Spurious Periodicity in Inappropriately Detrended Time Series

Charles R. Nelson; Heejoon Kang

Econometrica, Volume 49, Issue 3 (May, 1981), 741-751.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198105%2949%3A3%3C741%3ASPIIDT%3E2.0.CO%3B2-Y

Econometrica, Vol. 49, No. 3 (May, 1981)

SPURIOUS PERIODICITY IN INAPPROPRIATELY DETRENDED TIME SERIES[1]

BY CHARLES R. NELSON AND HEEJOON KANG

The expected sample autocorrelation function of residuals from regression of a random walk on time is shown to imply strongly pseudo-periodic behavior in the "detrended" series. The shape of the autocorrelation function is effectively independent of sample size and the corresponding spectral density function has a single peak at a period equal to .83 of sample size; thus the apparent dynamic properties of the residuals are artifactual. Sampling experiments are used to describe the distribution of the spectral peak and further suggest that nonzero autocorrelation in first differences of the raw data will have little effect on the spurious periodicity phenomenon.

1. INTRODUCTION

ECONOMETRIC ANALYSIS of time series data is frequently preceded by regression on time to remove a trend component in the data. The resulting residuals are then treated as a stationary series to which procedures requiring stationarity, such as spectral analysis, can be applied. The objective is often to investigate the dynamics of transitory movements in the system, for example, in econometric models of the business cycle.[2] When the data do consist of a deterministic function of time plus a stationary series, then regression residuals will clearly be unbiased estimates of the stationary component. However, Hannan [2] has shown that the estimated spectral density function will be distorted by trend removal and Hatanaka and Howrey [3] demonstrate the potential for spurious peaks arising therefrom. If, on the other hand, the data are generated by (possibly repeated) summation of a stationary and invertible process, then the series cannot be expressed as a deterministic function of time plus a stationary deviation, even though a least squares trend line and the associated residuals can always be calculated for any given finite sample. In a recent paper, Chan, Hayya, and Ord [1] (hereafter CHO) were able to derive the autocovariance function of residuals from linear regression of a realization of a random walk (the summation of a purely random series) on time. Since the autocovariances for a given lag are a function of time, the residuals are not stationary. Further, CHO established that the expected sample autocovariance function (the expected autocovariances for a given lag averaged over the time interval of the sample) depends on sample size and lag and is therefore an artifact of the detrending procedure. This function is characterized by CHO in their Figure 1 as being effectively linear in lag (although the exact function is a fifth degree polynomial) with the rate of decay from unity at the origin depending inversely on sample size. The first

[1]This research was carried out under a grant from the National Science Foundation. Much of the work was done while Nelson was a participant in the University of Warwick Summer Workshop (1979), sponsored by the Social Science Research Council (U.K.), and many helpful comments and suggestions were received from other participants, particularly R. Engle and C. W. J. Granger. We are likewise indebted to S. Beveridge, J. K. Ord, as well as referees and an editor of this journal. The authors are responsible for any errors.

[2]Inclusion of a time trend in a regression between variables of interest is of course equivalent to detrending of the variables prior to regression (Lovell [5]).

differences of a random walk are, of course, stationary with zero autocovariance at all lags. They concluded that "the low frequency portion of the spectrum will be exaggerated and the high frequency portion attenuated" relative to the appropriate first difference transformation.

The objective of this paper is to show that after the expression given by CHO for the expected sample autocovariance function is corrected for errors in the values of coefficients and is examined over a greater range of lags it is seen to imply strongly pseudo-periodic behavior in the time trend residuals. The corresponding spectral density function has a single peak at a period corresponding to .83 of the number of observations in the sample. The distribution of the peak in sample power spectra is studied in a Monte Carlo experiment and is shown to have a mean period corresponding to .65 of sample length with a standard deviation of .21 of sample size. The presence of serial correlation in first differences is shown to shift the location of this peak in the direction intuition would suggest but not alter its essential dominance of the sample spectrum. Our results suggest then that inappropriate detrending of time series will tend to produce apparent evidence of periodicity which is not in any meaningful sense a property of the underlying system. They further suggest that the dynamics of econometric models estimated from such data may well be wholly or in part an artifact of the trend removal procedure.

