtained by sampling the sinc-function rectangle in the frequency domain.

Consider a continuous function $x(t)$ defined on the interval $[0, T)$, where, without loss of generality, T is taken to be an integer. (We may define the unit of time accordingly, which will eventually correspond to an interval between sampled observations.) For brevity, we shall consider only the case where T is even. Since the function on $[0, T)$ can also be regarded as a single cycle of a periodic function, such that $x(t + T) = x(t)$, it can be represented via a Fourier series expansion

$$x(t) = \sum_{j=-\infty}^{\infty} \xi_j e^{i\omega_j t} \longleftrightarrow \frac{1}{T} \int_0^T x(t) e^{-i\omega_j t}, \quad (15)$$

where $\omega_j = 2\pi j/T$ is the j th harmonic frequency, which is a multiple of the fundamental frequency $\omega_1 = 2\pi/T$, which corresponds to a single cycle within the time interval of length T .

If the function $x(t)$ is limited by the Nyquist frequency, then the index j is bounded by the integer value $n = T/2$. Then, the summation of the series expansion runs from $-n = -T/2$ to $n - 1$ or, more conveniently, from 0 to $T - 1$. (The change of index is allowable in consequence of the T -periodicity of the complex exponential function.) In that case, the Fourier ordinates ξ_j are from the discrete Fourier transform of T points sampled in the time domain, and there is

$$x(t) = \sum_{j=0}^{T-1} \xi_j e^{i\omega_j t} \longleftrightarrow \xi_j = \frac{1}{T} \sum_{t=0}^{T-1} x_t e^{-i\omega_j t}. \quad (16)$$

Putting the expressions for the Fourier ordinates into the finite Fourier series expansion of the time function and commuting the summation signs gives

$$x(t) = \sum_{j=0}^{T-1} \left\{ \frac{1}{T} \sum_{k=0}^{T-1} x_k e^{i\omega_j k} \right\} e^{i\omega_j t} = \frac{1}{T} \sum_{k=0}^{T-1} x_k \sum_{j=0}^{T-1} e^{i\omega_j(t-k)}. \quad (17)$$

The inner summation gives rise to the Dirichlet Kernel:

$$\phi_n^\circ(t) = \sum_{t=0}^{T-1} e^{i\omega_j t} = \frac{\sin([n - 1/2]\omega_1 t)}{\sin(\omega_1 t/2)}. \quad (18)$$

Thus, the Fourier expansion can be expressed in terms of the Dirichlet kernel, which is a circularly wrapped sinc function:

$$x(t) = \frac{1}{T} \sum_{t=0}^{T-1} x_k \phi_n^\circ(t - k). \quad (19)$$

The functions $\{\phi^\circ(t - k); k = 0, 1, \dots, T - 1\}$ are appropriate for reconstituting a continuous periodic function $x(t)$ defined on the interval $[0, T)$ from its sampled ordinates x_0, x_1, \dots, x_{T-1} . However, according to (16), the function can also be reconstituted, in the manner of equation (1), from its Fourier ordinates as

$$x(t) = \sum_{j=0}^{T-1} \xi_j e^{i\omega_j t} = \sum_{j=0}^{[T/2]} \{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\}, \quad (20)$$

where $\xi_j = (\alpha_j - i\beta_j)/2$ and $\xi_{-j} = (\alpha_j + i\beta_j)/2$ and where $[T/2]$ is the integral part of $T/2$.

For economic data, the sampling interval is typically a month, a quarter or a year. For financial data, it may be less. Whatever the case, it is unlikely that this interval will correspond to an integer multiple of the critical rate of sampling of whatever band-limited process might underlie the data. Therefore, it may not be possible to achieve the critical rate by subsampling the data.

Nevertheless, if the Nyquist frequency exceeds the maximum frequency of the process, then it will always be possible to resample the data at the critical rate. This can be achieved, once the continuous function has been reconstituted, via the formula of (20), from the relevant nonzero Fourier ordinates of the available data.

4. The Processes Underlying the Data

In this section, we shall describe the one-to-one correspondence that exists between a continuous-time ARMA process supported on the Nyquist frequency interval $[-\pi, \pi]$ and the discrete-time ARMA process of which the ordinates are obtained by sampling the continuous process at the critical rate.

The forcing function of a stationary linear stochastic process in discrete time is a white-noise sequence $\{\varepsilon_t; t = 0, \pm 1, \pm 2, \dots\}$ of independently and identically distributed random variables. The corresponding continuous-time process $\varepsilon(t)$ is obtained by associating sinc functions to each of the sample ordinates in manner of equation (10):

$$\varepsilon(t) = \sum_{k=-\infty}^{\infty} \varepsilon_k \phi(t - k). \quad (21)$$

The original sequence will be recovered by re-sampling this function at the integer points. Sampling the function at other points that are not the integers, but

which are separated by unit intervals, will also give rise to a white-noise sequence, albeit one that will differ from the original sequence.

One should be aware that, if $\varepsilon(t)$ is sampled at points that are not separated by the unit intervals, then the resulting sequence will show serial correlation. Therefore, in describing $\varepsilon(t)$ as a continuous-time white-noise process, one must refer to the unit sampling interval.

The autocovariance function of the continuous process $\varepsilon(t)$ is $\gamma_\varepsilon(\tau) = E\{\varepsilon(t)\varepsilon(t + \tau)\}$. Without loss of generality, t can be taken to be an integer point, with the effect that $\varepsilon(t) = \varepsilon_t$. Then,

$$\begin{aligned}\gamma_\varepsilon(\tau) &= \sum_{k=-\infty}^{\infty} E(\varepsilon_t \varepsilon_k) \phi(t + \tau - k) \\ &= \sigma_\varepsilon^2 \phi(\tau),\end{aligned}\tag{22}$$

which follows from the fact that

$$E(\varepsilon_t \varepsilon_k) = \begin{cases} 0, & \text{if } j \neq k, \\ \sigma_\varepsilon^2, & \text{if } j = k. \end{cases}\tag{23}$$

A continuous-time ARMA process supported on the Nyquist interval is derived by associating a sinc function kernel to each of the ordinates of a discrete-time ARMA process. Let $\{y_t; t = 0, \pm 1, \pm 2, \dots\}$ be the ordinates of the process and let

$$y(t) = \sum_k y_k \phi(t - k)\tag{24}$$

be the corresponding continuous-time trajectory. Then, the equation for the continuous ARMA process is

$$\sum_{j=0}^p \alpha_j y(t - j) = \sum_{j=0}^q \mu_j \varepsilon(t - j),\tag{25}$$

where $\alpha_0 = 1$.

