A Method for Estimating Distributed Lags When
Observations are Randomly Missing

Melvin J. Hinich
and
Warren E. Weber

April 1981

Staff Report #65

Virginia Polytechnic Institute
and State University
and
Federal Reserve Bank of Minneapolis

This paper presents a frequency-domain technique for estimating distributed lag coefficients (the impulse-response function) when observations are randomly missed. The technique treats stationary processes with randomly missed observations as amplitude-modulated processes and estimates the transfer function accordingly. Estimates of the lag coefficients are obtained by taking the inverse transform of the estimated transfer function. Results with artificially created data show that the technique performs well even when the probability of an observation being missed is one-half and in some cases when the probability is as low as one-fifth. The approximate asymptotic variance of the estimator is also calculated in the paper.

Introduction:

Least-squares estimation of the lag coefficients of a distributed lag model is not a straight-forward regression problem when the sample has missing observations. Even though the normal equations can be computed using sums of the available cross products, the estimators in the solution of these equations are not necessarily unbiased. In any event, a special computer program is required to compute the normal equations of a distributed lag linear model for this case.

In this paper, we present a frequency domain approach to the estimation of lag coefficients when observations are randomly missed. The estimators are easy to compute if a discrete Fourier transform program is available, as is the case for most time series software libraries. Estimates of the lags computed by our approach have common large sample variance. This is a useful property for identifying model structure.

Our paper proceeds as follows. In the first section we review the frequency-domain approach to the estimation of distributed lag models. In the second, we discuss the estimation of spectra and distributed lags with missing observations. In the final section we present some results of using our suggested technique on artifically created data.

1. A Frequency Domain Approach:

Our discussion of the frequency domain approach to the estimation of distributed lags follows the commonly used non-parametric approach (Jenkins and Watts, 1968, Chapter 10). Let $\{x(t)\}$ and $\{\varepsilon(t)\}$ be two independent, mean zero, stationary time series, and let

$$E[x(t+\tau)x(t)] = \sigma_x(\tau) \qquad \text{for all } \tau$$

and

$$E[\varepsilon(t+\tau)\varepsilon(t)] = \begin{cases} \sigma_\varepsilon^2 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases},$$

so that $\{\varepsilon(t)\}$ is white noise. For simplicity, set the time unit equal to the sampling interval, so that t takes on integer values. The time series $\{y(t)\}$ is related to the other two series according to:

$$y(t) = \sum_{k=-\infty}^{\infty} h(k)x(t-k) + \varepsilon(t), \tag{1}$$

where $\{h(k)\}$ is an absolutely summable sequence whose transfer function is zero outside the band $-\pi < \omega < \pi$. We will also assume that $h(0) = 0$, so that changes in $x(t)$ cannot have an immediate impact on $y(t)$.

From (1) we obtain

$$S_{xy}(\omega) = H(\omega)S_x(\omega), \qquad -\pi < \omega < \pi, \tag{2}$$

where $S_x(\omega)$ and $S_{xy}(\omega)$ denote the own spectrum of $\{x(t)\}$ and the cross-spectrum between {x(t)} and {y(t)}, respectively, and

$$H(\omega) = \sum_{k=-\infty}^{\infty} h(k)\, \exp(-i\omega k) \qquad (3)$$

is the transfer function of {h(t)}.[1] Upon rearranging (2),

$$H(\omega) = S_{xy}(\omega)/S_x(\omega). \qquad (4)$$

Now, suppose that we observe the series {y(t)} and {x(t)} at the same times $t=0,1,\ldots,n-1$. From this set of observations we can obtain estimates of the spectrum of {x(t)} and the cross-spectrum between {x(t)} and {y(t)} as follows. Let $X(\omega_k)$ denote the discrete Fourier transform of {x(t)} for the observations at the angular frequency $\omega_k = 2\pi k/n$; i.e.,

$$X(\omega_k) = \sum_{t=0}^{n-1} x(t)\exp(-i\omega_k t). \qquad (5)$$

---

[1]Let $\gamma_a(k)$ be the absolutely summable covariance function of the mean zero, stationary time series {a(t)}. We use the convention that

$$S_a(\omega) = \sum_{k=-\infty}^{\infty} \gamma_a(k)e^{-i\omega k}, \qquad -\pi\leq\omega\leq\pi$$

defines the spectrum of {a(t)}. Under this convention

$$\gamma_a(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_a(\omega)e^{i\omega k}d\omega.$$

Thus,

$$\gamma_a(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_a(\omega)d\omega.$$

Thus, we can estimate $S_x(\omega_k)$ as

$$\hat{S}_x(\omega_k) = \frac{1}{s}\sum_{j=-d}^{d} |X(\omega_{k+j})|^2 , \tag{6}$$

where s=2d+1. Similarly, if we let $Y(\omega_k)$ denote the discrete Fourier transform of {y(t)} at the angular frequency $_k$, we can estimate $S_{xy}(\omega_k)$ as

$$\hat{S}_{xy}(\omega_k) = \frac{1}{s}\sum_{j=-d}^{d} X(\omega_{k+j})Y^*(\omega_{k+j}), \tag{7}$$

where an asterisk denotes complex conjugate.