2. THE EXPECTED SAMPLE AUTOCOVARIANCE FUNCTION AND APPROXIMATE EXPECTED AUTOCORRELATION AND SPECTRAL DENSITY FUNCTION FOR TIME TREND RESIDUALS FROM A RANDOM WALK

The autocovariance function for the residuals produced by regression of a realization of a random walk on time has been derived by CHO (their expression 3.10) and for a given lag it is a function of time and the number of observations in the sample. The residuals are therefore nonstationary, but it is straightforward to derive the expected value of the sample autocovariance function as the average across the sample of the autocovariances for given lag. The expression given by CHO (3.13) for this averaged autocovariance contains several numerical errors which materially affect its shape. The corrected expression in the notation of CHO is

$$
\begin{aligned}
(2.1)\qquad \overline{\text{cov}}(s,n) &= \frac{1}{N}\sum_{t=-n+s}^{n}(\hat{\epsilon}_t,\hat{\epsilon}_{t-s}) \\
&= \big\{(32n^5+80n^4+40n^3-20n^2-12n) \\
&\quad + s(-152n^4-304n^3-150n^2+2n+6) \\
&\quad + s^2(180n^3+270n^2+90n) \\
&\quad + s^3(-68n^2-68n-9) \\
&\quad + s^5(3)\big\}\,\frac{\sigma^2}{60n(n+1)(2n+1)^2}
\end{aligned}
$$

where $N = 2n + 1$ is the length of the sample and σ^2 is the variance of the steps $\{\epsilon_t\}$ in the underlying random walk. We follow CHO in approximating the expected sample autocorrelation function by $\bar{\rho}(s,n) \equiv \overline{\text{cov}}(s,n)/\overline{\text{cov}}(0,n)$ which can be written as

$$(2.2) \qquad \bar{\rho}(s,n) = 1 + \left(\frac{s}{n}\right)\left(\frac{-152 - 304n^{-1} - 150n^{-2} + 2n^{-3} + 6n^{-4}}{32 + 80n^{-1} + 40n^{-2} - 20n^{-3} - 12n^{-4}}\right)$$
$$+ \left(\frac{s}{n}\right)\left(\frac{180 + 270n^{-1} + 90n^{-2}}{32 + 80n^{-1} + 40n^{-2} - 20n^{-3} - 12n^{-4}}\right)$$
$$+ \left(\frac{s}{n}\right)^3\left(\frac{-68 - 68n^{-1} - 9n^{-2}}{32 + 80n^{-1} + 40n^{-2} - 20n^{-3} - 12n^{-4}}\right)$$
$$+ \left(\frac{s}{n}\right)^5\left(\frac{3}{32 + 80n^{-1} + 40n^{-2} - 20n^{-3} - 12n^{-4}}\right).$$

It is apparent that as n gets large the coefficients of the powers of (s/n) in (2.2) approach constants and hence $\bar{\rho}$ may be approximated as a fifth degree polynomial which will take on the same value at any given fraction of sample size. Maxima and minima will therefore occur at roughly the same fraction of sample length and with the same values for $\bar{\rho}$ regardless of sample length. The exact function is plotted in Figure 1 for sample length 101 ($n = 50$). Note that the

