The Annals of Statistics
1979, Vol. 7, No. 6, 1303–1320

SIGNAL EXTRACTION ERROR IN NONSTATIONARY TIME SERIES

BY DAVID A. PIERCE

Federal Reserve Board

It is supposed that an observable time series $\{x_t\}$ is representable as the sum of a "signal" s_t and a "noise" n_t, and that it is desired to extract the signal s_t, i.e., to obtain an estimate $\hat{s}_t$ of s_t. Corresponding to any such estimate is a signal extraction error, $\delta_t = s_t - \hat{s}_t$, which for nondeterministic stochastic processes possesses a nonzero mean square. For the class of homogeneously nonstationary processes, characterizations of the extraction error process are given, and it is shown that the mean square of the error does not exist unless the nonstationary autoregressive operators in the s- and n-processes have distinct roots. The MSE, autocorrelations and spectrum of the error, when it is stationary, are illustrated for some special cases, including two stochastic-model approximations to the Census X-11 seasonal adjustment procedure.

1. Introduction. The resolution of an observed time series into unobserved components has been a topic of interest in several disciplines for at least a century (see Nerlove (1967) for a review). Applications include seasonal adjustment of economic series, estimation of variables subject to measurement error, the formation of expectations, and signal extraction problems in engineering and other areas. An appropriately specified and estimated component series is generally more easily interpreted, being relatively freer from other, perhaps extraneous, influences. Alternatively, as in the case of seasonal adjustment, a given component series may itself be regarded as extraneous, its successful extraction facilitating a more meaningful study of the remainder of the series. Thus it is supposed that the observable series $\{x_t, -\infty < t < \infty\}$ is representable in the form

$$x_t = s_t + n_t \tag{1.1}$$

where s_t is the "signal" and n_t the "noise" (equivalent mnemonics, if the context is appropriate, are that s_t is the "seasonal" component of x_t and n_t the "nonseasonal" component). The two components are assumed to be mutually independent.

The estimation of the unobserved components s_t and n_t, given a realization of the observable series $\{x_t, -\infty < t < \infty\}$ (or a segment of such a realization), is the goal of signal extraction. (Generally one will want to utilize for this purpose whatever relevant information that exists, including relationships of s_t and n_t with other variables (e.g., Box, Hillmer and Tiao (1976), Granger (1976), Porter (1975), Wallis (1976)) as well as the obvious relationships between s_t or n_t and the observable series $\{x_t\}$. However, since for many situations such information may

Received January 1978; revised July 1978.

AMS 1970 *subject classifications.* Primary 62M10; secondary 0G35.

Key words and phrases. Time series, signal extraction, seasonal adjustment, unobserved component estimation.

be unknown, unreliable or disputed, or may be of only marginal value (e.g., Pierce (1977)), "univariate" approaches such as analyzed here will doubtless remain important.) For a large class of situations, knowledge of the model or generating mechanism of the series s_t and n_t enables a perfect separation of them to be attained; these component series can be estimated without error. This class includes regression functions and other deterministic mechanisms. For others, however, including most situations where the components are stochastic, complete knowledge of the model does not make possible an exact determination of the component series even given an infinite realization of the observable series. Any estimate $\hat{s}_t$ of s_t is necessarily in error by an amount

$$\delta_t = s_t - \hat{s}_t = \hat{n}_t - n_t \tag{1.2}$$

with a nonzero mean square.

For stationary time series the estimate $\hat{s}_t$ which minimizes the mean square of the extraction error δ_t, and the resulting variance and spectrum of δ_t, were given by Kolmogorov (1941) and Wiener (1949) and are also described in Whittle (1963) and Grether and Nerlove (1970). This estimate, which is the conditional expectation of s_t given the realization $\{x_t\}$, was extended to nonstationary time series by Cleveland and Tiao (1976). However, the properties of the estimation error (1.2) for nonstationary series, including even the question of whether its mean square is finite, have evidently not been investigated.

This paper examines the stochastic structure of the error (1.2) induced by signal extraction procedures for the class of homogeneously nonstationary time series, i.e., those whose suitable differences and/or sums are linear, stationary, nondeterministic processes. In the following section the class of estimation procedures, which includes all the "optimal" ones later considered, is described, and a general expression (valid irrespective of whether the series are stationary) for the extraction error process is given. The general nonstationary model for (1.1) is then set forth.

Section 3 examines the estimation error for optimal signal extraction where both the future and the past of the observable series are available, and Section 4 considers the cases where only current, recent, or past data are given. In all these situations it is found that the mean square of the error does not exist unless the nonstationary autoregressive operators in the s-process and in the n-process have distinct roots. The MSE, auto-correlations and spectrum of the extraction error, when they do exist, are illustrated for some special cases.

Section 5 focuses on seasonal adjustment, particularly seasonal adjustment based on two models for the Census X-11 procedure (Shiskin, Young and Musgrave, 1967), given by Cleveland (1972) and by Cleveland and Tiao (1976). The seasonal adjustment error spectrum and mean square error are obtained for the first of these, where s_t and n_t have distinct unit-modulus autoregressive roots. The error process is found to be mildly seasonal, as expected from the theory. Section 6 briefly considers some generalizations of the procedures and model employed.

2. Error process and component series models. All of the extraction/adjustment procedures considered in this paper consist of estimating the signal/seasonal s_t as a linear combination of $\{x_t\}$,

$$\hat{s}_t = \Sigma \nu_{sj} x_{t-j} = \nu_s(B) x_t, \tag{2.1}$$

where

$$\nu_s(B) = \Sigma \nu_{sj} B^j \tag{2.2}$$

is a polynomial in the lag operator B such that $\nu_s(z)$ is absolutely convergent in a neighborhood of $|z| = 1$. The limits of summation are deliberately left unspecified as they vary within the paper. Then

$$\hat{n}_t = x_t - \hat{s}_t = [1 - \nu_s(B)] x_t = \nu_n(B) x_t \tag{2.3}$$

is the corresponding estimate of n_t. This class of procedures includes both the "optimal" ones discussed in Sections 3 and 4 and others with a rich tradition such as the Census X-11 program.

It is actually the linear-filter approximation [16], [21] to the "additive" Census X-11 procedure that is of the form (2.1), as the procedure itself [14] includes additional features such as adjustments for outliers. Moreover there is also a "multiplicative" version of the X-11 procedure, which, however, is closely approximated by applying the additive version to the series' logarithms and then exponentiating.

