Intervention Analysis with Applications to Economic and Environmental Problems

G. E. P. BOX and G. C. TIAO*

This article discusses the effect of interventions on a given response variable in the presence of dependent noise structure. Difference equation models are employed to represent the possible dynamic characteristics of both the interventions and the noise. Some properties of the maximum likelihood estimators of parameters measuring level changes are discussed. Two applications, one dealing with the photochemical smog data in Los Angeles and the other with changes in the consumer price index, are presented.

1. INTRODUCTION

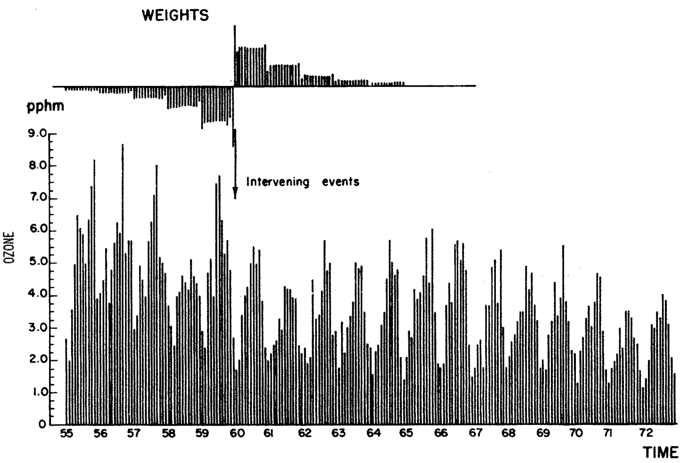
Data of potential value in the formulation of public and private policy frequently occur in the form of time series. Questions of the following kind often arise: "Given a known intervention,[1] is there evidence that change in the series of the kind expected actually occurred, and, if so, what can be said of the nature and magnitude of the change?"

For example, in early 1960 two events occurred, here referred to jointly as the intervention, which might have been expected to reduce the oxidant (denoted by O_3) pollution level in downtown Los Angeles. These events were the diversion of traffic by the opening of the Golden State Freeway and the coming into effect of a new law (Rule 63) which reduced the allowable proportion of reactive hydrocarbons in the gasoline sold locally. The expected effect of this intervention would be to produce a more or less immediate reduction (i.e., a step change) in the oxidant level in early 1960. Figure A shows the monthly averages of oxidant concentration level from 1955–72 in downtown Los Angeles [6]. Using this highly variable and seasonal time series, is there evidence for a change in level and, if so, what is its magnitude?

Many other problems of this kind have come to our attention in recent years. These have included the possible effect of the opening of a nuclear power station on measurements made on river samples, the possible effect of the Nixon Administration's Phases I and II on an economic indicator, and the possible effect of promotions, advertising campaigns and price changes on the sale of a product.

Available procedures such as Student's t test for estimating and testing for a change in mean have played an important role in statistics for a very long time.

A. Monthly Average of Hourly Readings of O_3 (pphm) in Downtown Los Angeles (1955–1972)[a]



[a] With the weight function for estimating the effect of intervening events in 1960.

However, the ordinary t test would be valid only if the observations before and after the event of interest varied about means μ_1 and μ_2, not only normally and with constant variance but *independently*. In the examples quoted, however, the data are in the form of time series in which successive observations are usually serially dependent and often nonstationary, and there may be strong seasonal effects. Thus the ordinary parametric or nonparametric statistical procedures which rely on independence or special symmetry in the distribution function are not available nor are the blessings endowed by randomization.

An approach we initiated earlier [2] was to build a stochastic model which included the possibility of change of the form expected. Such model building is necessarily iterative and, as discussed, e.g., in [3], involves inferences from a tentatively entertained model alternating with criticism of the appropriate tentative analysis. The process proceeds [1] by successive use of Identification (tentative specification of the model form), Fitting, and Diagnostic Checking. Using these ideas in the present context, we come to the following general strategy:

1. Frame a model for change which describes what is expected to occur given knowledge of the known intervention;

* G.E.P. Box is R. A. Fisher professor and G. C. Tiao is professor and chairman, Department of Statistics, University of Wisconsin, Madison, Wis. 53706. This research was supported in part by a grant from the American Petroleum Institute and in part by U.S. Army Research Office under Grant DA-ARO-D-31-124-72-G162.

[1] A term introduced in [5], based on our earlier work [2].

© Journal of the American Statistical Association
March 1975, Volume 70, Number 349
Invited Paper, Theory and Methods Section

2. Work out the appropriate data analysis based on that model;
3. If diagnostic checks show no inadequacy in the model, make appropriate inferences; if serious deficiencies are uncovered, make appropriate model modification, repeat the analysis, etc.

Suppose the data . . . Y_{t-1}, Y_t, Y_{t+1}, . . . are available as a series obtained at equal time intervals. Following, e.g., [1], we will employ models of the general form

$$y_t = f(\boldsymbol{\kappa}, \boldsymbol{\xi}, t) + N_t \tag{1.1}$$

where:

$y_t = F(Y_t)$ is some appropriate transformation of Y_t, say $\log Y_t$, $(Y_t)^{\frac{1}{2}}$ or Y_t itself;
$f(\boldsymbol{\kappa}, \boldsymbol{\xi}, t)$ can allow for deterministic effects of time, t, the effects of exogenous variables, $\boldsymbol{\xi}$, and in particular, interventions;
N_t represents stochastic background variation or noise;
$\boldsymbol{\kappa}$ is a set of unknown parameters.