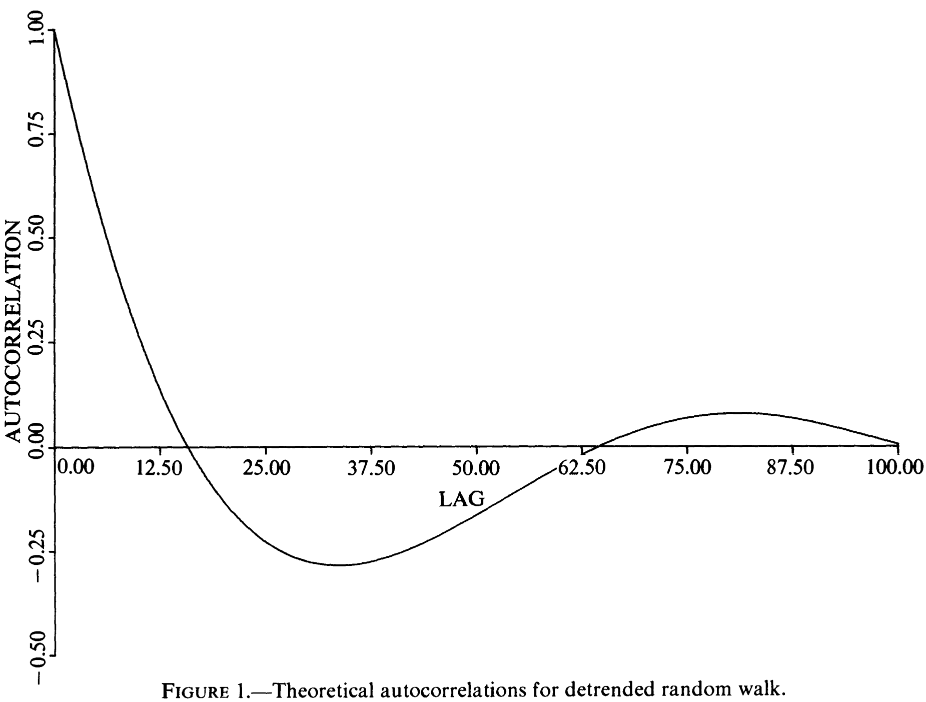

FIGURE 1.—Theoretical autocorrelations for detrended random walk.

TABLE I

APPROXIMATION TO EXPECTED SAMPLE AUTOCORRELATION FUNCTION FOR SAMPLE LENGTH $N = 101$ ($n = 50$): CORRECTED FOR CHO EXPRESSION

Lag s	1	2	3	4	5	10	20	30	50	75
$\bar{\rho}(s)$	.91	.82	.74	.66	.58	.26	−.13	−.28	−.16	.07
CHO	.93	.87	.81	.75	.70	.47	.19	.10	.16	.20

autocorrelations decline monotonically until a value of $-.28$ is reached at lag 34 (that is .34 of sample length), then increase until a value of .08 is reached at lag 81 (.80 of sample length) and finally decline toward zero. For comparison, in a sample of length 11 ($n = 5$) the minimum occurs at lag 4 (.36 of sample length) with a $\bar{\rho}$ value of $-.29$ and the second maximum at lag 9 (.82 of sample length) with a value of .08. For practical purposes then the shape of the $\bar{\rho}$ function is effectively independent of sample size.[3] The incorrect expression for $\bar{\rho}(s,n)$ given by CHO does not decline as rapidly at low lags, the slope for small values of (s/n) being roughly -3.5 for the CHO expression as opposed to -4.75 for the corrected expression. The CHO expression reaches a first minimum of $+.09$ at about .33 of sample length and a second maximum of $+.22$ at about .66 of sample length before declining toward zero. The figures provided by CHO show only the linear decline which characterizes the function for small values of (s/n). The corrected function is more strongly nonlinear even at low lags since the coefficient of $(s/n)^2$ is now larger. Calculated values for the CHO and the corrected $\bar{\rho}(s)$ at selected lags are compared in Table I for a sample size of 101.