Thus, we can form an estimate of $H(\omega)$ by substituting (6) and (7) into (4) to obtain

$$\hat{H}(\omega) = \hat{S}_{xy}(\omega)/\hat{S}_x(\omega). \tag{8}$$

From (10.3.14) in Jenkins and Watts we know that when n and s are large

$$E|\hat{H}(\omega_k) - H(\omega_k)|^2 \simeq \frac{1}{s} |H(\omega_k)|^2 \lfloor \gamma_{xy}^{-2}(\omega_k)-1 \rfloor, \tag{9}$$

where $\gamma_{xy}^2(\omega_k)$ is the squared coherency between {y(t)} at {x(t)} at frequency ω_k; i.e.,

$$\gamma_{xy}^2(\omega_k) = |S_{xy}(\omega_k)|^2/S_x(\omega_k)S_y(\omega_k).$$

An estimate of {h(k)} can be obtained by taking the discrete inverse transform of $\hat{H}(\omega)$; i.e., we can estimate {h(k)} as

$$\hat{h}(k) = \frac{1}{n}\sum_{j=0}^{n-1} \hat{H}(\omega_j)\exp(i\omega_k j). \tag{10}$$

Given the smoothing in the numerator and denominator of (8)

$$\hat{h}(k) \simeq \frac{1}{m}\sum_{j=0}^{m-1} \hat{H}(\omega_{js})\exp(i\omega_{js}k), \tag{10'}$$

where $m = n/s$. If $m \to \infty$ as $n \to \infty$, then the expected value and variance of the right hand side of (10') converge to the expected value and variance of (10). Since $\{\hat{H}(\omega_{js}): j=0,\ldots,m-1\}$ are asymptotically independent as $m \to \infty$, it follows from (9) that

$$\mathrm{Var}[\hat{h}(k)] \simeq \frac{1}{m^2} \sum_{j=0}^{m-1} E|\hat{H}(\omega_{js}) - H(\omega_{js})|^2 \tag{11}$$

$$= \frac{1}{m^2 s} \sum_{j=0}^{m-1} |H(\omega_{js})|^2 [\gamma_{xy}^{-2}(\omega_{js}) - 1]$$

$$\simeq \frac{1}{2\pi n} \int_0^{2\pi} |H(\omega)|^2 [\gamma_{xy}^{-2}(\omega) - 1] d\omega.$$

The estimators of $\{\hat{h}(k)\}$ have the same variance at all lags since this variance is independent of k. Equation (11) suggests that the variance of $\{\hat{h}(k)\}$ can be estimated as

$$\hat{\mathrm{V}}\mathrm{ar}[\hat{h}(k)] = \frac{1}{n} \sum_{j=0}^{n-1} |\hat{H}(\omega_j)|^2 [\hat{\gamma}_{xy}^{-2}(\omega_j) - 1]. \tag{12}$$

2. Missing Observations:

Our method of estimating $\{h(k)\}$ in the case of missing observations proceeds along lines similar as those above. However, it first involves finding a method of estimating $S_x(\omega)$ and $S_{xy}(\omega)$ for the case of randomly missing observations. Our approach is to follow Bloomfield (1970) who applies Parzen's (1963) approach to amplitude-modulated stationary processes to the estimation of single series spectra in the case of randomly missing observations.

Consider the case of the $\{x(t)\}$ series. The amplitude-modulated series $\{x'(t)\}$ is constructed by replacing the missing observations in the original series by zero, the mean of $\{x(t)\}$. In other words, we define the amplitude modulating series $\{z(t)\} = \{x'(t)/x(t)\}$ as

$$z(t) = \begin{matrix} 1 \text{ if } x(t) \text{ is observed} \\ 0 \text{ otherwise.} \end{matrix} \tag{13}$$

Thus,

$$x'(t) = z(t)x(t). \tag{14}$$

Following Bloomfield and Scheinok (1965), we assume that the process which causes observations to be missing is a stochastic process which is independent of $\{x(t)\}$. In addition, we assume that $\{z(t)\}$ has the properties that for all t,

$$P\lfloor z(t) = 1\rfloor = p_x$$

$$P\lfloor z(t) = 0\rfloor = 1 - p_x$$

and

$$\mathrm{Cov}\lfloor z(t), z(t+\tau)\rfloor = \sigma_z(\tau).$$

Thus $\{z(t)\}$ is weakly stationary, and

$$E[z(t)z(t+\tau)] = \sigma_z(\tau) + p_x^2 \qquad \text{for all } \tau.$$

By the independence of $\{x(t)\}$ and $\{z(t)\}$, it follows that $E\{x'(t)\} = E\{x(t)\}E\{z(t)\} = 0$. Consequently,

$$\begin{aligned} E\{x'(t)x'(t+\tau)\} &= E\{x(t)x(t+\tau)\}\, E\{z(t)z(t+\tau)\} \\ &= \sigma_x(\tau)\, [p_x^2 + \sigma_z(\tau)] \end{aligned} \tag{15}$$

It then follows that the spectrum of $\{x'(t)\}$ is

$$S_{x'}(\omega) = p_x^2 S_x(\omega) + (2\pi)^{-1} \int_{-\pi}^{\pi} S_z(\omega-\omega')\, S_x(\omega')d\omega'. \tag{16}$$

Similarly, we assume that there is an amplitude-modulating series $\zeta(t)$ such that

$$\zeta(t) = \begin{cases} 1 \text{ if } y(t) \text{ is observed} \\ 0 \text{ otherwise} \end{cases} \tag{17}$$

which gives rise to the amplitude-modulated series

$$y'(t) = \zeta(t)y(t). \tag{18}$$

We assume that $\{\zeta(t)\}$ is independent of both $\{x(t)\}$ and $\{y(t)\}$ and that the series has the properties that for all t,

$$P[\zeta(t) = 1] = p_y$$

$$P[\zeta(t) = 0] = 1 - p_y$$

$$\mathrm{Cov}[\zeta(t), z(t+\tau)] = \sigma_{z\zeta}(\tau).$$

Thus,

$$E[\zeta(t)z(t+\tau)] = \sigma_{z\zeta}(\tau) + p_x p_y.$$

Since $\{x(t)\}$ and $\{y(t)\}$ are independent of both $\{z(t)\}$ and $\{\zeta(t)\}$, it follows that

$$\begin{aligned} E\{x'(t)y'(t+\tau)\} &= E\{x(t)y(t+\tau)\}\ E\{z(t)\zeta(t+\tau)\} \\ &= \sigma_{xy}(\tau)[p_x p_y + \sigma_{z\zeta}(\tau)]. \end{aligned} \tag{19}$$