The moving-average representation of this process is

$$y(t) = \sum_{j=0}^{\infty} \psi_j \varepsilon(t - j),\tag{26}$$

where the coefficients are from the series expansion of the rational function $\mu(z)/\alpha(z) = \psi(z)$, wherein $\alpha(z) = \alpha_0 + \alpha_1 z + \dots + \alpha_p z^p$ and $\mu(z) = \mu_0 + \mu_1 z + \dots + \mu_q z^q$ are the autoregressive and moving-average polynomials respectively.

The autocovariance function of the continuous ARMA process is given by

$$\begin{aligned}\gamma(\tau) &= E\left\{ \left[\sum_{i=0}^{\infty} \psi_i \varepsilon(t - \tau - i) \right] \left[\sum_{j=0}^{\infty} \psi_j \varepsilon(t - j) \right] \right\} \\ &= \sum_{i=-\infty}^{\infty} \gamma_i \phi(\tau - i),\end{aligned}\tag{27}$$

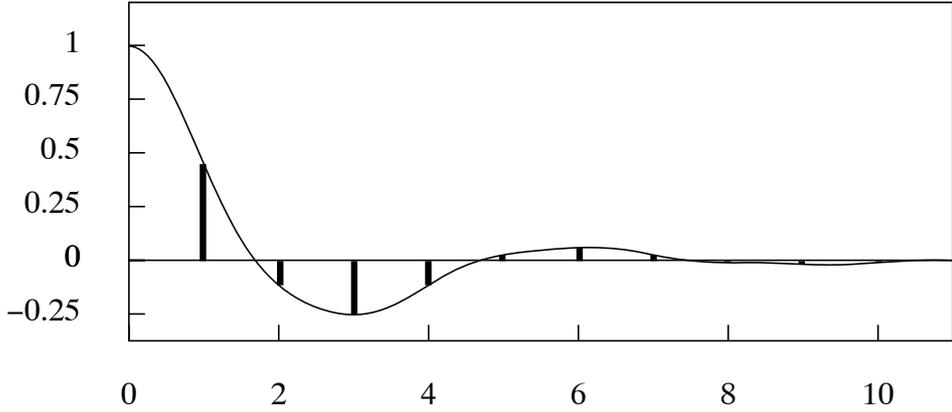


Figure 9. A continuous autocorrelation function of an AR(2) process, obtained via the inverse Fourier transform of the spectral density function, together with the corresponding discrete-time autocorrelations, calculated from the AR parameters.

where $\gamma_i = \sigma_\varepsilon^2 \sum_j \psi_j \psi_{j+i}$ is the i th autocovariance of the discrete-time process. Thus, in theory, the continuous-time autocovariance function is obtained from the discrete-time function by sinc-function interpolation. It can be seen that $\gamma(\tau) = \gamma_\tau$, when τ takes an integer value,

The autocovariance function $\gamma(\tau)$ is related to the spectral density function $f(\omega)$ via the following Fourier integral transforms, which take the same essential forms as those of equation (4):

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} f(\omega) d\omega \longleftrightarrow f(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} \gamma(\tau) d\tau. \quad (28)$$

Notice, however, that, here, the factor $1/2\pi$ is applied in the frequency domain as opposed to the time domain. (Since the purpose of the factor is ensure that the product of the Fourier transform and its inverse is unity, it can be placed in either domain.)

In the absence of aliasing, the spectral density function of the continuous process is identical, over the Nyquist interval $[-\pi, \pi]$, to that of the discrete-time process, and it is zero-valued outside the interval. By contrast, the spectrum of the discrete-time process is a periodic function, which replicates the function defined over the Nyquist interval in every preceding and succeeding interval of length 2π .

It follows that the spectrum $f(\omega)$ is also the discrete-time cosine Fourier transform of the sequence $\{\gamma_\tau; \tau = 0, \pm 1, \pm 2, \dots\}$ of the discrete autocovariances:

$$f(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\omega\tau} = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{\tau=1}^{\infty} \gamma_\tau \cos(\omega\tau) \right\}. \quad (29)$$

In practice, this is liable to be calculated by setting $z = \exp\{-i\omega\}$ within the autocovariance generating function

$$\gamma(z) = \sigma_\varepsilon^2 \frac{\mu(z)\mu(z^{-1})}{\alpha(z)\alpha(z^{-1})}. \quad (30)$$

The continuous autocovariance function $\gamma(\tau)$, which is the Fourier integral transform of $f(\omega)$, must be evaluated, in practice, via a discrete Fourier transform. By applying the discrete transform to a large number of ordinates sampled from $f(\omega)$ evenly over the interval $[-\pi, \pi]$, an accurate approximation of $\gamma(\tau)$ may be obtained. An example is provided by Figure 9.

A continuous function $y(t)$ that is limited by the frequency value $\omega_c < \pi$ can be expressed in the manner of (2) as

$$y(t) = \sum_{j=-\infty}^{\infty} y_j \frac{\sin\{\omega_c(t-j)\}}{\omega_c(t-j)} = \sum_{j=-\infty}^{\infty} y_j \phi_c(t-j), \quad (31)$$

where j is an index that demarcates time intervals of $s = \pi/\omega_c > 1$ units. In such circumstances, sampling at unit intervals, instead of sampling at the wider intervals of s units, is liable to be described as oversampling.