The corrected expected autocovariance function has the appearance of a damped sine wave which is indicative of pseudo-periodic behavior in the residual series with a period equal to .80 of the length of the sample or equivalently a frequency of $1.25N^{-1}$. The sample spectral density function corresponding to $\bar{\rho}(s,n)$ is defined as $\mathrm{SDF}(f,n) = 1 + 2\sum_{s=1}^{N-1}\bar{\rho}(s,n)\cos(2\pi fs)$, $0 < f < \frac{1}{2}$, and is plotted in Figure 2 for $N = 51$ and 101 over $0 < f < .10$. It has its maximum at frequency .024 for $N = 51$ and at .012 for $N = 101$ each corresponding to $1.25N^{-1}$. Note that as the sample size is increased the value at the maximum of the spectral density increases and the entire distribution becomes more concentrated at lower frequencies. It is shown in the Appendix that the value of the function at frequency $f = 0$ is identically zero, while for the corresponding function implied by the CHO expression the value at $f = 0$ is of order n with leading term $(5/6)n$ and thus rises roughly in proportion to sample size. The density function values at $f = \frac{1}{2}$ in both cases decrease roughly in proportion to sample size, leading terms being $(5/2)n^{-1}$ and $(15/8)n^{-1}$ for corrected and CHO versions respectively. The graphical characterization displayed by CHO in

[3] Stephen Beveridge has pointed out to us that since for given n $\bar{\rho}(s,n)$ is a fifth degree polynomial in s it follows that $(1 - L)^6\bar{\rho}(s,n) = 0$, $s \geqslant 7$, where L is understood to operate on s. Therefore $\bar{\rho}(s,n)$ follows a difference equation with coefficients that are independent of n; sample size only determines the initial conditions. Note also that the $\hat{\epsilon}$ must follow essentially the same difference equation with random error.

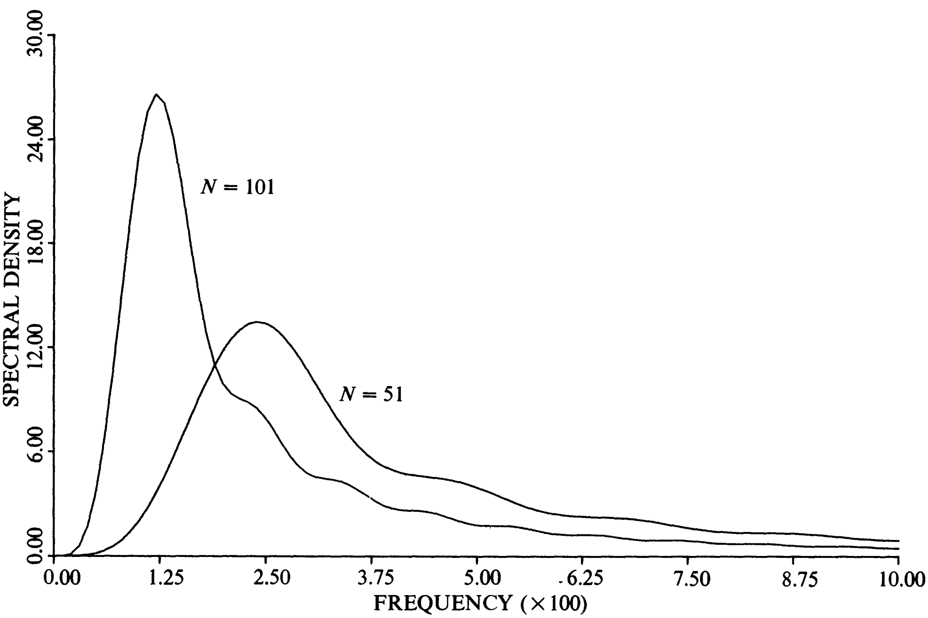

FIGURE 2.—Theoretical spectral density functions for detrended random walk.

their Figure 1 is of monotonic decline from the origin although a more detailed plot would show a prominent peak at a frequency of about $1.37N^{-1}$ in addition to a peak at the origin.