The error process $\delta_t = s_t - \hat{s}_t$ can be usefully expressed in terms of the extraction filters $\nu_s(B)$ and $\nu_n(B)$ and the two component series. Since

$$\hat{s}_t = \nu_s(B)(s_t + n_t),$$

we have, from (1.2) and (2.3), the following

PROPOSITION. *For any unobserved-component time series $x_t = s_t + n_t$ and any estimate $\hat{s}_t$ of the form* (2.1) *for s_t, the estimation error $\delta_t = s_t - \hat{s}_t$ is given by*

$$\delta_t = \nu_n(B) s_t - \nu_s(B) n_t. \tag{2.4}$$

The two components of δ_t are mutually independent if and only if the components of x_t are mutually independent.

The nature of the error δ_t thus depends on both the chosen estimation or extraction procedure $\nu_s(B)$ and on the models for s_t and n_t. Concerning the latter, it will be assumed that $\{s_t\}$ and $\{n_t\}$ are mutually independent and each representable in the form

$$\Delta_s(B) s_t = \psi_s(B) \varepsilon_t \tag{2.5}$$

$$\Delta_n(B) n_t = \psi_n(B) \eta_t \tag{2.6}$$

where ε_t and η_t are orthogonal white noise sequences with finite variances σ_ε^2 and σ_η^2; the one-sided polynomials

$$\psi_s(z) = \Sigma_0^\infty \psi_{sj} z^j, \qquad \psi_n(z) = \Sigma_0^\infty \psi_{nj} z^j$$

are absolutely convergent and nonzero for $|z| \leqslant 1$; and $\Delta_s(B)$ and $\Delta_n(B)$ are "differencing operators" such that the zeroes of $\Delta_s(z)$ and $\Delta_n(z)$ are on the unit circle. Examples of such operators are the ordinary and "seasonal" differencing operators, $1 - B$ and $1 - B^{12}$ respectively, and the summation operator $S_{12}(B)$ in (5.2). It is also assumed that suitable initial conditions (see, e.g., Box and Jenkins (1970), pages 114–119) are given for s_t and n_t.

The models (2.5) and (2.6) for s_t and n_t are known to imply a model for the observable series x_t of the same form,

$$\Delta(B)x_t = \psi(B)a_t, \tag{2.7}$$

so that $\Delta(B)x_t$ is a linear, stationary, nondeterministic time series, as are $\Delta_s(B)s_t$ and $\Delta_n(B)n_t$. If all differencing and summing operators are identically unity the series x_t, s_t and n_t are stationary; if $\Delta(z) \not\equiv 1$ then x_t, and at least one of s_t and n_t, are nonstationary.

In general, $\Delta(z)$ is the least common multiple of $\Delta_s(z)$ and $\Delta_n(z)$. Let $\Delta_c(z)$ denote the factors common to $\Delta_s(z)$ and $\Delta_n(z)$, and let $\Delta_s^*(z) = \Delta_s(z)/\Delta_c(z)$, with $\Delta_n^*(z)$ being similarly defined. Then

$$\Delta(z) = \Delta_c(z)\Delta_s^*(z)\Delta_n^*(z), \tag{2.8}$$

and the *stationary* observable series is, in terms of its components,

$$\Delta(B)x_t = \Delta_n^*(B)\psi_s(B)\varepsilon_t + \Delta_s^*(B)\psi_n(B)\eta_t \tag{2.9}$$

or

$$x_{1t} = s_{1t} + n_{1t}. \tag{2.10}$$

The following points are of interest:

(i) s_{1t} and n_{1t} may be noninvertible, but

(ii) since the zeroes of Δ_n^* and Δ_s^* are distinct, their sum x_{1t} is invertible if $\psi_s(z)$ and $\psi_n(z)$ are invertible.

(iii) If $\Delta_c(z) \equiv 1$ then s_t and n_t have no common root of unit modulus. Whether this is so is important to the behavior of δ_t in Sections 3 and 4.

3. Estimation error of two-sided procedures. This section examines the error δ_t in signal extraction/seasonal adjustment on the basis of a complete realization $\{x_t, -\infty < t < \infty\}$. In practice this is relevant for the estimation of s_t in "historical" data where a sufficiently large number of observations on either side of the given s_t are available to provide a good approximation to the doubly infinite situation. The stationary case is first considered, followed by the nonstationary model of Section 2.

3.1 Observable series stationary. Suppose first that all differencing operators in (2.5)–(2.7) are unity, so that x_t has the representation

$$x_t = \psi_s(B)\varepsilon_t + \psi_n(B)\eta_t = \psi(B)a_t. \tag{3.1}$$

For this model it is shown in Whittle (1963, page 57) that the estimate $\hat{s}_t$ which minimizes $E(\delta_t^2)$ is given by (2.1) with

$$\nu_s(z) = \frac{\sigma_\varepsilon^2|\psi_s(z)|^2}{\sigma_a^2|\psi(z)|^2} = \frac{f_s(z)}{f_{s+n}(z)} \tag{3.2}$$

where, e.g.,

$$f_s(z) = \sigma_\varepsilon^2|\psi_s(z)|^2 \tag{3.3}$$

and where the convention

$$|h(z)|^2 = h(z)h(z^{-1})$$

is employed. Thus the filter is symmetric, $\nu_j = \nu_{-j}$, as expected from the reversibility of the x-process. The numerator and denominator of (3.2) are the autocovariance generating functions (acgf's), or spectra at $z = e^{i\omega}$, of the component and over-all processes $\{s_t\}$ and $\{x_t\}$.

Whether $\nu_s(z)$ is given by (3.2) or not, the stationarity of s_t and n_t implies that the estimation error $\delta_t = s_t - \hat{s}_t$ is also stationary. From (2.4), δ_t follows the process

$$\delta_t = \nu_n(B)\psi_s(B)\varepsilon_t - \nu_s(B)\psi_n(B)\eta_t \tag{3.4}$$

with acgf/spectrum

$$f_\delta(z) = \sigma_\varepsilon^2|\nu_n(z)\psi_s(z)|^2 + \sigma_\eta^2|\nu_s(z)\psi_n(z)|^2. \tag{3.5}$$

Specializing to the optimal case (3.2), and letting $f_s(z) = \sigma_\varepsilon^2|\psi_s(z)|^2$, etc.,

$$f_\delta(z) = \frac{f_n^2(z)}{f_{s+n}^2(z)} \cdot f_s(z) + \frac{f_s^2(z)}{f_{s+n}^2(z)} \cdot f_n(z) = \frac{f_s(z)f_n(z)}{f_{s+n}(z)} \tag{3.6}$$

$$= \frac{\sigma_\varepsilon^2|\psi_s(z)|^2 \cdot \sigma_\eta^2|\psi_n(z)|^2}{\sigma_a^2|\psi(z)|^2}. \tag{3.7}$$

The result (3.6) is derived in Whittle (1963, page 58) using a different method; the present method will be seen to extend directly to the case of nonstationary series. The error spectrum is simply the ratio of the product and the sum of the component spectra. The coefficient of z^k in (3.6) is the lag-k autocovariance of δ_t; the mean square error of the signal extraction is thus the coefficient of z^0, or equivalently the integral of $f_\delta(e^{i\omega})$ over ω.