The set of sinc functions $\{\phi_c(t-j); k \in \mathcal{Z}\}$ constitutes an orthogonal basis for functions limited to the interval $[-\omega_c, \omega_c]$. Let the sequence $\{h_t; t = 0, \pm 1, \pm 2, \dots\}$ be the ordinates sampled from the function $\phi_c(t)$ at unit intervals of time. Then, according to the Shannon–Nyquist theorem, there is

$$\phi_c(t) = \sum_{k=-\infty}^{\infty} h_k \phi(t-k). \quad (32)$$

Putting this expression into (31) gives the expression for $y(t)$ in terms of the orthogonal basis corresponding to the Nyquist interval:

$$y(t) = \sum_{j=-\infty}^{\infty} y_j \left\{ \sum_{k=-\infty}^{\infty} h_k \phi(t-j-k) \right\}. \quad (33)$$

The comparison of this expression with the more parsimonious expression of (24) shows the advantage of employing a set of basis functions that cover the same frequency interval as the function $y(t)$.

The detriment of oversampling will be illustrated both via an empirical example and in sampling experiments. Its consequences will also be analysed from the point of view of the time domain with reference to the autocovariance function.

For the present, we may imagine that the process $y(t)$ is supported on the frequency interval $[-\pi/2, \pi/2]$, and that it has been sampled at unit intervals. Without loss of generality, we may locate the sinc functions $\phi_2(t-k) = \sin(\pi([t-k]/2)\pi t)$ of the appropriate basis on the points $\{k = 0, \pm 2, \pm 4, \dots\}$. This is illustrated in Figure 10.

When the observations are taken at successive integer points, the values sampled at $k \in \{0, \pm 2, \pm 4, \dots\}$, will correspond to the amplitudes of individual sinc functions, whereas those sampled at $k \in \{1, \pm 3, \pm 5, \dots\}$ will comprise values sampled from all of the basis functions.

The effect of oversampling, represented by the samples taken at the odd integers, may be described as interference. Interference is a counterpart to aliasing. Aliasing arises when the data are sampled at too low a rate, which is when it can be said that they are undersampled.

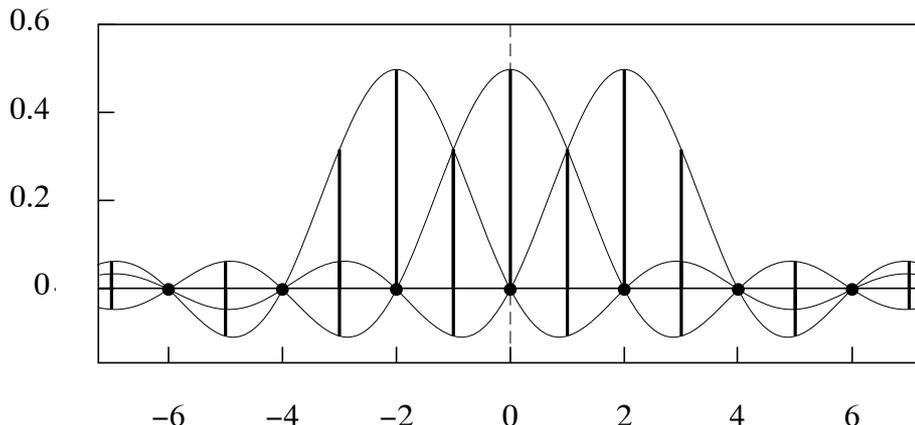


Figure 10. The wave packets $\phi_2(t)$ and $\phi_2(t - k)$ suffer no interference when $k \in \{\pm 2, \pm 4, \pm 6, \dots\}$.

The problems of interference and of aliasing, taken together, suggest that, to obtain estimates of the ARMA parameters that are free of biases, it will be necessary to sample the data at exactly the critical rate. However, the problem of oversampling can be overcome by resampling the data.

First, the underlying continuous process may be reconstituted from the nonzero Fourier ordinates of the band-limited process by the method of Fourier interpolation, which is described by equation (1) and which is illustrated in Figure 2. Then, the reconstituted function can be sampled at intervals of $s = \pi/\omega_c$ time units, where ω_c is the maximum frequency within the underlying process.

5. Alternative Estimates of an AR(2) Process

In this section, we shall demonstrate, via a practical example, the consequences of ignoring the band-limited nature of the underlying process. In the first instance, we shall also omit to cleanse the data of some residual noise contamination before applying a conventional estimator of a second-order autoregressive model. The consequence is that estimator delivers an autoregressive operator with real-valued roots, where we would expect to find conjugate complex roots.

In the second instance, we shall cleanse the data completely by eliminating all elements with frequencies in excess of the maximum value within the low-frequency component of interest. In this case, the estimated autoregressive operator does contain complex roots. The argument of the roots corresponds roughly to the periodicity of the dominant cycles within the data. However, their modulus assumes a value that is close to unity, which underestimates the damping of the cycles.

We shall complete the section by demonstrating an appropriate estimation procedure, which comprises both the cleansing of the data and their resampling. We shall also offer an explanation for the conflicting outcomes of the two cases, described above, where the data are not resampled.

The data in question will be the detrended consumption sequence that is depicted in Figure 2. The autoregressive estimator of choice is the least-squares estimator of Whittle (1951, 1962). This is just a case of the Yule–Walker estimator with circular empirical autocovariances.