It then follows that the cross spectrum of $\{x'(t)\}$ and $\{y'(t)\}$ is

$$S_{x'y'}(\omega) = p_x p_y S_{xy}(\omega) + (2\pi)^{-1} \int_{-\pi}^{\pi} S_{z\zeta}(\omega-\omega')S_{xy}(\omega')d\omega'. \tag{20}$$

Thus, using estimators similar to (6) and (7) we could estimate $S_{x'}(\omega)$, $S_{x'y'}(\omega)$, $S_z(\omega)$, and $S_{z\zeta}(\omega)$ from the available data and use (16) and (20) to solve for $S_x(\omega)$ and $S_{xy}(\omega)$. These estimates could then be substituted into (8) and $\{h(k)\}$ estimated by taking the inverse transform of the result, as in the case when no observations were missing. Note that our analysis does not require either that $\{x(t)\}$ and $\{y(t)\}$ be observed at the same times or that the same random process generates the missing observations of the two series. In addition, there is no requirement that the random processes governing the observability of the underlying series be independent, so that the analysis does not preclude the case in which $\{x(t)\}$ and $\{y(t)\}$ must be observed at the same times.

In general, (16) and (20) can be solved for $S_x(\omega)$ and $S_{xy}(\omega)$ using the method given by Bloomfield. However, in the present paper we will only discuss the special case in which the series $\{z(t)\}$ and $\{\zeta(t)\}$ are white noise. In this case, (16) and (20) can be considerably simplified and the solutions for $S_x(\omega)$ and $S_{xy}(\omega)$ become much easier.

When the binomial distributions generating the missing observations are white noise,

$$\sigma_z(\tau) = \begin{cases} p_x(1-p_x) & \tau = 0 \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

and

$$\sigma_{z\zeta}(\tau) = \begin{cases} \sigma_{z\zeta} & \tau = 0 \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

Under these assumptions (16) and (20) become

$$S_{x'}(\omega) = p_x^2 S_x(\omega) + p_x(1-p_x)\sigma_x^2 \tag{23}$$

and

$$S_{x'y'}(\omega) = p_x p_y S_{xy}(\omega) + \sigma_{z\zeta}\sigma_{xy}, \tag{24}$$

respectively. Thus, we find that in the case in which $\{z(t)\}$ and $\{\zeta(t)\}$ are white noise, the spectra of the amplitude-modulated series are linear transformations of the original series. Substituting (23) and (24) into (4) yields

$$H(\omega) = (p_x/p_y)\ \{\lfloor S_{x'y'}(\omega) - \sigma_{z\zeta}\sigma_{xy}\rfloor / \lfloor S_{x'}(\omega) - p_x(1-p_x)\sigma_x^2\rfloor\}. \tag{25}$$

Thus, if we estimate $S_{x'}(\omega_k)$ as

$$\hat{S}_{x'}(\omega_k) = \frac{1}{s}\sum_{j=-d}^{d} |X'(\omega_{k+j})|^2 \tag{6'}$$

and $S_{x'y'}(\omega_k)$ as

$$\hat{S}_{x'y'}(\omega_k) = \frac{1}{s} \sum_{j=-d}^{d} X'(\omega_{k+j})Y'^*(\omega_{k+j}), \tag{7'}$$

we can estimate $H(\omega_k)$ as

$$\hat{H}(\omega_k) = (\hat{p}_x/\hat{p}_y)\{[\hat{S}_{x'y'}(\omega_k) - \hat{\sigma}_{z\zeta}\hat{\sigma}_{xy}]/[\hat{S}_{x'}(\omega_k)-\hat{p}_x(1-\hat{p}_x)\hat{\sigma}_x^2]\}. \tag{26}$$

Further, if $\sigma_{z\zeta} = 0$ or $\sigma_{xy} = 0$, it is shown in the Appendix that

$$E|\hat{H}(\omega_k) - H(\omega_k)|^2 \approx \frac{1}{s} |H(\omega_k)|^2\{\gamma_{x'y'}^{-2}(\omega_k) + [(\frac{1-p_x}{p_x})\frac{\sigma_x^2}{S_x(\omega_k)}]^2 -1\}. \tag{27}$$

This approximation is good when $p\sqrt{s} >> 1$, where $p = \min(p_x, p_y)$.

Once again an estimate of $\{h(k)\}$ can be obtained by taking the discrete inverse transform of $\hat{H}(\omega)$ as given by (26). Further, we could obtain the approximate variance of $\{\hat{h}(k)\}$ by repeating the steps used to obtain (11) using (27) in place of (9). This variance is

$$\mathrm{Var}[\hat{h}(k)] \approx \frac{1}{2\pi n} \int_0^{2\pi} |H(\omega)|^2\{\gamma_{x'y'}^{-2}(\omega) + [(\frac{1-p_x}{p_x})\frac{\sigma_x^2}{S_x(\omega_k)}]^2 - 1\}d\omega. \tag{28}$$

It can be estimated in the same way that (11) was estimated above.