Sample estimates of spectral density functions are usually computed in practice from a sample autocorrelation function which is truncated at a lag considerably smaller than the maximum $(N - 1)$ lags computable. The effect of truncation of the $\bar{\rho}(s, n)$ function on the corresponding spectral density function is to reduce the prominence of the major peak, shift it to a higher frequency, and to introduce secondary peaks at higher frequencies which are not present in the complete function. For a sample of length 101 the primary peak is still in evidence using as few as 25 lags in the autocorrelation function and has shifted from a period of 83 to one of 62. When only 16 lags are used the prominent peak disappears and the function has its maximum value at the origin. In effect, the nonlinearity of $\bar{\rho}(s, n)$ and therefore its pseudo-periodicity is decreasingly evident at very low lags where the function is roughly exponential, resembling that of a first order autoregression. The spectral density function in that case also resembles that of a first order autoregression except for weak secondary peaks as noted above.

3. THE SAMPLING DISTRIBUTION OF PEAKS IN SAMPLE SPECTRAL DENSITY FUNCTION

The corrected theoretical results of CHO imply that one would expect to find a predominant peak at a low frequency in the sample spectrum of residuals obtained by regression of a random walk on time but do not lead readily to a description of the sampling distribution of the period or frequency at which the predominant peak occurs. We have conducted a Monte Carlo experiment to provide some idea of what this distribution looks like. Realization of length 101 observations were generated by a random walk process, regressed on time, and the sample autocorrelations and spectrum calculated. To investigate the effect of lag truncation on the sample spectrum, it was calculated in each case using 16, 25, 50, and all 100 available lags. Each sample spectrum was searched to find its maximum on a grid of .001 intervals starting at .010, the lowest frequency of practical interest. This was repeated for 500 independent realizations.

The columns of Table II report the mean and standard deviation of the frequency and period at which the maxima were observed in individual spectra, the per cent of cases in which the maximum occurred at the .01 lower boundary, and the characteristics of the mean sample SDF across the 500 runs. When all 100 available lags are used the mean frequency at which maxima occurred was .018, corresponding to a period of 56, while the mean period was 66, corresponding to a frequency of .015. The standard deviations were .008 for frequencies and 21.5 for periods. The distribution of peaks is centered then around higher frequencies and shorter periods than would be suggested by the theoretical approximation which had its peak at $f = .012$ and period 83. In 4 per cent of the cases the maximum occurred at $f = .010$. The highest frequency at which a maximum was recorded was at .065, a period of 15. The averaged sample autocorrelation function reached a low point of $-.23$ at lags 29 through 34 and a secondary maximum of $+.06$ at lags 78 through 87 compared with $-.28$ at lag 34 and $+.08$ at lag 81 for the theoretical approximation $\bar{\rho}$ given in (2.2). The averaged sample SDF reaches its peak at the same frequency (.012) and period (83) as the theoretical approximation, but with a lower value, 21 compared with 27. As the number of lags used to calculate the SDF is reduced to 50 and then 25 the distribution of peaks shifts further toward higher frequencies and shorter periods as the theoretical approximation had suggested. The averaged SDF confirms those tendencies. When only 16 lags are used, however, a large fraction of maxima are found at the low frequency bound. Hence the spurious periodicity phenomenon is greatly attenuated if lag truncation is carried far enough.

TABLE II

SUMMARY STATISTICS FOR SAMPLE SDF OF DETRENDED RANDOM WALKS

	Mean (Std. Dev.) of		Per Cent of Cases Peak	Averaged Sample SDF	
Lags Used	Freq. at Peak	Period at Peak	at $f = .01$	Period Freq. At Peak	Value At Peak
100	.018(.008)	66(21)	4%	.012/83	21.2
50	.018(.008)	63(18)	0%	.014/71	19.1
25	.019(.008)	61(25)	24%	.019/53	11.7
16	.016(.010)	79(30)	67%	.010/100	10.6