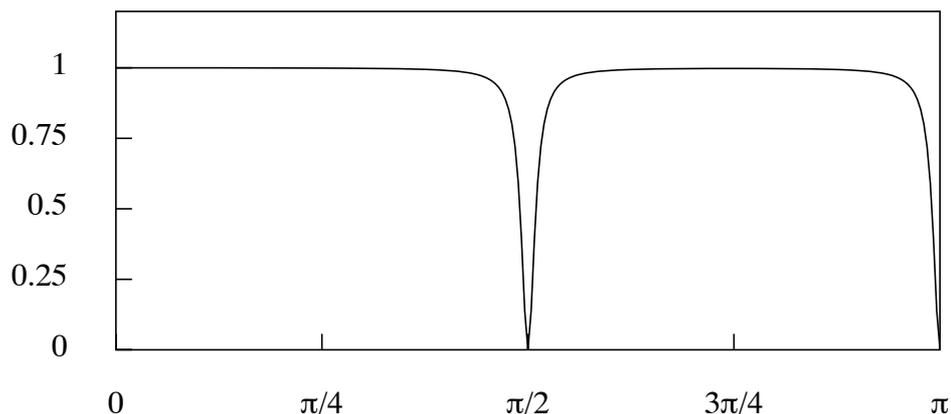


Figure 11. The squared gain of a seasonal adjustment filter to be applied to the quarterly detrended logarithmic consumption data.

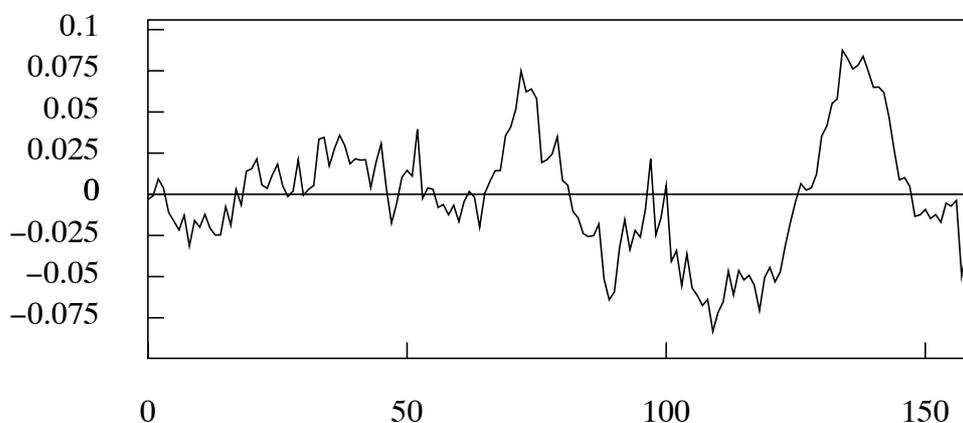


Figure 12. The seasonally-adjusted detrended logarithmic consumption data.

In the first approach, we work with seasonally-adjusted data. To obtain such data, a filter is used that mimics the processes of seasonal adjustment that occur within national central statistical offices. The frequency response of this filter is shown in Figure 11. Figure 12 shows the seasonally-adjusted version of the data sequence of Figure 2, and Figure 13 shows the extracted seasonal component.

The remarkable regularity of the seasonal component, which is not unusual among such estimates, is an artefact of the filter that is complementary to the seasonal-adjustment filter. This filter extracts from the data the elements at the seasonal frequencies $\pi/2$ and π , together with a small proportion of what lies at the adjacent frequencies.

The periodogram of the seasonally-adjusted data is shown in Figure 14. In the interval $(\pi/8, \pi]$, which lies beyond the upper limit of the spectrum of the low-frequency business cycle component, there are minor traces of a contaminating noise, which serves to roughen the profile of the seasonally-adjusted data. The smooth profile of the business cycle, from which this noise is absent, is shown in Figure 2.

When it is applied to the seasonally-adjusted data, the Whittle estimator delivers an AR(2) model of which the parametric spectrum is shown in Figure 15. The

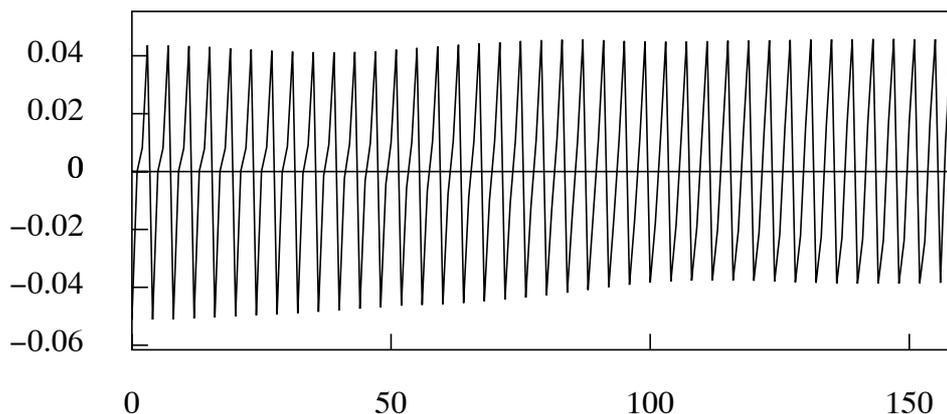


Figure 13. The seasonal component extracted from the detrended logarithmic consumption data in the process of seasonal adjustment.

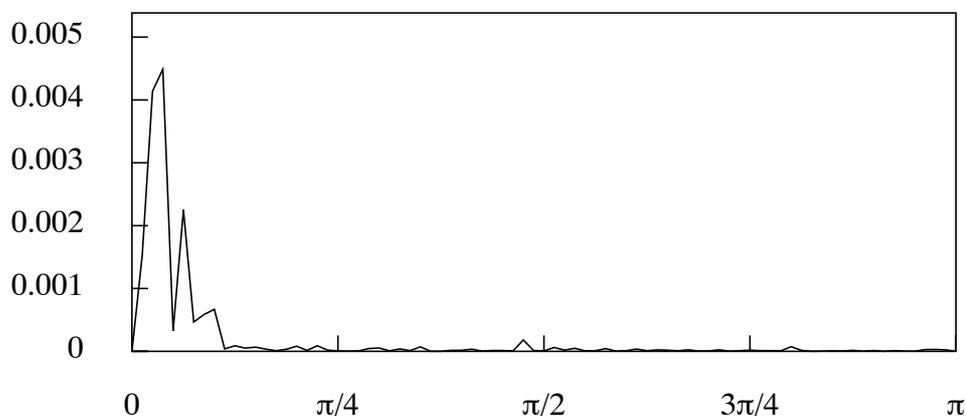


Figure 14. The periodogram of the seasonally-adjusted data.

roots of the estimated autoregressive operator, which are real-valued, are shown on the left side of Figure 17. This outcome is due to the presence of the noise contamination in the seasonally-adjusted data; and its removal leads to very different estimates.