Once again, we note that the asymptotic variance of $\hat{h}(k)$ is the same for all lag coefficients. Further, note that if $\sigma_x^2/S(\omega)$ and $\sigma_y^2/S_x(\omega)$ are close to unity for all ω and $p_x = p_y = p$, the variance in the case missing observations will be more than p^2 times that for the case in which no observations are missing.[2] Further, when p gets small, the standard errors of $\{\hat{h}(k)\}$ are of the order $(p\sqrt{n})^{-1}$.

---

[2]If $\sigma_{z\zeta} = 0$ or $\sigma_{xy} = 0$, it is easy to show that

$$\gamma^2_{x'y'}(\omega) = \gamma^2_{xy}(\omega)/\{\left[1 + \frac{(1-p_x)}{p_x}\,\frac{\sigma_x^2}{S_x(\omega)}\right]\left[1 + \frac{(1-p_y)}{p_y}\,\frac{\sigma_y^2}{S_y(\omega)}\right]\} \,. \qquad (29)$$

Thus, if $S_x(\omega) \approx \sigma_x^2$ and $S_y(\omega) \approx \sigma_y^2$, the terms in brackets in (29) are approximately equal to p_x^{-1} and p_y^{-1}, respectively. Thus, for this for this case $\gamma^2_{x'y'}(\omega) \approx (p_x p_y)\gamma^2_{xy}(\omega)$.

3. Results with Artifically Created Data:

In order to evaluate our frequency-domain approach to estimating distributed lags when observations are randomly missing, we used it on some artifically created data. Specifically, we generated $\{y(t)\}$ according to (1) using the weights

$$h(k) = \begin{cases} 0.25 & k = 1, 5 \\ 0.75 & k = 2, 4 \\ 1.0 & k = 3 \\ 0.0 & \text{otherwise.} \end{cases} \tag{30}$$

The errors $\{\varepsilon(t)\}$ were computed by a normal N(0,1) pseudo-random number generator. For all experiments the number of observations was 12000, which is approximately the number of days since 1947. The independent Bernoulli trials used to create $\{z(t)\}$ and $\{\zeta(t)\}$ were obtained using a uniform (0,1) pseudo-random number generator. In the case of $\{z(t)\}$, if the value of the t-th random number exceeded p_x, then $z(t)$ was set equal to zero. Otherwise, $z(t) = 1$. A similar procedure was used to generate $\{\zeta(t)\}$ with p_y used in place of p_x. In all experiments $\{z(t)\}$ and $\{\zeta(t)\}$ were independent.[3] All spectral estimates were obtained by smoothing periodograms. The weights used in the smoothing of $\hat{S}_{x'y'}(\omega_k)$ and $\hat{S}_{x'}(\omega_k)$ are discussed below. In the

---

[3]The specific normal pseudo-random number generator used was the FTGEN subroutine in the International Mathematics and Statistics Library (IMSL). The specific uniform (0,1) pseudo-random number generator used was the GGUBS subroutine of IMSL.

calculation of (26) actual values of p_x, p_y, and $\sigma_{z\zeta}$ were used instead of estimated ones.[4] The variance of $\{x(t)\}$ was estimated using only the non-zero values of $\{x'(t)\}$.

The first set of experiments were performed letting $\{x(t)\}$ be white noise normal random variables with unit variance. The $\{x(t)\}$ were generated with the same computer subroutine used to generate $\{\varepsilon(t)\}$. The results for each of up to ten runs are presented in Table 1. The values used for p_x and p_y are given in the column labeled "Percentage Observation." In all runs a 999 point moving average was used to smooth the periodogram to obtain the spectral estimates. The row labeled "Squared Error Fit" contains the values of

$$\sum_{k=0}^{11999} [h(k) - \hat{h}(k)]^2,$$

which is the sum of squared differences between the actual and estimated lag coefficients.

There are three major points concerning these results. The first is that the estimated lag coefficients track the actual lag coefficients quite well as long as p_x and p_y are 0.2 or above. This can be seen both in the low values of the squared error of fit and the high degree of correspondence between the means of the estimated lag coefficients and their actual values.

The second point is that the variances of the parameter estimates increase as the percentage of missing observations increasesas (28) and (29) indicates they should. This can be seen in the increasing standard errors of all of the estimates as the "Percentage Observation" decreases. Further, it

---

[4]Since the values which would have been obtained if p_x,p_y and $\sigma_{z\zeta}$ had been estimated were close to the true values, the results would not change much if estimated values of those parameters had been used instead.

is interesting to note as long as p_x and p_y are 0.2 or above, these standard errors are roughly constant cross coefficients as (27) indicates they should be.

The third point is that there is a downward bias in the estimates of h(1) to h(5), the non-zero lag coefficients. Our conjecture is that this downward bias is caused by the bias imparted to the spectral estimates due to including 999 terms in the moving average used to smooth the periodogram. To test this conjecture, we redid several of the runs with $S_{x'y'}(\omega_k)$ and $S_x(\omega_k)$ estimated using only 99 terms in the moving average. These results are contained in Table 2. As can be seen, this modification does decrease the downward bias for the cases $p_x = p_y = .9$ and $p_x = p_y = .5$. However, it is also apparent that for this case the estimated lag coefficients track the true lag distributions well only when no more than 50 percent of the observations are missing. Thus, we have the classic trade-off between bias and variance due to the fact that increasing the number of terms used to smooth the periodogram decreases the variance but increases the bias of the estimates of the spectrum.