4. THE EFFECT OF SERIAL CORRELATION IN FIRST DIFFERENCES ON THE SAMPLE SPECTRUM

Analysis of the pure random walk case has indicated that as sample length increases the sample spectrum of time trend residuals will become increasingly concentrated around a peak at a frequency which decreases in proportion to the inverse of sample size. Descriptively, the residuals will tend to exhibit cycles of increasing length and increasing amplitude around a fitted trend line as the length of the sample is increased. Will these results still hold if first differences are serially correlated? Clearly, periodicities due to underlying autocorrelation in first differences will be of fixed period and amplitude and, hence, will contribute relatively less to total variance in longer samples. In the mixing of actual periodicities with the spurious periodicity associated with regression detrending we would expect as a rough characterization that positive autocorrelation in first differences would reinforce the spurious low frequency peak and negative autocorrelation would dilute it. This is confirmed in a second Monte Carlo experiment in which thirty independent samples of length 101 observations were generated by the integrated process

$$(4.1) \qquad X_t = X_{t-1} + D_t,$$

$$D_t = \phi D_{t-1} + \epsilon_t,$$

using successively 0, +.3, −.3, +.8, −.8 as values for ϕ. The random number generator was restarted with the same seed for each successive value of ϕ so that there was no sampling variation across values of ϕ.

The results of the experiment are summarized in Table III. They indicate that the spurious peak will dominate the spectrum in a sample of realistic length and for values of ϕ of the magnitude often encountered in practice. They confirm that the effect of underlying positive autocorrelation will be to shift the distribution of the peak towards lower frequencies (longer periods). The average of the sample spectra shows greater concentration of variance around its peak and slight shifting of the peak to lower frequencies. The correspondingly reverse implications are confirmed for the case of negative autocorrelation.

TABLE III

SUMMARY STATISTICS FOR SAMPLE SDF OF DETRENDED INTEGRATED PROCESSES

	Mean (Std. Dev.) of		Averaged Sample SDF	
ϕ	Freq. at Peak	Period at Peak	Freq./Period at Peak	Value at Peak
0.0	.021(.012)	57(22)	.014/71	15.7
0.3	.020(.012)	59(22)	.014/71	16.7
− 0.3	.021(.012)	57(22)	.014/71	14.8
0.8	.019(.009)	63(23)	.013/77	22.1
− 0.8	.022(.012)	56(22)	.014/71	10.9

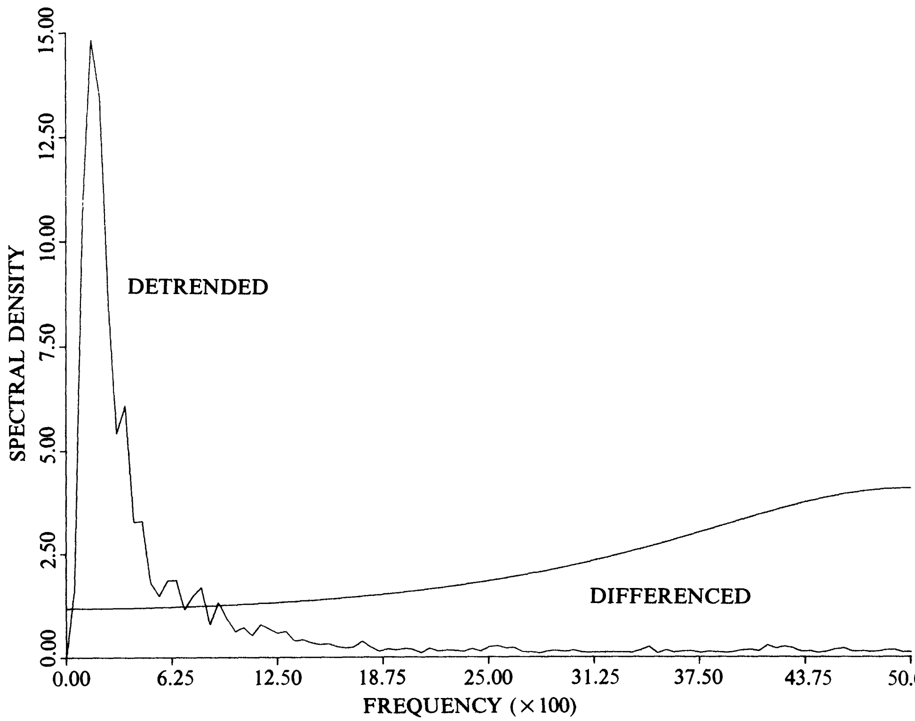

FIGURE 3.—SDF of integrated AR(1) process. First differences and trend residuals.