The noise contamination can be removed from the data by setting to zero the Fourier ordinates that lie in the interval $(\pi/8, \pi]$. The remaining ordinates are employed in the Fourier synthesis that has given rise to the continuous business cycle trajectory of Figure 2. Sampling this trajectory at the integer points gives rise to the cleansed data to which an AR(2) is fitted.

The parametric spectrum of this model, which is shown in Figure 16, has a prominent spike at a frequency that corresponds to the business cycle. The complex roots of the autoregressive operator are shown on the right side of Figure 17, where it can be seen that they are located virtually on the boundary of the unit circle. The cycles generated by such a model would be subject to very little damping, and they would show properties of regularity and persistence that are not found within the data.

The appropriate estimator of the business cycle parameters is one that takes account only of the information within the corresponding spectral structure and

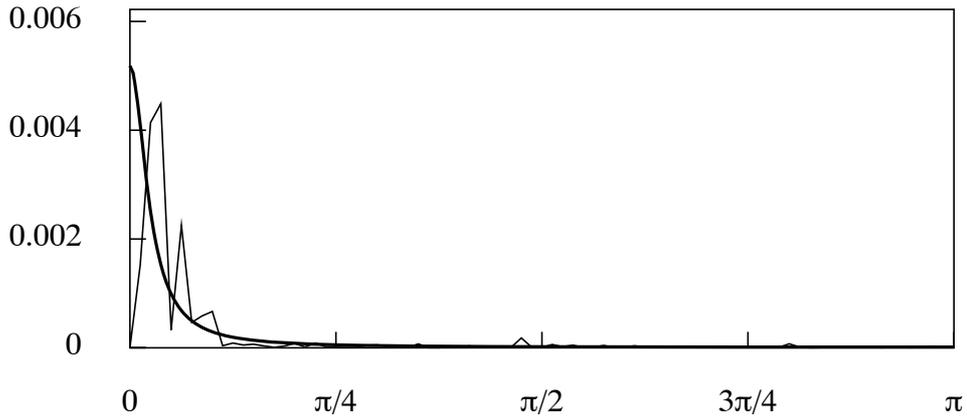


Figure 15. The spectrum of an AR(2) model fitted to the detrended, seasonally-adjusted logarithmic consumption data, superimposed on the periodogram.

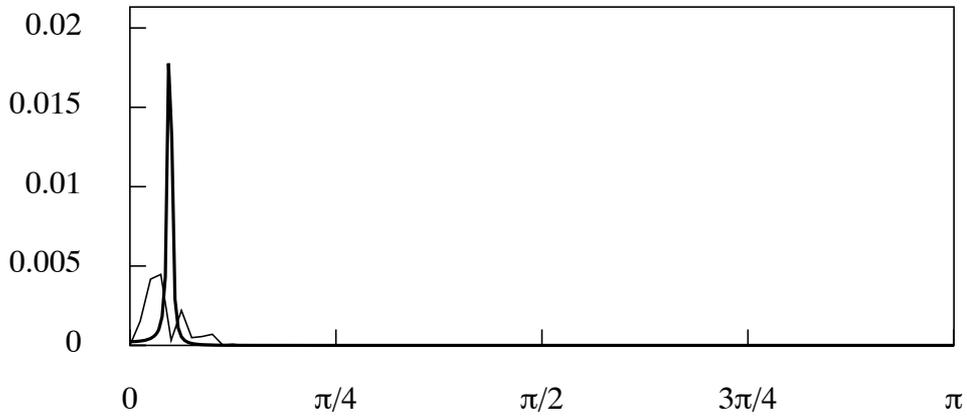


Figure 16. The spectrum of an AR(2) model fitted to the detrended consumption data cleansed of the elements of frequencies in excess of $\pi/8$.

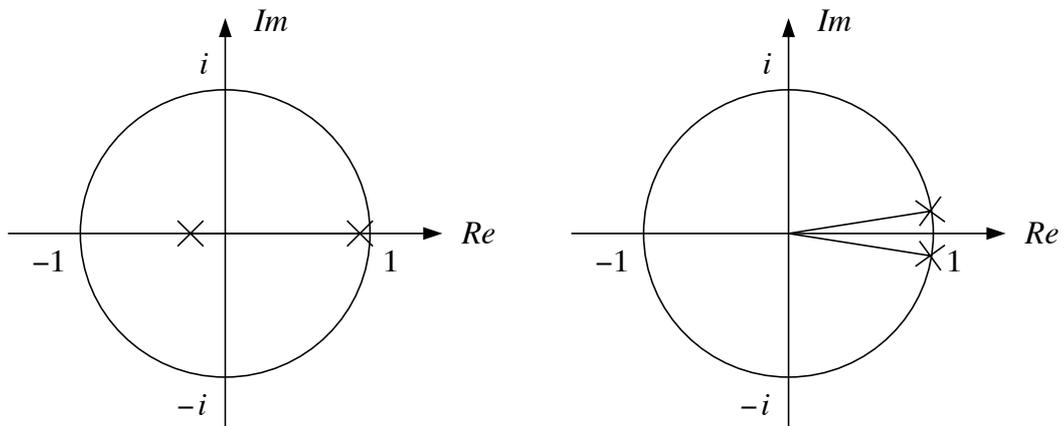


Figure 17. The poles of the AR(2) models fitted to the detrended logarithmic consumption data. (a) is from unrestricted estimator and (b) is from the band-limited estimator.