By factoring $f_\delta(z)$, the error δ_t can be expressed as $\psi_\delta(B)\xi_t$ where $\{\xi_t\}$ is a white noise sequence and ψ_δ is one-sided (see Theorem 2).

3.2 Nonstationary models. Consider now the model (2.5)–(2.7) where x_t is stationary only after application of the operator $\Delta(B) \not\equiv 1$. Cleveland and Tiao (1976) have shown that the conditional expectation of s_t given $\{x_s, -\infty < s < \infty\}$ is again of the form (2.1), where now, referring to (2.9)–(2.10),

$$\nu_s(z) = \frac{\sigma_\varepsilon^2|\Delta_n^*(z)\psi_s(z)|^2}{\sigma_a^2|\psi(z)|^2} \doteq \frac{f_{s1}(z)}{f_{n1+s1}(z)}. \tag{3.8}$$

This filter applied to x_t gives, when the MSE exists, the minimum MSE estimate of s_t; moreover,

$$\hat{s}_{1t} = \Delta(B)\hat{s}_t = \nu_s(B)x_{1t} \tag{3.9}$$

is the minimum MSE estimate of s_{1t}.

The estimate $\hat{s}_t$ of s_t is clearly nonstationary whenever s_t is nonstationary. However, the difference $s_t - \hat{s}_t = \delta_t$ may or may not be nonstationary in this case. From (3.8) and (2.9) it follows that

$$\begin{aligned} \Delta(B)\delta_t &= s_{1t} - \hat{s}_{1t} = \nu_n(B)s_{1t} - \nu_s(B)n_{1t} \\ &= \frac{\sigma_\eta^2|\Delta_s^*(B)\psi_n(B)|^2}{\sigma_a^2|\psi(B)|^2}\Delta_n^*(B)\psi_s(B)\varepsilon_t \\ &\quad - \frac{\sigma_\varepsilon^2|\Delta_n^*(B)\psi_s(B)|^2}{\sigma_a^2|\psi(B)|^2}\Delta_s^*(B)\psi_n(B)\eta_t. \end{aligned} \tag{3.10}$$

Now for δ_t to be stationary, it is necessary and sufficient (the latter since δ_t has no deterministic components) that the function $\Delta(z)$, corresponding to the noninvertible operator on the left hand side of (3.10), be a factor of $\Delta_n^*(z)\Delta_s^*(z)$, the analogous function for the right hand side of (3.10). But this can occur if and only if $\Delta_c(z) = 1$ in (2.8), that is, $\Delta_s(z)$ and $\Delta_n(z)$ have no common roots. We thus have proved

THEOREM 1. *In the model* (2.5)–(2.7), *if* $\hat{s}_t$ *is the conditional expectation of* s_t *given* $\{x_s, -\infty < s < \infty\}$, i.e., *if* $\hat{s}_t = \nu_s(B)x_t$ *with* $\nu_s(z)$ *given by* (3.8), *then the estimation error* $\delta_t = s_t - \hat{s}_t$ *is stationary* (*and has finite mean square*) *if and only if the difference/summation operators* $\Delta_s(B)$ *and* $\Delta_n(B)$ *in the component processes of* x_t *have no common roots of unit modulus.*

COROLLARY. *If* $\Delta_c(z)$ *denotes the common factors of* $\Delta_n(z)$ *and* $\Delta_s(z)$, *then the process* $\Delta_c(B)\delta_t$ *is stationary*; *and no proper factor of* $\Delta_c(B)$ *renders* δ_t *stationary.*

In common usage the term "optimal" estimation (signal extraction, seasonal adjustment) has usually referred to minimum mean square error estimation of s_t. If $\Delta_c(z) \neq 1$ then $\hat{s}_t$ cannot be a minimum MSE estimator of s_t since in this case the MSE of $\hat{s}_t$ does not exist. Thus the notion of optimality needs to be extended to cover this case. Note that the estimate $\hat{s}_t$ is well defined if only the *conditional* expectation $E(s_t|\{x_{t'}\})$ exists, which is so even when the MSE is infinite. Note also that if $\hat{s}_t = E(s_t|\{x_{t'}\})$ and if $\Delta_c(z)$ is common to $\Delta_s(z)$ and $\Delta_n(z)$, then (see the corollary above and the discussion following (3.9)) the series

$$\Delta_c(B)\hat{s}_t = \nu_s(B)\Delta_c(B)x_t$$

is the minimum MSE estimate of $\Delta_c(B)s_t$. These considerations motivate the following

DEFINITION. By an *optimal* estimate of s_t in the model (2.5)–(2.7) is meant the conditional expectation of s_t given $\{x_{t'}\}$ whether or not the mean square error of this conditional expectation exists.

Specific results on the structure of δ_t can be obtained from (3.10), from which it follows that

$$\Delta_c(B)\delta_t = \frac{\sigma_\eta^2 \Delta_2^*(F)|\psi_n(B)|^2 \psi_s(B)}{\sigma_a^2 |\psi(B)|^2} \varepsilon_t - \frac{\sigma_\varepsilon^2 \Delta_n^*(F)|\psi_s(B)|^2 \psi_n(B)}{\sigma_a^2 |\psi(B)|^2} \eta_t. \tag{3.11}$$

Thus the acgf/spectrum of $\Delta_c(B)\delta_t$ is

$$\begin{aligned} f_{\Delta\delta}(z) &= \frac{\sigma_\varepsilon^2 \sigma_\eta^2 |\psi_n(z)\psi_s(z)|^2 \{\sigma_\eta^2 |\Delta_s^*(z)\psi_n(z)|^2 + \sigma_\varepsilon^2 |\Delta_n^*(z)\psi_s(z)|^2\}}{\sigma_a^4 |\psi(z)|^4} \\ &= \frac{\sigma_\varepsilon^2 |\psi_s(z)|^2 \cdot \sigma_\eta^2 |\psi_n(z)|^2}{\sigma_a^2 |\psi(z)|^2}, \end{aligned} \tag{3.12}$$

since the bracketed quantity is the spectrum of $s_{1t} + n_{1t}$ or of x_{1t}.

The result (3.12) is important in determining an explicit expression for the generation of δ_t in terms of its innovations, say ξ_t, which, as in Section 3.1 for the stationary case, exists by virtue of the spectral factorization (e.g., Hannan, 1970, page 129). We summarize this result as

THEOREM 2. *The estimation error δ_t follows the stochastic process*

$$\Delta_c(B)\delta_t = \psi_\delta(B)\xi_t \tag{3.13}$$

where

$$\psi_\delta(z) = \psi_n(z)\psi_s(z)/\psi(z),$$
$$\sigma_\xi^2 = \sigma_\eta^2 \sigma_\varepsilon^2 / \sigma_a^2,$$

ξ_t is white noise and $\psi_\delta(z)$ is one-sided. The acgf/*spectrum of $\Delta_c(B)\delta_t$ is given by* (3.12).