While the sample spectra do then exhibit characteristics traceable to the underlying autocorrelation structure of the first difference process, it seems unlikely that the latter would be detected in practice. For example, consider the comparison in Figure 3 of the averaged sample spectra of detrended data from (4.1) for $\phi = -.3$ with the theoretical spectral density of the first differences. Evidently, the high frequency power present in the theoretical spectrum is obliterated by the low frequency power introduced by regression detrending. It seems plausible to us that a major cost of inappropriate detrending may be that genuine dynamics will often be overlooked. Perhaps this explains in part why Granger [4] was able to characterize economic time series as having a "typical spectral shape" which was closely akin to that described in this paper as being typical of integrated time series which have been detrended by regression on time.

University of Washington

Manuscript received October, 1979; final revision received April, 1980.

APPENDIX

The following shows the detailed derivation of equation (2.1) and evaluation of the spectral density function at $f = 0$ and $f = 0.5$.

Rearranging the equation (3.10) in CHO, we have

(A.1)
$$\operatorname{cov}(\hat{\epsilon}_t, \hat{\epsilon}_{t-s}) = \frac{\sigma^2}{10n(n+1)(2n+1)} \Big\{ 10t^4 + (-20s)t^3 + (-8n^2 - 8n - 4 + 15s^2)t^2 + (8sn^2 + 8sn + 4s - 5s^3)t + \Big(\frac{10}{3}n^4 - 10sn^3 + \frac{20}{3}n^3 + \frac{10}{3}n^2 - 15sn^2 + 5s^2n^2 - 5sn + 5s^2n \Big) \Big\}.$$

To have

$$\overline{\operatorname{cov}}(s) = \frac{1}{N} \sum_{t=-n+s}^{n} \operatorname{cov}(\hat{\epsilon}_t, \hat{\epsilon}_{t-s}), \qquad \text{where} \quad N = 2n + 1,$$

we need the following:

$$\sum_{t=-n+s}^{n} t^0 = (2n + 1 - s),$$

$$\sum t^1 = \frac{1}{2} s(2n + 1 - s),$$

$$\sum t^2 = \frac{1}{6}(2n + 1 - s)(2n^2 + 2n - 2ns - s + 2s^2),$$

$$\sum t^3 = \frac{1}{4} s(2n + 1 - s)(2n^2 + 2n - 2ns - s + s^2),$$

$$\sum t^4 = \frac{1}{30}(2n + 1 - s)\{2n(3n^3 + 6n^2 + 2n - 1) + (-12n^3 - 18n^2 - 4n + 1)s + (24n^2 + 24n + 1)s^2 - 9(2n + 1)s^3 + 6s^4\}.$$

Substituting these into (A.1), we have equation (2.1) after some tedious but straightforward algebra. The spectral density function, SDF, is defined as

$$\operatorname{SDF}(f) = 1 + 2 \sum_{s=1}^{2n} \bar{\rho}(s, n)\cos(2\pi f s),$$

where $\bar{\rho}(s)$ is given in (2.2). SDF at $f = 0$ is

$$\text{(A.2)} \qquad \mathrm{SDF}(0) = 1 + 2\sum_{s=1}^{2n} \bar{\rho}(s,n)$$

$$= 1 + 2\sum_{s=1}^{2n}\left[1 + \frac{X_1}{nX}s + \frac{X_2}{n^2X}s^2 + \frac{X_3}{n^3X}s^3 + \frac{X_5}{n^5X}s^5\right]$$