We also evaluated our procedure when {x(t)} was colored. Specifically, using the same nNormal N(0,1) pseudo-random number generator as before we generated

$$x(t) = \rho x(t-1) + \nu(t), \qquad (19)$$

with $\rho = 0.5$ and $\sigma_\nu^2 = 1$. We also generated enough observations on x(t) prior to those used to estimate the lag coefficients, so that the choice of initial x(t) should not influence the results. Our results for this case are presented in Table 3. Once again we find that we are able to track the lag distribution well as long as no more than 50 percent of the observations are missing. Further, the results for this case continue to show that variance of the estimates increase as p_x and p_y decrease and that including more points in the moving average used to smooth the periodograms increases the bias, but decreases the variance of the estimated coefficients.

Acknowledgement: Professor Hinich's research was supported by the Office of Naval Research (Statistics and Probability Program) under contract.

Table 1 - Results with x(t) White Noise and 999 Point Moving Average

PARAMETER	RUN 1	2	3	4	5	6	7	8	9	10	MEAN	STANDARD ERROR	PERCENTAGE OBSERVATION
h(0) (0.0)*	-0.0058	-0.0140	-0.0173	0.0012	0.0039	-0.0139	-0.0160	0.0033	-0.0170	0.0041	-0.0072	0.0094	.9
	-0.0272	0.0207	-0.0180	-0.0072	-0.0293	-0.0157	-0.0142	-0.0061	-0.0508	0.0476	-0.0100	0.0273	.5
	0.0682	0.0592	-0.0362	0.0353	0.0309						0.0315	0.0410	.25
	0.1399	0.0089	-0.0877	0.1599	0.0422						0.0686	0.0987	.2
	0.7663	2.513	-1.225	0.1010	-0.2340	0.0040	-0.5135	-0.0636	-1.141	-0.0118	0.0194	1.056	.1
h(1) (0.25)	0.2304	0.2351	0.2203	0.2465	0.2496	0.2427	0.2519	0.2306	0.2419	0.2502	0.2399	0.0105	.9
	0.2640	0.2295	0.2912	0.2857	0.1872	0.2088	0.2513	0.2345	0.2763	0.2811	0.2510	0.0352	.5
	0.2746	0.2305	0.2462	0.3205	0.1890						0.2552	0.0492	.25
	0.3315	0.3103	0.2904	0.3209	0.2619						0.3030	0.0275	.2
	0.9100	3.179	-0.1667	0.2722	0.3383	0.0984	0.5212	0.1811	-0.1632	0.2933	0.5464	0.9771	.1
h(2) (0.75)	0.7177	0.7074	0.7292	0.7121	0.7114	0.6994	0.7053	0.7214	0.7255	0.7117	0.7141	0.0093	.9
	0.7110	0.7193	0.7218	0.7103	0.6844	0.7265	0.7391	0.7140	0.7214	0.7883	0.7236	0.0267	.5
	0.8571	0.6637	0.7134	0.7432	0.6182						0.7191	0.0907	.25
	0.9183	0.7799	0.6377	0.8201	0.5893						0.7491	0.1347	.2
	2.287	3.811	1.059	0.6999	0.5314	0.6807	1.582	0.6852	1.387	0.5518	1.328	1.038	.1
h(3) (1.0)	0.9023	0.8889	0.9173	0.8879	0.8973	0.9179	0.9022	0.9044	0.9278	0.8877	0.9034	0.0139	.9
	0.9164	0.9040	0.8938	0.9118	0.8432	0.8903	0.8683	0.8339	0.9569	0.9254	0.8944	0.0376	.5
	0.9520	0.9119	0.9578	0.7716	0.7857						0.8758	0.0906	.25
	0.9514	0.9699	0.9573	0.7439	0.7768						0.8799	0.1099	.2
	2.004	3.827	2.602	0.8589	0.6081	0.5409	1.419	0.7126	2.131	0.9799	1.568	1.068	.1
h(4) (0.75)	0.6339	0.6265	0.6477	0.6222	0.6350	0.6274	0.6227	0.6502	0.6386	0.6240	0.6328	0.0101	.9
	0.6666	0.6615	0.6631	0.6280	0.6073	0.6066	0.6151	0.6392	0.6884	0.6639	0.6440	0.0287	.5
	0.7022	0.6778	0.7690	0.6065	0.5324						0.6576	0.0910	.25
	0.7046	0.6586	0.7930	0.6116	0.5779						0.6691	0.0842	.2
	1.658	3.311	2.150	0.5719	0.5138	0.2642	1.539	0.4677	1.217	0.5001	1.219	0.9688	.1
h(5) (0.25)	0.1950	0.1842	0.1878	0.1925	0.1898	0.1975	0.2103	0.1861	0.1995	0.2154	0.1958	0.0103	.9
	0.1955	0.1797	0.1888	0.1622	0.1866	0.2077	0.2026	0.1965	0.2137	0.2211	0.1954	0.0172	.5
	0.2534	0.2004	0.2481	0.1686	0.2173						0.2176	0.0350	.25
	0.3316	0.2150	0.3247	0.1831	0.2184						0.2546	0.0686	.2
	1.466	2.730	0.605	0.2413	0.2355	0.2792	0.5428	0.0037	-0.8051	0.3049	0.5603	0.9461	.1
h(6) (0.0)	0.0077	0.0106	0.0343	0.0217	0.0050	0.0247	0.0163	0.0224	0.0094	0.0060	0.0158	0.0100	.9
	0.0190	-0.0059	0.0033	-0.0022	-0.0053	0.0110	0.0185	0.0580	0.0106	0.0400	0.0147	0.0206	.5
	0.0516	0.0059	0.0243	-0.0239	-0.0271						0.0062	0.0332	.25
	0.0738	0.0388	0.0729	-0.0212	0.0289						0.0386	0.0390	.2
	0.7462	2.261	-1.022	0.0579	-0.0721	0.1707	-0.0206	-0.2139	-1.362	-0.0777	0.0468	0.9796	.1
Squared Error of Fit	0.0540	0.0619	0.0695	0.0692	0.0579	0.0606	0.0681	0.0613	0.0598	0.0679	0.0063	0.0054	.9
	0.0468	0.0578	0.0851	0.0724	0.0871	0.0805	0.0832	0.0966	0.0621	0.0647	0.0736	0.0155	.5
	0.1427	0.0886	0.1313	0.1787	0.1752						0.1433	0.0368	.25
	0.2418	0.1194	0.1941	0.2730	0.2059						0.2068	0.0579	.2
	0.3810^4	0.3212^5	0.1488^4	0.2871	0.4158	0.6835	3.426	0.4660	0.3564^4	0.2597	0.4099^4	0.9962^4	.1