If δ_t is stationary then $\Delta_c(z) = 1$ and asterisks are removed in the right hand side of (3.11), but otherwise the foregoing is unchanged.

Equation (3.12) is formally identical to (3.7) for the stationary case; however, note that the two factors of the numerator of (3.10) are the spectra of $\Delta_s(B)s_t$ and $\Delta_n(B)n_t$, and the denominator is the spectrum of the sum of $\Delta(B)s_t$ and $\Delta(B)n_t$. It is of interest that the error δ_t itself in (3.11) is the sum of two generally noninvertible components (containing $\Delta_s^*(F)$ and $\Delta_n^*(F)$ respectively), but that the spectrum of $\Delta_c(B)\delta_t$ is strictly positive as the zeroes of $\Delta_s^*(z)$ and $\Delta_n^*(z)$ are distinct. Analogous features are characteristic of $\Delta(B)x_t$ itself, as noted following (2.10).

3.3 Illustrations. To illustrate these results, three examples are now considered, each involving component processes which sum to a first order integrated moving average process. A more complex example concerning seasonality is discussed in Section 5.

EXAMPLE 3.1. Suppose the signal s_t is the random walk

$$(1 - B)s_t = \varepsilon_t \tag{3.14}$$

and the noise $n_t = \eta_t$ is white. Then x_t follows the first order integrated moving average

$$(1 - B)x_t = (1 - \theta B)a_t. \tag{3.15}$$

From (3.8)

$$\hat{s}_t = \frac{\sigma_\varepsilon^2}{\sigma_a^2|1 - \theta B|^2} x_t,$$

and since $\Delta(z) = \Delta_s(z) = 1 - z$ and $\Delta_n(z) = 1$, δ_t is stationary.

From (3.12) and (3.13),

$$f_\delta(z) = \frac{\sigma_\eta^2 \sigma_\varepsilon^2}{\sigma_a^2|1 - \theta z|^2} \tag{3.16}$$

and δ_t follows the first order *autoregressive* process

$$(1 - \theta B)\delta_t = \xi_t.$$

Thus, for this example, the mean square error of the optimal signal extraction procedure is

$$E(\delta^2) = \frac{\sigma_\varepsilon^2 \sigma_\eta^2}{\sigma_a^2(1 - \theta^2)}. \tag{3.17}$$

EXAMPLE 3.2. *A "canonical" decomposition.* (The idea of this decomposition, which is to choose, from among the component model specifications consistent with the overall model, the one which minimizes the variance of s_t, is due independently to Pierce (1976) and to Box, Hillmer and Tiao (1976). The term "canonical" is taken from Tiao and Hillmer (1977); in Pierce (1976) it is referred to as the "principle of minimum extraction.") The models for s_t and n_t are in general unidentified given only $\{x_t\}$ and its model. If x_t follows the first order IMA process (3.15), the most general signal-plus-white-noise model for s_t and n_t consistent with this is

$$(1 - B)s_t = (1 - \Theta B)\varepsilon_t, \qquad n_t = \eta_t,$$

and for any Θ value $-1 \leqslant \Theta \leqslant \theta$ there exist values of σ_ε^2 and σ_η^2 such that $s_t + n_t$ yields (3.15). It is shown by Pierce (1976), Box, Hillmer and Tiao (1976), and Wecker (1976) that the value $\Theta = -1$ minimizes the variance of s_t and ε_t. Choosing this value, $\psi_s(B) = 1 + B$, and

$$\hat{s}_t = \frac{\sigma_\varepsilon^2|1 + B|^2}{\sigma_a^2|1 - \theta B|^2} x_t,$$

$$f_\delta(z) = \frac{\sigma_\eta^2\sigma_\varepsilon^2|1 + z|^2}{\sigma_a^2|1 - \theta_z|^2}$$

and

$$(1 - \theta B)\delta_t = (1 + B)\xi_t. \tag{3.18}$$

The extraction error process is thus noninvertible, reflecting the noninvertibility of s_t.

EXAMPLE 3.3. Also consistent with the IMA model (3.15) for x_t is the sum of two random walks, or more generally the sum of two IMA's

$$(1 - B)s_t = (1 - \theta_s B)\varepsilon_t, \qquad (1 - B)n_t = (1 - \theta_n B)\eta_t.$$

The case where sufficient a priori information to identify these component models is available is probably unrealistic; but if the models were known, the optimal estimate (à la Definition 1) is

$$\hat{s}_t = \frac{\sigma_\varepsilon^2|1 - \theta_s B|^2}{\sigma_a^2|1 - \theta B|^2} x_t$$

with infinite mean square error; in fact from Theorem 2 δ_t follows the integrated "ARIMA (1, 1, 2)" process

$$(1 - B)(1 - \theta B)\delta_t = (1 - \theta_n B)(1 - \theta_s B)\xi_t.$$

4. Estimation error of one-sided procedures. In the preceding section it was assumed that in estimating s_t, the future as well as the past of $\{x_t\}$ was available. In many situations it is necessary to estimate s_t given only data on x_s up through $s = t - m$, for finite m. This includes the problems of signal extraction/seasonal adjustment based either on current data ($m = 0$) or on recent data ($m < 0$), and the problem of forecasting the signal/seasonal ($m > 0$). This section examines the nature of the error introduced by the optimal estimation/extraction of s_t given a semi-infinite realization $\{x_s, s \leqslant t - m\}$. The development parallels Section 3 in considering in turn the stationary and nonstationary cases and examples.

4.1 Stationary observable series. Referring to equation (3.1), let

$$\hat{s}_t^{(m)} = \nu_s^{(m)}(B)x_t \tag{4.1}$$

be the minimum MSE estimate of x_t given $\{x_{t'}, t' \leqslant t - m\}$. Then (Whittle, 1963, pages 66–67).

$$\nu_s^{(m)}(z) = \frac{1}{\sigma_a^2\psi(z)}\left[\frac{\sigma_\varepsilon^2|\psi_s(z)|^2}{\psi(z^{-1})}\right]_m = 1 - \nu_n^{(m)}(z) \tag{4.2}$$

where, for any m and $h(z) = \Sigma_{-\infty}^{\infty} h_j z^j$,

$$[h(z)]_m = \Sigma_{j=m}^{\infty} h_j z^j,$$

and $|h(z)|^2 = h(z)h(z^{-1})$ as in Section 3.