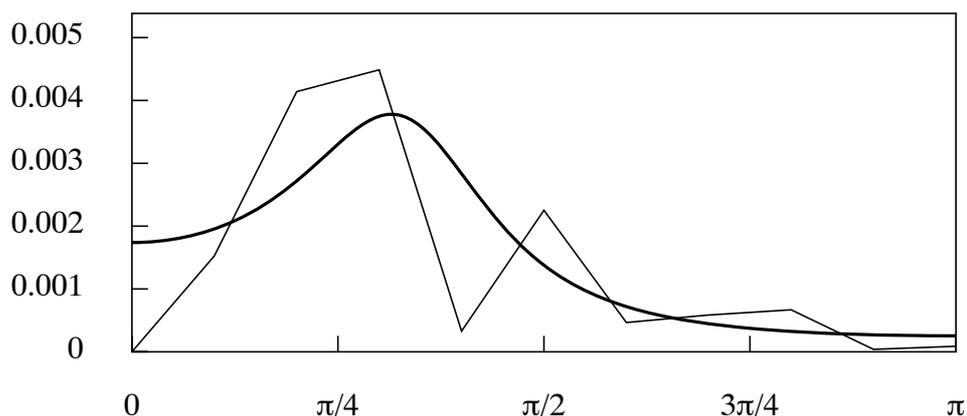


Figure 18. The periodogram of the subsampled anti-aliased data with the parametric spectrum of an estimated AR(2) model superimposed.

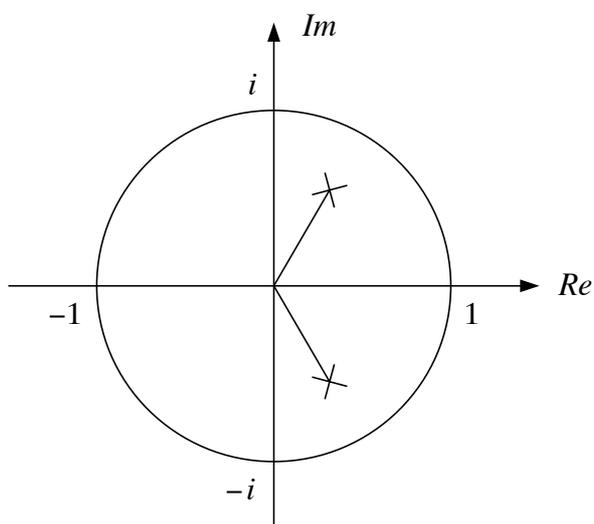


Figure 19. The poles of the AR(2) model fitted to 20 points subsampled at the rate of 1 in 8 from data that has been subjected to an anti-aliasing lowpass filter with cut off at $\pi/8$ radians.

which maps this structure into the interval $[0, \pi]$, which is the domain of an ordinary ARMA model. To achieve this outcome, the data are sampled from the trajectory of Figure 2 at 1/8th of the original rate of observation.

The periodogram of the cleansed and sub-sampled data is show in Figure 18, with the parametric spectrum of an estimated AR(2) model superimposed. The periodogram represents a rescaled version of the part of the periodogram of Figure 6, pertaining to the original data, that occupies the frequency range $[0, \pi/8]$; and it appears to be well represented by the parametric spectrum.

The parameters of the fitted model are recorded in Table 1. The roots of the autoregressive operator, which are also displayed in Figure 19, appear to give an appropriate representation of the dynamic properties of the business cycle, both as regards its frequency and its damping characteristics.

Table 1. The parameters of the operator $(1 + \alpha_1 z + \alpha_2 z^2) = (1 - \lambda_1 z)(1 - \lambda_2 z)$, of an AR(2) model estimated from an empirical data sequence that has been subsampled at the rate of 1 in 8 points after the application of a lowpass anti-aliasing filter with a cut-off point at $\pi/8$ radians or 22.5 degrees. The complex-valued, the roots may be expressed as $\lambda_1, \lambda_2 = \gamma \pm i\delta = \rho \exp\{\pm\theta\}$.

$$\begin{array}{lll} \alpha_1 = -1.63000 & \lambda_1 = 0.3150 + i0.5456 & \rho = 0.63 \\ \alpha_2 = 0.3969 & \lambda_2 = 0.3150 - i0.5456 & \theta^\circ = 60.0 \end{array}$$

The parametric spectrum of the fitted AR(2) model differs in one significant respect from the periodogram of the data. As a consequence of the successful detrending of the data, the periodogram has a zero-valued ordinate at zero frequency. By contrast, the intercept of the spectrum with the vertical axis has a large value, which tends to misrepresent the data.

The aberrant results of the first two experiments, in which the data are subject to oversampling, can be explained by reference to the autocovariance function. When the rate of sampling is excessive, the autocovariances will be sampled at points that are too close to the origin, where the variance is to be found. In the absence of noise contamination, their values will decline at a diminished rate. The reduction in the rate of convergence is reflected in the modulus of the estimated complex roots, which understates the rate of damping.

When there is a contamination that extends across the range of frequencies, its variance will be added to the variance of the underlying process. Virtually nothing will be added to the adjacent sampled ordinates of autocovariance function. Therefore, the initial sampled autocovariances will decline at an enhanced rate. If this rate of convergence exceeds the critical value, then there will be a transition from cyclical convergence to monotonic convergence, and the estimated autoregressive roots will be real-valued.

6. Simulation Experiments

The results that have been demonstrated with empirical data can be reaffirmed by some sampling experiments based on pseudo-random data. The outcomes of the experiments are readily intelligible when they are presented in a graphical form.

For these experiments, we adopt a band-limited model that corresponds to the one that is represented by the parametric spectrum of Figure 18, of which the parameters are recorded in Table 1. This AR(2) model, which has been estimated from biannual data, has conjugate complex roots with arguments of $\pm\pi/3$ radians or $\pm 60^\circ$ and a modulus of 0.63.

The continuous process, generated from these parameters, may be sampled at a rate corresponding to one observation per quarter to produce a band-limited sequence that is supported on the frequency interval $[0, \pi/8]$. The band-limited process has an angular velocity of $\pi/24$ radians or 7.5° per quarter and an effective modulus or damping factor of 0.944. Each of the three experiments represented by Figures 20–22 is based on 20 sequences of 160 observations, which have been generated using a white-noise forcing function of unit variance.