*true value of parameter in parenthesis

Table 2 - Results with x(t) White Noise and 99 Point Moving Average

PARAMETER	RUN 1	2	3	4	5	6	7	8	9	10	MEAN	STANDARD ERROR	PERCENTAGE OBSERVATION
h(0) (0.0)*	0.0005	-0.0097	-0.0014	0.0112	0.0136	-0.0033	-0.0122	0.0063	-0.0095	0.0112	0.0008	0.0095	.9
	-0.0124	0.0244	-0.0178	0.0082	-0.0063	-0.0056	-0.0046	-0.0442	0.0468	-0.0196	-0.0046	0.0246	.5
	0.4048	0.0212	-0.0951	0.0131	0.4143						0.1517	0.2399	.25
	-2.191	0.0737	-0.5083	0.6516							-0.4935	1.227	.1
h(1) (0.25)	0.2391	0.2389	0.2301	0.2531	0.2589	0.2535	0.2574	0.2389	0.2497	0.2559	0.2476	0.0100	.9
	0.2886	0.2474	0.3040	0.1940	0.2244	0.2707	0.2455	0.2906	0.2910	0.2701	0.2626	0.0346	.5
	0.5766	0.1805	0.3007	0.2868	0.6266						0.3942	0.1957	.25
	-1.798	-0.8449	-1.658	0.9261							-0.8435	1.252	.1
h(2) (0.75)	0.7515	0.7428	0.7653	0.7466	0.7459	0.7359	0.7398	0.7581	0.7586	0.7450	0.7490	0.0093	.9
	0.7598	0.7742	0.7673	0.7245	0.7787	0.7909	0.7701	0.7720	0.8277	0.7486	0.7714	0.0268	.5
	1.300	0.7462	0.8841	0.9403	1.138						1.001	0.2182	.25
	-2.832	0.3427	-0.1110	1.044							-0.3891	1.697	.1
h(3) (1.0)	0.9983	0.9856	1.013	0.9830	0.9938	1.017	0.9998	1.001	1.025	0.9814	0.9998	0.0146	.9
	1.028	1.024	1.008	0.9406	0.9991	0.9753	0.9336	1.075	1.031	1.021	1.004	0.0433	.5
	1.439	1.182	1.167	0.9303	1.369						1.217	0.1990	.25
	-1.145	-1.128	1.705	1.081							0.1283	1.482	.1
h(4) (0.75)	0.7587	0.7500	0.7723	0.7452	0.7573	0.7481	0.7438	0.7790	0.7616	0.7473	0.7563	0.0119	.9
	0.8070	0.8103	0.8017	0.7332	0.7331	0.7347	0.7842	0.8405	0.8008	0.7596	0.7805	0.0381	.5
	1.168	0.8787	1.107	0.8438	1.090						1.018	0.1461	.25
	-1.116	0.4266	1.958	0.8674							0.5340	1.274	.1
h(5) (0.25)	0.2523	0.2376	0.2388	0.2494	0.2380	0.2493	0.2697	0.2383	0.2546	0.2775	0.2506	0.0139	.9
	0.2645	0.2437	0.2402	0.2504	0.2614	0.2582	0.2617	0.2922	0.2841	0.2035	0.2560	0.0246	.5
	0.5787	0.3075	0.3270	0.2103	0.7263						0.4300	0.2144	.25
	1.694	-0.0817	1.056	0.6281							0.8241	0.7460	.1
h(6) (0.0)	-0.0077	0.0037	0.0259	0.0161	-0.0205	0.0144	0.0080	0.0182	-0.0087	-0.0094	0.0040	0.0150	.9
	-0.0509	-0.0231	-0.0210	-0.0299	-0.0045	0.0075	0.0631	0.0040	-0.0418	-0.0257	-0.0122	0.0324	.5
	0.2751	-0.0920	-0.0614	-0.1025	0.3153						0.0669	0.2094	.25
	0.9052	-1.005	-0.2584	0.6541							0.0740	0.8760	.1
Standard Error	0.0184	0.0174	0.0207	0.0181	0.0192	0.0162	0.0194	0.0172	0.0156	0.0203	0.0183	0.0017	.9
	0.1763	0.1547	0.1547	0.1306	0.1337	0.1515	0.1585	0.1652	0.1255	0.1354	0.1486	0.0166	.5
	49.91	121.5	4.253	3.418	8762.						1788.	3899.	.25
	0.3016^5	0.6235^5	0.2195^5	0.2162^4							0.2916^5	0.2506^5	.1