The estimation error is now denoted by

$$\delta_t^{(m)} = s_t - \hat{s}_t^{(m)} \tag{4.3}$$

and the series $\delta_t^{(m)}$ is a stationary stochastic process given by (3.4), with acgf/spectrum given by (3.5) (with the superscript (m) adjoined to $\nu_s(B)$ and $\nu_n(B)$). For the MMSE form (4.2), the latter becomes

$$f_\delta^{(m)}(z) = \frac{\sigma_\varepsilon^2|\psi_s(z)|^2}{\sigma_a^4|\psi(z)|^2} N^{(m)}(z) + \frac{\sigma_\eta^2|\psi_n(z)|^2}{\sigma_a^4|\psi(z)|^2} S^{(m)}(z)$$

where

$$S^{(m)}(z) = \left| \left[\frac{\sigma_\varepsilon^2|\psi_s(z)|^2}{\psi(z^{-1})} \right]_m \right|^2$$

and analogously $N^{(m)}(z)$. Note that the simplification (3.6) for the two-sided case does not occur here.

4.2 Nonstationary observable series. Returning to the general model (2.5)–(2.7), the form of (4.2) for nonstationary series is

$$\nu_s^{(m)}(z) = \frac{1}{\sigma_a^2\Psi(z)} \left[\frac{\sigma_\varepsilon^2|\Psi_s(z)|^2}{\Psi(z^{-1})} \right]_m \tag{4.4}$$

where, e.g.,

$$\Psi(z) = \Delta^{-1}(z)\psi(z) = (\Delta^*(z)\Delta_s^*(z)\Delta_n^*(z))^{-1}\psi(z)$$

is the (nonconvergent) infinite moving average representation of x_t. To show that the formal expression (4.4) is, in fact, a well defined stable filter, it is necessary (and sufficient) to show that after cancellation of common factors in (4.4) no forms such as $\Delta^{-1}(z)$ remain. Using (2.8), (4.4) can be rewritten

$$\begin{aligned} \nu_s^{(m)}(z) &= \left\{ \frac{\Delta_n^*(z)}{\sigma_a^2\psi(z)} \right\} \{\Delta_s^*(z)\Delta_c(z)\} \left[\frac{\sigma_\varepsilon^2|\Delta_s^*(z)|^{-2}|\Delta_c(z)|^{-2}|\psi_s(z)|^2}{(\Delta_s^*(z^{-1})\Delta_n^*(z^{-1})\Delta_c(z^{-1}))^{-1}\psi(z^{-1})} \right]_m \\ &= \kappa(z)\cdot\lambda(z)\cdot[\mu(z)]_m \end{aligned} \tag{4.5}$$

where $\kappa(z)$ and $\lambda(z)$ correspond to the quantities in braces and, simplifying further,

$$[\mu(z)]_m = \left[\frac{\sigma_\varepsilon^2\Delta_n^*(z^{-1})|\psi_s(z)|^2}{\psi(z^{-1})} (\Delta_s^*(z))^{-1}(\Delta_c(z))^{-1} \right]_m$$

(The operator $\kappa(z)$ will be used later to combine with the n_t-process in the second term of the error process as in (2.4)). The present task is to show that the product

$$H(z) = \lambda(z)[\mu(z)]_m \tag{4.6}$$

is stable. This will be done by showing that this product differs by a finite amount in a finite number of terms from the operator

$$h(z) = [\lambda(z)\mu(z)]_m = [\sigma_\varepsilon^2\psi^{-1}(z^{-1})\Delta_n^*(z^{-1})|\psi_s(z)|^2]_m \tag{4.7}$$

which is clearly stable. (For the case $m = 0$ and $\lambda(z) = (1 - z)^p$ an equivalent result is derived in Whittle (1963 page 93).) The coefficient of z^k in (4.7) is

$$h_k = \sum_{j=0}^{d} \lambda_j \mu_{k-j}, \qquad k \geq m$$

and 0 for $k < m$. This is a finite sum because $\lambda(z)$ is finite (assumed to be of degree d). Further, in (4.6)

$$H_k = \sum_{j=0}^{D} \lambda_j \mu_{k-j}, \qquad k \geqslant m$$

where

$$D = \min(d, k - m);$$

thus

$$\text{(4.8)} \qquad \begin{aligned} h_k - H_k &= 0, & k < m \\ &= \sum_{j=D+1}^{d} \lambda_j \mu_{k-j}, & m \leqslant k < m + d \\ & & = 0, k \geqslant m + d \end{aligned}$$

and thus $H(z)$, and therefore $\nu_s^{(m)}(z)$, are absolutely convergent.

It remains to investigate whether the mean square error exists, i.e., whether the estimation error $\delta_t^{(m)}$, as defined in (4.3) except that s_t and $\hat{s}_t$ are now nonstationary, is stationary. From (2.4) the second term in $\delta_t^{(m)}$ is, formally,

$$\text{(4.9)} \qquad \nu_s^{(m)}(B)n_t = \lambda(B)[\mu(B)]_m \kappa(B)\left[(\Delta_n^*(B)\Delta_c(B))^{-1}\psi_n(B)\right]\eta_t.$$

Now from (4.5), (4.7) and (4.8) the only differencing/summing operator contained in $\nu_s^{(m)}(z)$ is $\Delta_n^*(z)$. But $\nu_s^{(m)}(B)n_t$ is stationary if and only if $[\Delta_n^*(z)\Delta_c(z)]$ is a factor of $\nu_s(B)$, which, therefore, occurs if and only if $\Delta_c(z)$ is identically 1. Analogous derivations show this is also necessary and sufficient for stationarity of $\nu_n^{(m)}(B)s_t$, and thus of $\delta_t^{(m)}$. Otherwise only $\Delta_c(B)\delta_t^{(m)}$ is stationary. We thus have proved

THEOREM 3. *In the model* (2.5)–(2.7), *if* $\hat{s}_t^{(m)} = \nu_s^{(m)}(B)x_t$ *with* $\nu_s^{(m)}(z)$ *given by* (4.4) *or* (4.5), *the estimation error* $\delta_t^{(m)} = s_t - \hat{s}_t^{(m)}$ *is stationary* (*and has finite mean square*) *if and only if the difference/summation operators* $\Delta_s(B)$ *and* $\Delta_n(B)$ *in* (2.5) *and* (2.6) *have no common roots. If* $\Delta_c(z)$ *is the common factor of* $\Delta_n(z)$ *and* $\Delta_s(z)$, *then the process* $\Delta_c(B)\delta_t^{(m)}$ *is stationary and no proper factor of* $\Delta_c(B)$ *renders* $\delta_t^{(m)}$ *stationary*.