In the first experiment, AR(2) models have been fitted to data that are free of contamination. The results are shown in Figure 20. In this case, the majority of the fitted models deliver conjugate complex roots. The complex roots have

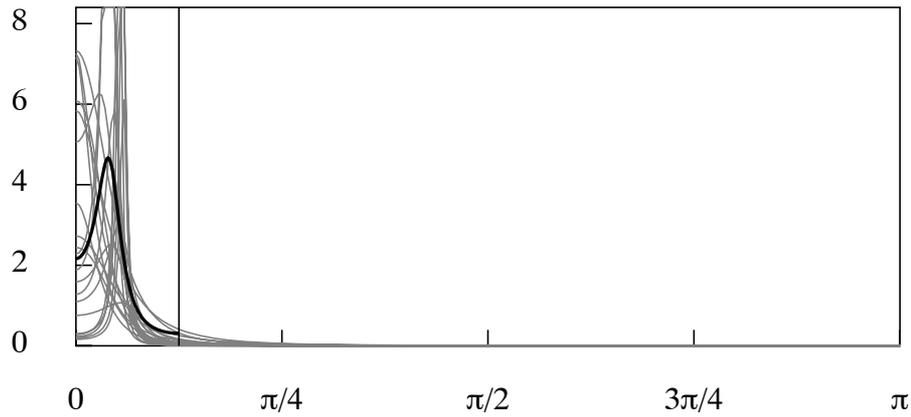


Figure 20. The spectra of 20 AR(2) models estimated from band-limited data supported on the interval $[0, \pi/8]$. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

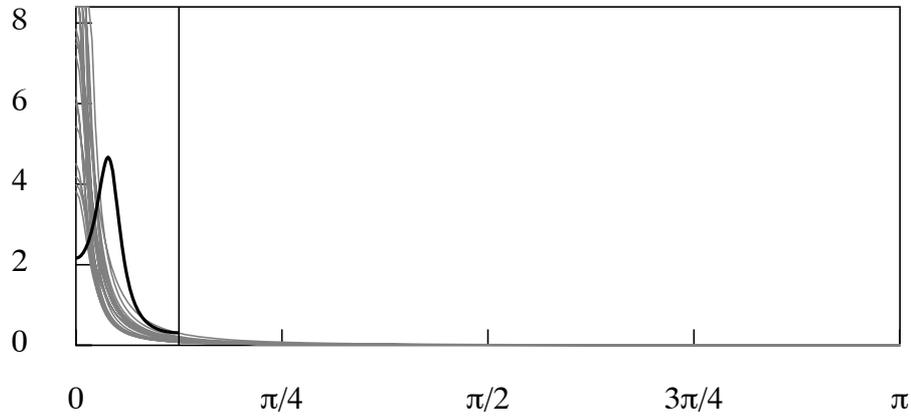


Figure 21. The spectra of 20 AR(2) models estimated from band-limited data contaminated by white noise. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

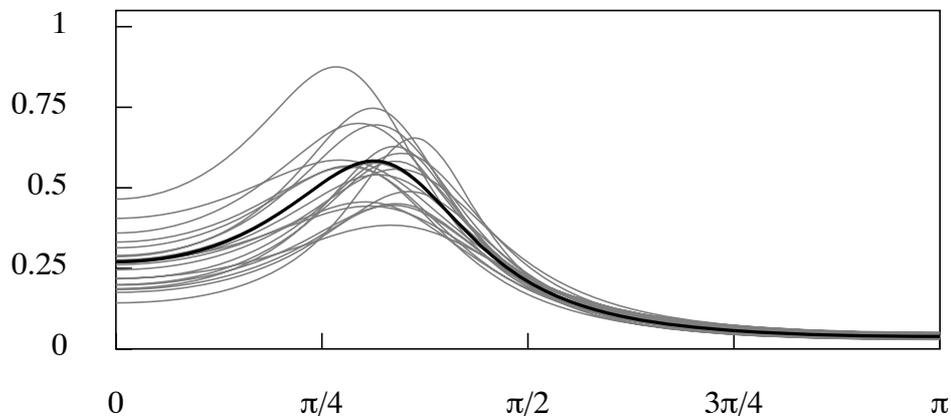


Figure 22. The spectra of 20 AR(2) models estimated from band-limited data subsampled at the critical rate. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

moduli that considerably exceed the value of the damping factor of the band-limited process. Figure 20 may be compared with Figure 16, where the data in that case are supported of the interval $[0, \pi/8]$, which is the pass band of a filter that was applied, by implication, to the seasonally-adjusted data.

Amongst the estimated models, there are a fair number that have real-valued roots. Their presence is a testimony to the fact that there is a considerable spectral mass in the vicinity of zero frequency in the parametric spectrum of the model that has been used in generating the band-limited pseudo-random data.

In the second experiment, white-noise contaminations have been added to the pseudo-random band-limited data. Their variance is equal to 10% of the variance of the band-limited AR(2) process. The parametric spectra of the fitted AR(2) models are plotted in Figure 21 and the spectrum of the band-limited model used in generating the data is superimposed upon them. None of the estimated models delivers complex roots. Figure 21 may be compared with Figure 15, which shows the effect of fitting an AR(2) model to some seasonally-adjusted empirical data.

In the third experiment, AR(2) models are fitted directly to the sets of 20 subsampled data points that underlie the band-limited sequences and that summarise their information. Such points would also be obtained by taking one in every eight of the points of the data sequences of the first experiment. On average, the fitted models slightly underestimate the modulus of the complex roots of the AR(2) process that has been used in generating the data. This reflects the well know bias of the least-squares estimator in small samples.

These various outcomes are reflected in some further sampling experiments of a numerical nature, of which the results are partially recorded in Table 2. Here, the estimation procedures have been applied to 10,000 mutually independent pseudo-random sequences of band-limited data generated by the AR(2) model.