*true value of paramenter in parenthesis

Table 3 - Results with x(t) as AR(1), ρ = 0.5

PARAMETER	RUN 1	2	3	4	5	6	7	8	9	10	MEAN	STANDARD ERROR	PERCENTAGE OBSERVATION
h(0) (0.0)*	0.0014	-0.0172	-0.0035	0.0006	0.0103	0.0072	-0.0182	0.0056	-0.0013	0.0248	0.0010	0.0126	.9(99)**
	0.0192	0.0527	-0.1124	-0.1040	-0.0013	0.0681	-0.0344	0.0112	-0.0101	0.0932	-0.0018	0.0677	.5(99)
	-0.0688	0.0626	-0.0708	-0.0545	-0.0030	0.0885	0.0263	0.0556	0.0146	0.1063	0.0163	0.0642	.5(999)
	-0.0858	0.1115	-0.1064	-0.3370	-0.0352	0.1829	0.1199	0.1225	0.0128	0.0581	0.0043	0.1535	.25(999)
h(1) (0.25)	0.2446	0.2502	0.2344	0.2539	0.2583	0.2570	0.2782	0.2468	0.2568	0.2433	0.2524	0.0118	.9(99)
	0.2144	0.2474	0.3855	0.3655	0.0956	0.2101	0.3055	0.2191	0.2617	0.1465	0.2451	0.0901	.5(99)
	0.3896	0.2810	0.4043	0.4190	0.1818	0.2332	0.2806	0.2590	0.3102	0.2123	0.2971	0.0828	.5(999)
	0.2252	0.3652	0.4081	0.8137	0.3734	0.3226	0.1547	0.3519	0.3549	0.3930	0.3763	0.1728	.25(999)
h(2) (0.75)	0.7503	0.7521	0.7733	0.7399	0.7396	0.7368	0.7276	0.7493	0.7451	0.7511	0.7465	0.0122	.9(99)
	0.8352	0.7004	0.7717	0.7031	0.8263	0.7450	0.7229	0.8676	0.7529	0.8945	0.7820	0.0695	.5(99)
	0.6933	0.7136	0.7711	0.6918	0.7788	0.7414	0.7987	0.8343	0.7451	0.8752	0.7643	0.0600	.5(999)
	0.9834	0.4005	0.6569	0.3362	0.5723	0.7388	0.8715	0.5478	0.6684	0.5352	0.6311	0.1987	.25(999)
h(3) (1.00)	1.006	0.9837	1.010	0.9946	1.006	1.017	1.008	0.9952	1.035	0.9776	1.003	0.0165	.9(99)
	0.9304	1.111	0.9214	0.9841	0.8568	0.9573	1.083	0.8011	1.057	0.9755	0.9778	0.0898	.5(99)
	0.9678	1.005	0.8646	0.9172	0.8256	0.8910	0.9546	0.7651	0.9805	0.9213	0.9093	0.0746	.5(999)
	0.9498	1.154	1.021	1.064	0.9147	0.7430	0.9708	0.8135	0.8604	1.121	0.9612	0.1328	.25(999)
h(4) (0.75)	0.7572	0.7549	0.7773	0.7442	0.7689	0.7427	0.7492	0.7851	0.7514	0.7519	0.7583	0.0141	.9(99)
	0.8890	0.7836	0.8761	0.8492	0.7733	0.7545	0.6690	0.8851	0.7912	0.7849	0.8056	0.0696	.5(99)
	0.6131	0.6241	0.6771	0.6621	0.6091	0.5794	0.5278	0.7003	0.6419	0.6030	0.6229	0.0488	.5(999)
	0.5576	0.5614	0.7367	0.2455	0.5003	0.6615	0.6317	0.6315	0.8744	0.5988	0.5999	0.1628	.25(999)
h(5) (0.25)	0.2530	0.2334	0.2294	0.2508	0.2322	0.2487	0.2690	0.2184	0.2574	0.2919	0.2484	0.0215	.9(99)
	0.0799	0.2374	0.1606	0.1332	0.3493	0.2212	0.2270	0.2057	0.3183	0.2444	0.2176	0.0805	.5(99)
	0.1182	0.1159	0.1011	0.0818	0.1856	0.1545	0.1673	0.1174	0.1782	0.1462	0.1366	0.0348	.5(999)
	0.1442	0.3214	0.2242	0.4743	0.3548	0.2146	-0.2000	0.2391	-0.0114	0.1597	0.1917	0.1904	.25(999)
h(6) (0.0)	-0.0175	0.0011	0.0201	0.0062	-0.0264	0.0214	0.0179	0.0264	-0.0012	-0.0205	0.0028	0.0190	.9(99)
	0.0959	-0.0211	0.0481	0.0236	-0.1922	0.0338	0.0724	0.1168	-0.0214	0.0640	0.0220	0.0877	.5(99)
	-0.0577	-0.0175	-0.0016	-0.0152	-0.1038	-0.0038	0.0004	0.0694	-0.0234	0.0166	-0.0137	0.0453	.5(999)
	-0.1181	-0.0686	-0.0791	-0.4091	-0.2320	-0.0296	0.1415	-0.1117	0.0466	0.1018	-0.0760	0.1613	.25(999)
Squared Error of Fit	0.0245	0.0255	0.0294	0.0232	0.0279	0.0222	0.0272	0.0300	0.0254	0.0312	0.0266	0.0030	.9(99)
	39.28	0.5503	2.384	43.56	2.334	0.6041	5.048	0.4812	0.5078	0.5400	9.529	16.90	.5(99)
	0.1123	0.0899	0.1536	0.1352	0.1058	0.1175	0.1188	0.1562	0.0872	0.1237	0.1200	0.0234	.5(999)
	0.2724	0.3542	0.3238	8.0194	0.3155	0.1902	0.4786	0.2341	0.2957	0.2678	1.077	2.441	.24(999)

*true value of parameter in parenthesis
**number of points used to smooth periodogram in parenthesis

REFERENCES

Bloomfield, P., "Spectral Analysis with Randomly Missing Observations," Journal of the Royal Statistical Society, Series B, 32, 1970, pp. 369-80.