The generalization of the notion of optimality which was given in the definition of Section 3 can be extended to the present case: $\hat{s}_t^{(m)} = \nu_s^{(m)}(B)x_t$ denotes the *optimal* estimate of s_t, given $x_{t'}$, $t' \leqslant t - m$, whether its MSE is infinite or finite. A caution is in order though, for, unlike the two-sided case, this conditional expectation is not preserved under filtering, since the conditioning set changes. For example,

$$(1 - B)E[s_t | x_{t'}, t' \leqslant t - m] = \hat{s}_t^{(m)} - \hat{s}_{t-1}^{(m)}$$

whereas

$$E[(1 - B)s_t | x_{t'} t' \leqslant t - m] = \hat{s}_t^{(m)} - \hat{s}_{t-1}^{(m-1)}.$$

The difference disappears as $m \to -\infty$.

The acgf/spectrum of the estimation error $\delta_t^{(m)}$, when it is stationary, is (omitting the argument)

$$(4.10) \qquad f_\delta^{(m)} = \frac{\sigma_\varepsilon^2 \sigma_\eta^2}{\sigma_a^4} \left\{ \frac{\sigma_\varepsilon^2 |\Delta_n \psi_s|^2}{|\psi|^2} \left| \left[\frac{\tilde{\Delta}_s |\psi_n|^2}{\tilde{\psi}} \Delta_n^{-1} \right]_m \right|^2 + \frac{\sigma_\eta^2 |\Delta_s \psi_n|^2}{|\psi|^2} \left| \left[\frac{\tilde{\Delta}_n |\psi_s|^2}{\tilde{\psi}} \Delta_s^{-1} \right]_m \right|^2 \right\}$$

where $\tilde{h}(z) = h(z^{-1})$, which apparently does not simplify as in the two-sided case (3.12), making the error stochastic process more difficult to characterize. (With common roots in Δ_s and Δ_n, (4.10) gives the spectrum of $\Delta_c(B)\delta_t^{(m)}$, provided asterisks are adjoined to Δ_s and Δ_n). The mean square error of the procedure (of the estimate $\hat{s}_t^{(m)}$) can be found by integrating (4.10), evaluated at $z = e^{i\omega}$, over ω. In practice one would make repeated use of the fast Fourier transform and sum the computed $f_\delta^{(m)}(e^{i\omega_j})$ over the Fourier frequencies ω_j.

4.3 Illustration. It is informative to analyze the estimated signal and its error for the random-walk-plus-white-noise example in which the signal is estimated with a one-sided projection. Given the model (3.14)–(3.15), from (4.4) the optimal estimate of s_t given $\{x_{t'}, t' \leqslant t - m\}$ is

$$\hat{s}_t^{(m)} = \left\{ \frac{\sigma_\varepsilon^2(1 - B)}{\sigma_a^2(1 - \theta B)} \left[\frac{(1 - B)^{-1}}{(1 - \theta F)} \right]_m \right\} x_t.$$

In the notation of Section 4.2, $\lambda(z) = 1 - z$ and

$$\begin{aligned} h(z) &= \left[(1 - z) \frac{(1 - z)^{-1}}{(1 - \theta z^{-1})} \right]_m = \left[\Sigma_0^\infty \theta^j z^{-j} \right]_m \\ &= 0, \qquad m > 0 \\ &= \Sigma_0^{-m} \theta^j z^{-j}, \quad m \leqslant 0 \end{aligned}$$

and thus, from (4.8) or directly,

$$\begin{aligned} H(z) &= (1 - z) \left[\frac{(1 - z)^{-1}}{1 - \theta z^{-1}} \right]_m \\ &= \frac{z^m}{1 - \theta}, \qquad m \geqslant 0 \\ &= \frac{\theta^{-m} z^m}{1 - \theta} + \Sigma_{m+1}^0 \theta^{-k} z^{-k}, \qquad m < 0. \end{aligned}$$

Thus

$$\nu_s^{(m)}(z) = \sigma_\varepsilon^2 H(z)/\sigma_a^2(1-\theta z) \tag{4.11}$$

$$= \frac{\sigma_\varepsilon^2 z^m}{\sigma_a^2(1-\theta)} \Sigma_0^\infty (\theta z)^k, \qquad m \geqslant 0$$

$$= \theta^{-m} \frac{\sigma_\varepsilon^2 z^m}{\sigma_a^2(1-\theta)} \Sigma_0^\infty (\theta z)^k + \frac{1}{1-\theta z} \Sigma_0^{-m-1} \theta^k z^{-k}, \tag{4.12}$$

$$m < 0.$$

Consider first the problems of estimating s_t for the current time period ($m = 0$) and of forecasting $s_t (m > 0)$. Equation (4.11) shows that for either case one applies an exponentially weighted moving average to the observed series, beginning with the most recent data available, but not otherwise depending on the value of m. In other words, the forecast of s_{t+m} at origin t is a constant, just as in the case of an *observable* random walk; but this constant is the current estimated (rather than observed) value, a value determined from the entire record $\{x_s, s \leqslant t\}$. A further analogy is that forecasts of both x and s are EWMA's of present and past x.

For $m < 0$, that is for estimating s_t based on some but not all of the relevant future of x, the filter as above is applied to the furthest forward observation but with a declining weight (θ^{-m}), with the second term in (4.12) playing a stronger role.

From (4.10) the error $\delta_t^{(m)}$ is stationary and posseses acgf/spectrum

$$f_\delta^{(m)}(z) = \frac{\sigma_\varepsilon^2 \sigma_\eta^2}{\sigma_a^4 |1-\theta z|^2} \left\{ \sigma_\varepsilon^2 \left| \left[\frac{1-z^{-1}}{1-\theta z^{-1}} \right]_m \right|^2 + \sigma_\eta^2 |1-z|^2 \left| \left[\frac{(1-z)^{-1}}{1-\theta z^{-1}} \right]_m \right|^2 \right\} \tag{4.13}$$

which approaches (3.16), and its integral (3.17), as $m \to -\infty$.

5. Two seasonal adjustment models. Many seasonal adjustment procedures currently in use, including the widespread Census X-11 program (Shiskin, Young and Musgrave, 1967), estimate the seasonal component s_t of an observed series (1.1) as a symmetric filter (or moving average) of the form (2.1). The historical development of such filters has apparently occured without conscious reference to the class of time series models (2.5)–(2.7) for which they are optimal; but it is probably more than coincidence that the particular models for which they are optimal are fairly well approximated by models fitted to many of the economic and social time series on which they are used.

In particular the linear filter approximation to X-11 (Young, 1968; Wallis, 1974) has been studied by Cleveland (1972) and Cleveland and Tiao (1976). In each work a stochastic model of the form (2.5)–(2.7) is presented which implies a seasonal adjustment filter $\nu_s(z)$ close to that used in X-11. In one of these models (Cleveland, 1972; here referred to as the "C" model) the roots of $\Delta_s(z)$ and $\Delta_n(z)$ are distinct, and in the other (Cleveland and Tiao, 1976; the "CT" model) the factor $(1 - z)$ is common, so that the two models provide a good illustration of the possible behavior of the seasonal adjustment error δ_t. In addition, insofar as observed time series are generated by models similar to these, a study of δ_t can increase our understanding of the X-11 procedure.