Table 2. The averages of 10,000 replications of the parameter estimates of an AR(2) model, derived from 160 sample points of a band-limited process.

<i>The Data Processes</i>	<i>No. of points</i>	<i>Roots</i>	<i>Modulus</i>	<i>Argument Degrees</i>
Band limited AR(2) with subsampling	20	$0.2569 \pm i0.5510$	0.6161	± 65.36
Band limited AR(2) without subsampling	160	$0.8545 \pm i0.1039$	0.8608	6.9349
Band limited AR(2) with noise	160	0.9531, 0.1289	—	—

The results that are obtained when an ordinary AR(2) model is fitted to the uncontaminated data bear some further investigation. Of the 10,000 fitted models, 6,412 comprised complex-valued autoregressive roots and 3,588 comprised real-valued roots. By averaging the estimated parameters in the two categories, some highly contrasting results are obtained, which are show in Table 3.

Table 3. The averaged results of 10,000 estimations of the parameters of an AR(2) operator $(1 + \alpha_1 z + \alpha_2 z^2) = (1 - \lambda_1 z)(1 - \lambda_2 z)$, derived from 160 sample points of a band-limited process. When they are complex-valued, the roots may be expressed as $\lambda_1, \lambda_2 = \gamma \pm i\delta = \rho \exp\{\pm i\theta\}$.

The Average of all 10,000 Estimates

$$\begin{array}{lll} \alpha_1 = -1.7089 & \lambda_1 = 0.8545 + i0.1039 & \rho = 0.8606 \\ \alpha_2 = 0.7409 & \lambda_2 = 0.8545 - i0.1039 & \theta^\circ = 6.939 \end{array}$$

The Average of 6,412 Estimates with Complex-Valued Roots

$$\begin{array}{lll} \alpha_1 = -1.8451 & \lambda_1 = 0.9226 + i0.1508 & \rho = 0.9348 \\ \alpha_2 = 0.8739 & \lambda_2 = 0.9226 - i0.1508 & \theta^\circ = 9.2846 \end{array}$$

The Average of 3,588 Estimates with Real-Valued Roots

$$\begin{array}{ll} \alpha_1 = -1.4655 & \lambda_1 = 0.9161 \\ \alpha_2 = 0.5033 & \lambda_2 = 0.5494 \end{array}$$

Only a limited set of experiments have been reported in this section. Those who wish to explore matters further may do so with the help of a computer program, BLIMDOS.PAS, which can be downloaded from a website at the following address:

<http://www.le.ac.uk/users/dsgp1/>

7. Conclusions and Remarks

In this paper, we have presented a model for a band-limited stochastic process in continuous time and we have demonstrated that, if the rate of sampling corresponds to the maximum frequency within the data, then a sequence of sampled ordinates can be described by an ordinary discrete-time ARMA model. In these circumstances, it may be said that the sampling is at the critical rate.

We have concentrated on the case where the sampling is over-rapid and where we can afford to reduce it, either by a process of subsampling or by resampling a reconstituted version of the continuous signal. There are also cases to be considered where the sampling rate is less than the maximum frequency of the data and where there is an inevitable problem of aliasing. Such circumstances are already accounted for in the existing literature under the rubric of temporal aggregation.

The available theory indicates that, if the data can be described, at the critical rate of sampling, by an AR(p) autoregressive model of order p , then the subsampled data will be described by an ARMA($p, p-1$) model. The only proviso here is that, in both cases, the sampling rate should be sufficiently rapid to accommodate the dynamics implied by the autoregressive operator. Thus, the maximum value ω_m of the arguments associated with the roots of the operator must be less than the Nyquist frequency of π radians per sample period.

Telser (1967) derived this result in the context of what he described as the skip sampling of a discrete-time autoregressive process, whereas Phadke and Wu (1974) and Pandit and Wu (1975) did so by considering the relations between continuous-time linear stochastic processes driven by the increments of a Wiener process and the models that can be fitted to the sampled data.

Our empirical example of a band-limited process is complicated by the fact that it is a component of a composite process, which includes a trend. In this

case, we have been able to represent the trend by a quadratic function in the logarithmic data, which is virtually a linear function corresponding to a trajectory of constant exponential growth. However, in general, we can expect that the orderly development of the data will be disrupted, occasionally, by structural breaks and outliers.

Such breaks will have a spectral traces that extend over the entire frequency range; and these will tend to obscure the band-limited nature of the predominant processes that underlie the data. This may be one of the reasons why econometricians have tended to overlook an essential feature of many of their data sequences, which is the localisation of the frequency bands in which their component parts reside.

References

- Pandit, S.M., and S.M. Wu, (1975), Unique Estimates of the Parameters of a Continuous Stationary Stochastic Process, *Biometrika*, 62, 497–501.
- Pagan, A., (1997), Towards an Understanding of Some Business Cycle Characteristics, *The Australian Economic Review*, 30, 1–15.
- Phadke, M.S., and S.M. Wu, (1974), Modelling of Continuous Stochastic Processes from Discrete Observations with Applications to Sunspot Data, *Journal of the American Statistical Association*, 69, 325–329.
- Shannon, C.E., (1949), Communication in the Presence of Noise, *Proceedings of the Institute of Radio Engineers*, 37, 10–21. Reprinted in 1998, *Proceedings of the IEEE*, 86, 447–457.
- Telser, L.G., (1967), Discrete Samples and Moving Sums in Stationary Stochastic Processes, *Journal of the American Statistical Association*, 62, 318, 484–499.
- Whittaker, J.M., (1935), *Interpolatory Function Theory*, Cambridge Tracts in Mathematics and Mathematical Physics, no. 33. Cambridge, U.K., Cambridge University Press.
- Whittle, P., (1951), *Hypothesis Testing in Time Series Analysis*, Almqvist and Wiksells Boktryckeri, Uppsala.
- Whittle, P., (1962), Gaussian Estimation in Stationary Time Series, *Bulletin of the International Statistical Institute*, 39, 105–129.