where

$$X = 32 + 80n^{-1} + 40n^{-2} - 20n^{-3} - 12n^{-4},$$

$$X_1 = -152 - 304n^{-1} - 150n^{-2} + 2n^{-3} + 6n^{-4},$$

$$X_2 = 180 + 270n^{-1} + 90n^{-2},$$

$$X_3 = -68 - 68n^{-1} - 9n^{-2},$$

$$X_5 = 3.$$

Since none of the X's contain an s term, in order to evaluate (A.2) we need

$$\sum_{s=1}^{2n} s^0 = 2n,$$

$$\sum s^1 = n(2n+1),$$

$$\sum s^2 = \frac{1}{3}n(2n+1)(4n+1),$$

$$\sum s^3 = n^2(2n+1)^2,$$

$$\sum s^5 = \frac{1}{3}n^2(2n+1)^2(8n^2+4n-1).$$

Substituting these, but again after some laborious algebra, we have

$$\mathrm{SDF}(0) = 0.$$

Finally, SDF at $f = 0.5$ is

$$\text{(A.3)} \qquad \mathrm{SDF}(0.5) = 1 + 2\sum_{s=1}^{2n} \bar{\rho}(s,n)\cos(\pi s) = 1 + 2\left[\sum_{\text{even}} \bar{\rho}(s,n) - \sum_{\text{odd}} \bar{\rho}(s,n)\right].$$

From the summation formula for various powers of s from $s = 1$ to $s = 2n$, given above, it is not difficult to derive, for even s:

$$\sum s^0 = n,$$

$$\sum s^1 = n(n+1),$$

$$\sum s^2 = \frac{2}{3}n(n+1)(2n+1),$$

$$\sum s^3 = 2n^2(n+1)^2,$$

$$\sum s^5 = \frac{8}{3}n^2(n+1)^2(2n^2+2n-1),$$

and for odd s:

$$\sum s^0 = n,$$

$$\sum s^1 = n^2,$$

$$\sum s^2 = \frac{1}{3} n(2n+1)(2n-1),$$

$$\sum s^3 = n^2(2n^2-1),$$

$$\sum s^5 = \frac{1}{3} n^2(16n^4 - 20n^2 + 7).$$

Substituting all these into (A.3), we have

$$\mathrm{SDF}(0.5) = \frac{20n(n+1)}{8n^3 + 12n^2 - 2n - 3},$$

which becomes approximately $5/2n$ as n gets large.

On the other hand, if we use CHO's formula for $\bar{\rho}(s, n)$ given by (3.13) and (3.14) in their paper, we have

$$\mathrm{SDF}(0) = \frac{n(10n^3 + 35n^2 + 46n + 21)}{3(4n^3 + 10n^2 + 3n - 3)}$$

and

$$\mathrm{SDF}(0.5) = \frac{15n(n+1)^2}{8n^2 + 20n^3 + 10n^2 - 5n - 3}.$$

When n is large, in CHO, we have

$$\mathrm{SDF}(0) \simeq \frac{5}{6} n \quad \text{and} \quad \mathrm{SDF}(0.5) \simeq \frac{15}{8n}.$$

REFERENCES

[1] Chan, K. H., J. C. Hayya, and J. K. Ord: "A Note on Trend Removal Methods: The Case of Polynomial Regression Versus Variate Differencing," *Econometrica*, 45(1977), 737–744.

[2] Hannan, E. J.: "The Estimation of the Spectral Density After Trend Removal," *Journal of the Royal Statistical Society (B)*, 20(1958), 323–333.

[3] Hatanaka, M., and E. P. Howrey: "Low Frequency Variation in Economic Time Series," *Kyklos*, 22(1969), 752–766.

[4] Granger, C. W. J.: "The Typical Spectral Shape of An Economic Variable," *Econometrica*, 34(1966), 150–161.

[5] Lovell, M. C.: "Seasonal Adjustment of Economic Times Series and Multiple Regression Analysis," *Journal of the American Statistical Association*, 58(1963), 993–1010.