The C and CT models, in common with the X-11 procedure itself, regard the nonseasonal (n_t) component as being the sum of trend (p_t) and white noise irregular (e_t) components, so that the observable series is

$$x_t = s_t + p_t + e_t. \tag{5.1}$$

Both decompositions imply overall models for x_t of the form (2.7) with

$$\Delta(B) = (1 - B)^2 S_{12}(B) = (1 - B)(1 - B^{12}) \tag{5.2}$$

where $S_k(B) = \Sigma_{i=0}^{k-1} B^i$. Additionally $\psi(B)$, though of degree 14 in C and 25 in CT, is numerically similar in each model. Thus the models for the observable series are very much alike; the essential difference is rather in the nature of the decomposition.

Consider first the C model. The trend is taken to be of the form

$$(1 - B)^2 p_t = \psi_p(B)\pi_t = (1 + .26B + .30B^2 - .32B^3)\pi_t$$

so that

$$(1 - B)^2 n_t = \psi_p(B)\pi_t + (1 - B)^2 e_t = \psi_n(B)\eta_t \tag{5.3}$$

with $\psi_n(B)$ of degree 3. The seasonal component is of the form

$$S_{12}(B)s_t = \psi_s(B)\varepsilon_t = (1 + .26B^{12})\varepsilon_t. \tag{5.4}$$

Thus $\Delta_s(z) = S_{12}(z)$ and $\Delta_n(z) = (1 - z)^2$, so that their roots are distinct. The mean square of the error δ_t for the C model is therefore finite, and moreover δ_t obeys the stochastic process (3.11).

The spectrum of δ_t was computed based on the particular parameter values given in Cleveland (1972), and it is graphed in Figure 1. Note that the seasonal adjustment error is itself seasonal (peaks at the seasonal frequencies). This is to be expected, as the apparent "overadjustment" from an optimal procedure (Grether and Nerlove, 1970) results in an "underadjustment" in the error of that procedure. (For stationary component series the spectrum of the nonseasonal series n_t is the sum of the spectra of the SA series $\hat{n}_t$ and the error series δ_t). The mean square error, relative to that of the differenced series x_{1t}, was found to be

$$\frac{E(\delta_t^2)}{E(x_{1t}^2)} = (.035)^2.$$

For example, for the log of the money supply $M1$, $(E(\delta_t^2))^{\frac{1}{2}} = .00084$, so that assuming that the overall model for $M1$ is approximately the one given in Cleveland (1972), the standard error of seasonally adjusted monthly M_1 is about .09% of its current level (or 1.1% for an annualized growth rate).

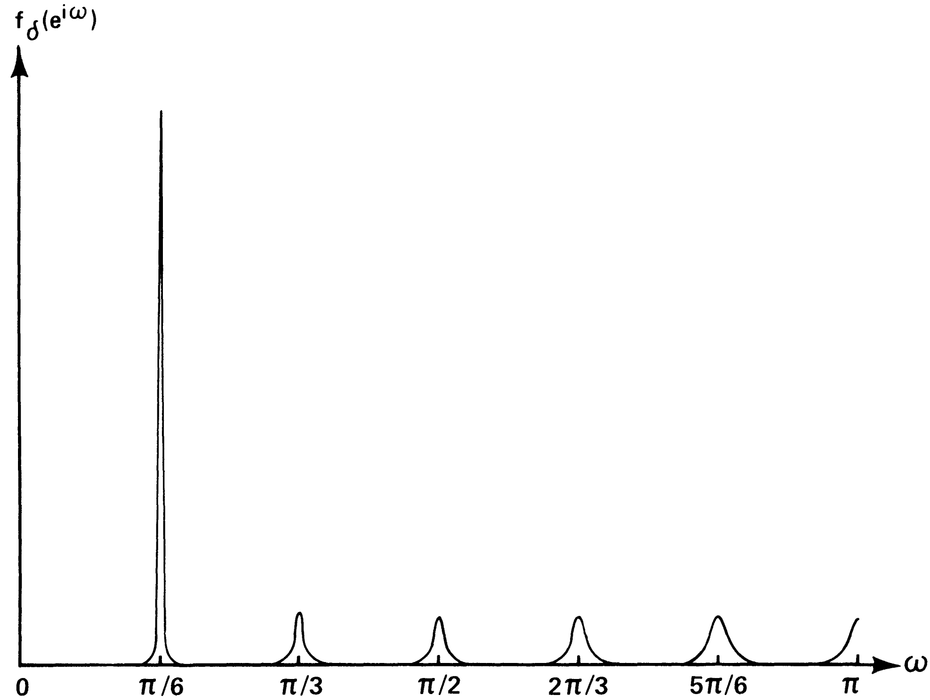

FIG. 1. *Error spectrum for Cleveland* (1972) *model of Census X-II seasonal adjustment procedure.*

The CT model differs from the C model in that the seasonal component s_t is assumed to be generated according to

$$(5.5) \qquad (1 - B)^{12} s_t = \psi_s(B)\varepsilon_t$$

and thus, since $(1 - z^{12}) = (1 - z)S_{12}(z)$, the s_t and n_t models have the factor $(1 - z)$ in common. It follows that the mean square error is infinite, and that the error itself obeys a model of the form

$$(1 - B)\delta_t = [\psi_n(B)\psi_s(B)/\psi(B)]\xi_t.$$

6. Discussion. One can no more observe the extraction error in an unobserved components estimation procedure than one can observe the unobserved components. It is nevertheless important to know something about the probabilistic structure of that error, to understand what can be expected of these procedures. Insofar as many time series in practice are evidently nonstationary but well approximated by linear stationary models after suitable differencing, the results of this paper should serve as a useful step in this direction. For example, they can be used to construct confidence intervals for the signal s_t, or for a "true" seasonally adjusted series. They are only a first step, however, for such reasons as: (i) we almost never know the true model; (ii) for this and other reasons we almost never employ "optimal" procedures; and (iii) the model (2.5)–(2.7) probably does not encompass some of the important time series occurring. We conclude by discussing possible generalizations in each of these three directions.

6.1 Knowledge of the model. "Knowing" the true model implies knowing (a) the model for x and (b) given this, the models for s and n. Exact knowledge of the model for x is almost never available. However, given a sufficient length T of the observable time series, several methodologies are available for estimating this model to a degree of accuracy, roughly speaking, of order $n^{-\frac{1}{2}}$, compared with an accuracy only of order 1 for estimating s_t. Thus, while work is clearly needed on the effects of imperfect estimation of the x-process (perhaps along the lines of Section 6.2), these effects can probably be neglected, except in short series, in a first approximation.