Brillinger, D. Time Series, Data Analysis and Theory, New York: Holt, Rinehart and Winston, 1975.

Jenkins, G.M. and D.G. Watts, Spectral Analysis and Its Applications, Holden-Day, San Francisco, 1968.

Parzen, E., "On Spectral Analysis with Missing Observations and Amplitude Modulation," Sankhya, Series A, 25, 1963, pp. 383-92.

Scheinok, P.A. "Spectral Analysis with Randomly Missing Observations: The Binomial Case," Annals of Mathematical Statistics, 36, 1965, pp. 971-77.

APPENDIX

The Approximation (27) for the Mean Squared Error of $\hat{H}(\omega)$

Let $\varepsilon_{x'y'}(\omega) = \hat{S}_{x'y'}(\omega) - S_{x'y'}(\omega)$ and $\varepsilon_{x'}(\omega) = \hat{S}_{x'}(\omega) - S_{x'}(\omega)$. Assume that the cumulants of $\{x(t)\}$ and $\{y(t)\}$ satisfy condition (4.3.10) in Brillinger (1975).[5] From Theorems 4.3.2 and 4.4.1,

$$E\varepsilon_{x'y'}(\omega) \approx 0, \quad E\varepsilon_{x'}(\omega) \approx 0, \tag{A1}$$

$$E|\varepsilon_{x'y'}(\omega)|^2 \approx \frac{1}{s} S_{x'}(\omega)S_{y'}(\omega), \tag{A2}$$

$$E\varepsilon_{x'}{}^2(\omega) \approx \frac{1}{s} S_{x'}{}^2(\omega), \tag{A3}$$

and

$$E\overset{*}{\varepsilon}_{x'y'}(\omega)\varepsilon_{x'}(\omega) \approx \frac{1}{s} \overset{*}{S}_{x'y'}(\omega)S_{x'}(\omega) \tag{A4}$$

for large n (the approximations are of order n^{-1}). Since the variances of $\varepsilon_{x'y'}$ and $\varepsilon_{x'}$ are of order s^{-1}, whereas the variances of $\hat{p}_x$, $\hat{p}_y$, and $\hat{\sigma}_x{}^2$ are of order n^{-1}, we can substitute p_x, p_y, and $\sigma_x{}^2$ for their estimates in (26) without affecting the order of the approximation when $n >> s$.

Since $\sigma_{\zeta z} = 0$ or $\sigma_{xy} = 0$,

$$\hat{H}(\omega) = (p_x/p_y)\lfloor S_{x'y'}(\omega) + \varepsilon_{x'y'}(\omega)\rfloor / \lfloor S_{x'}(\omega) + \varepsilon_{x'}(\omega) - c\rfloor \tag{A5}$$

---

[5]This condition holds if $\{x(t)\}$ and $\{\varepsilon(t)\}$ are Guassian ARMA processes.

where $c = p_x(1 - p_x)\sigma_x^2$. The first order term in a Taylor series approximation of (A5) is

$$\hat{H}(\omega) - H(\omega) \simeq p_x p_y^{-1}(S_{x'}(\omega) - c)^{-1} \lfloor \varepsilon_{x'y'}(\omega) - (S_{x'}(\omega) - c)^{-1} S_{x'y'}(\omega)\varepsilon_{x'}(\omega) \rfloor \tag{A6}$$

It follows from (23), (24), (A2), and (A3) that this approximation is good if $1/p\sqrt{s}$ is small, where $p = \min(p_x, p_y)$. Applying (A2) and (A3) to (A6),

$$E|\hat{H}(\omega) - H(\omega)|^2 \simeq s^{-1}\lfloor p_x p_y^{-1}(S_{x'}(\omega) - c)^{-2}\rfloor$$

$$\lfloor S_{x'}(\omega)S_{y'}(\omega) - 2|S_{x'y'}(\omega)|^2(S_{x'}(\omega) - c)^{-1}S_{x'}(\omega)$$

$$+ |S_{x'y'}(\omega)|^2(S_{x'}(\omega) - c)^{-2}S^2_{x'}(\omega)\rfloor$$

$$\simeq s^{-1}\lfloor p_x p_y^{-1}(S_{x'}(\omega) - c)^{-2}|S_{x'y'}(\omega)|\rfloor^2 \lfloor \gamma^{-2}_{x'y'}(\omega)$$

$$-2(S_{x'}(\omega) - c)^{-1}S_{x'}(\omega) + (S_{x'} - c)^{-2}\ S^2_{x'}(\omega)\rfloor$$

$$= s^{-1}|H(\omega)|^2\{\gamma^{-2}_{x'y'}(\omega) - 1 + \lfloor 1 - (S_{x'}(\omega) - c)^{-1}S_{x'}(\omega)\rfloor^2$$

$$= s^{-1}|H(\omega)|^2\{\gamma^{-2}_{x'y'}(\omega) - 1 + \lfloor c(S_x(\omega) - c)^{-1}\rfloor^2\}$$

$$= s^{-1}|H(\omega)|^2\{\gamma^{-2}_{x'y'}(\omega) - 1 + \lfloor\left(\frac{1-p_x}{p_x}\right)\frac{\sigma_x^2}{S_x(\omega)}\rfloor^2\}.$$