A separate problem is the determination of the decomposition of the observable series x_t into the components s_t and n_t, i.e., the specification of the component models (2.5) and (2.6), given the overall model (2.7). When, as is usually the case, the component models are unidentified from (2.7) alone, additional information such as subject-matter knowledge and the purpose and goals of the decomposition must be considered. The minimal extraction principle ("canonical decomposition") mentioned in connection with Example 2, Section 3.2, is an example of this. Certainly a strong case can be made, from this paper's results, for requiring that the components have distinct unit-modulus autoregressive roots.

6.2 Nonoptimal procedures. Failure to know the true model means that the signal extraction procedure employed (based on an estimated model) will differ from the optimal procedure. Moreover, when a large number of time series need to be seasonally adjusted or otherwise smoothed, it may be infeasible to conduct a thorough investigation into the stochastic structure of each. The Census X-11 procedure, for example, consists of a very limited set of filters used on a wide variety of economic data series. In these situations the extraction procedure employed will deviate from the optimal one, and the resulting mean square error of $\hat{s}_t$ will be increased by model specification/estimation error.

Perhaps equation (2.4) provides a starting point for analyzing the (mean square) error in such situations, though even there, while $\nu_n(z)$ and $\nu_s(z)$ will be known exactly, something must be assumed about s_t and n_t. But a study of the effects of plausible alternative assumptions may add to our understanding of the robustness of the procedure employed. And concerning stationarity of δ_t for nonoptimal procedures, a generalization of Theorems 1 and 3 is

THEOREM 4. *For the estimation error of a (possibly nonoptimal) procedure to be stationary, it is necessary and sufficient that in equation* (2.4) $\Delta_s(z)$ *be a factor of* $\nu_n(z)$ *and* $\Delta_n(z)$ *be a factor of* $\nu_s(z)$. *In this case the* acgf/*spectrum of* δ_t *is given by*

$$f_\delta(z) = \sigma_\varepsilon^2\left|\frac{\nu_n(z)}{\Delta_s(z)}\psi_s(z)\right|^2 + \sigma_\eta^2\left|\frac{\nu_s(z)}{\Delta_n(z)}\psi_n(z)\right|^2. \tag{6.1}$$

6.3 Other models. In Section 1 reasons were given for concentrating on expectations given $\{x_t\}$ only; however, *provided* that viable specifications exist, the case where x_t, s_t and n_t are vectors should be amenable to treatment along the lines of this paper.

Another extension is to allow for *deterministic* effects in the model so that

$$\begin{aligned} x_t &= (s_{1t} + n_{1t}) + (s_{2t} + n_{2t}) \\ &= D_t + S_t \end{aligned}$$

is the sum of a deterministic component D_t and a stochastic component S_t. This model was treated by Pierce (1976) in the context of seasonal adjustment. If the model (6.2) is known, all the results of this paper extend to x_t and are identical to those obtaining for S_t. However, the more complex the model the more difficult it is to elucidate and the larger the estimation/specification error is likely to be (entering into both D_t and S_t) so that this effect may be harder to ignore for moderate-length series.

Acknowledgment. Thanks are due to Gerhard Fries for computing and to the editors and referee for helpful comments. Any views expressed are not necessarily related to those of the Federal Reserve System.

REFERENCES

[1] Box, George E. P., Hillmer, Steven C. and Tiao, George C. (1976). Analysis and modelling of seasonal time series. Presented at the NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C., September 9–10.

[2] Box, George E. P. and Jenkins, Gwilym M. (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

[3] Cleveland, William P. (1972). Analysis and forecasting of seasonal time series. Ph.D. dissertation, Depart. Statist., Univ. Wisconsin.

[4] Cleveland, William P. and Tiao, George T. (1976). A model for the Census X-11 seasonal adjustment program. *J. Amer. Statist. Assoc.*, **71** 581–587.

[5] Geweke, John (1976). The temporal and sectoral aggregation of seasonally adjusted time series. Presented at the NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C., September 9–10.

[6] Granger, Clive W. J. (1976). Seasonality: causation, interpretation and implications. Presented at the NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C., September 9–10.

[7] Grether, David M. and Nerlove, Marc (1970). Some properties of 'optimal' seasonal adjustment. *Econometrica* **38** 682–703.

[8] Hannan, Edward J. (1970). *Multiple Time Series.* Wiley, New York.

[9] Kolmogorov, A. (1941). Interpolation und extrapolation von stationaren zufalligen folgen. *Bull. Acad. Sci., U.S.S.R. Ser. Math.* **5** 3–14.

[10] Nerlove, Marc (1967). Distributed lags and unobserved components in economic time series. In *Ten Essays in Economic Theory*, (W. Fellner, et.al., eds) Wiley, New York.

[11] Pierce, David A. (1976). Seasonal adjustment when both deterministic and stochastic seasonality are present. Presented at NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C., September 9–10.

[12] Pierce, David A. (1977). Relationships—and the lack thereof—between economic time series, with special reference to money and interest rates. *J. Amer. Statist. Assoc.* **72** 11–26.

[13] PORTER, RICHARD D. (1975). Multiple time series containing unobserved components. Special Studies Paper No. 65, Federal Reserve Board.

[14] SHISKIN, JULIUS, YOUNG, ALLAN H. and MUSGRAVE, JOHN C. (1967). The X-11 variant of the Census Method II seasonal adjustment program. Technical Paper No. 15, U.S. Bureau of the Census, February.

[15] TIAO, GEORGE C. and HILLMER, STEVEN C. (1977). Some consideration of decomposition of a time series. Technical Report No. 482, Depart. Statist., Univ. Wisconsin.

[16] WALLIS, KENNETH F. (1974). Seasonal adjustment and relations between variables. *J. Amer. Statist. Assoc.* **69** 18–31.

[17] WALLIS, KENNETH F. (1976). Seasonal adjustment and multiple time series analysis. Presented at the NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C. September 9–10.

[18] WECKER, WILLIAM E. (1976). Discussion. Presented at the NBER-CENSUS Conference on Seasonal Analysis of Economic Time Series, Washington, D.C., September 9–10.

[19] WHITTLE, PETER (1963). *Prediction and Regulation by Linear Least Squares Methods*. English Universities Press, London.

[20] WIENER, NORBERT (1949). *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. Wiley, New York.

[21] YOUNG, ALLAN H. (1968). Linear approximations to the census and BLS seasonal adjustment methods. *J. Amer. Statist. Assoc.* **63** 445–457.

SPECIAL STUDIES SECTION
DIVISION OF RESEARCH AND STATISTICS
BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM
WASHINGTON, D.C. 20551