In Section 2 we discuss a general integrated mixed autoregressive moving average model for representing the noise N_t. A class of general dynamic models capable of representing the effect of interventions is given in Section 3. The associated parameter estimation procedures are given in Section 4. In Section 5 two illustrative examples of intervention analysis are presented. The first concerns the Los Angeles oxidant data, and the second considers possible effects on the consumer price index of recent government actions. Finally, in Section 6, the nature of the maximum likelihood estimators for some specific level-change parameters is discussed in some detail.

2. A STOCHASTIC MODEL FOR THE NOISE

We suppose that the noise $N_t = y_t - f(\mathbf{k}, \xi, t)$ may be modeled by a mixed autoregressive moving average process

$$\varphi(B)N_t = \theta(B)a_t \tag{2.1}$$

where:

1. B is the backshift operator such that $By_t = y_{t-1}$;
2. $\ldots a_{t-1}, a_t, a_{t+1}, \ldots$ is a sequence of independently distributed normal variables having mean zero and variance $(\sigma_a)^2$ which for brevity we refer to as "white" noise;
3. $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \cdots - \theta_q B^q$, $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 \cdots - \varphi_p B^p$ are "moving average" and "autoregressive" polynomials in B of degrees q and p, respectively;
4. the roots of $\theta(B)$ lie outside, and those of $\varphi(B)$ lie on or outside the unit circle.

For the representation of certain kinds of homogeneous nonstationary series, the operator $\varphi(B)$ is factored so that

$$\varphi(B) = (1 - B)^d \boldsymbol{\phi}(B) \tag{2.2}$$

where the roots of $\boldsymbol{\phi}(B)$ all lie outside the unit circle. This corresponds to the use of a stationary model in the dth difference. Also, for seasonal data with period s (e.g., monthly data with $s = 12$), it is often helpful to write $\varphi(B) = \varphi_1(B)\varphi_2(B^s)$ and $\theta(B) = \theta_1(B)\theta_2(B^s)$ with $\varphi_2(B^s) = (1 - B^s)^D \boldsymbol{\phi}_2(B^s)$ to allow for seasonal nonstationarity.

Finally, we entertain a class of noise model of the form

$$\boldsymbol{\phi}_1(B)\boldsymbol{\phi}_2(B^s)(1 - B)^d(1 - B^s)^D N_t = \theta_1(B)\theta_2(B^s)a_t \tag{2.3}$$

where the polynomials $\boldsymbol{\phi}_1(B)$, $\boldsymbol{\phi}_2(B^s)$, $\theta_1(B)$, $\theta_2(B^s)$ are of degrees p_1, p_2, q_1, q_2, respectively.

3. A DYNAMIC MODEL FOR INTERVENTION

Frequently the effects of exogenous variables $\boldsymbol{\xi}$ can be represented by a dynamic model of the form

$$f(\boldsymbol{\delta}, \boldsymbol{\omega}, \boldsymbol{\xi}, t) = \sum_{j=1}^{k} \mathcal{Y}_{tj} = \sum_{j=1}^{k} \{\omega_j(B)/\delta_j(B)\}\xi_{tj} \tag{3.1}$$

where:

1. The $\mathcal{Y}_{tj}$ represent the dynamic transfer from ξ_{tj};
2. The parameters $\boldsymbol{\kappa}$ previously lumped together are now denoted by $\boldsymbol{\delta}$ and $\boldsymbol{\omega}$;
3. The polynomials in B

$$\delta_j(B) = 1 - \delta_{1j}B - \cdots - \delta_{r_j j}B^{r_j} \quad \text{and}$$
$$\omega_j(B) = \omega_{0j} - \omega_{1j}B - \cdots - \omega_{s_j j}B^{s_j}$$

are of degrees r_j and s_j, respectively;
4. We shall normally assume that $\omega_j(B)$ has roots outside, and $\delta_j(B)$, outside or on, the unit circle.

In general, the individual ξ_{tj} could be exogenous time series whose influence needs to be taken into account. For the present purpose, however, some or all of them will be indicator variables taking the values 0 and 1 to denote the nonoccurrence and occurrence of intervention.

For illustration, suppose for a single exogenous variable $(k = 1)$ the model is

$$y_t = \mathcal{Y}_t + N_t = (\omega(B)/\delta(B))\xi_t + (\theta(B)/\varphi(B))a_t \;; \tag{3.2}$$

then the transfer $\mathcal{Y}_t$ to the output from ξ_t is generated by the linear difference equation

$$\delta(B)\mathcal{Y}_t = \omega(B)\xi_t \;.$$

Figures B(a), B(b) and B(c) show the response $\mathcal{Y}_t$ transmitted to the output for various simple dynamic systems by an indicator variable representing a step. We can denote such an indicator by $\xi_t = S_t^{(T)}$ where

$$S_t^{(T)} = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} . \tag{3.3}$$

Similarly, we use $P_t^{(T)}$ for a pulse indicator where

$$P_t^{(T)} = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases} . \tag{3.4}$$

Referring to the figure for the case we have discussed for the Los Angeles 1960 intervention, we would expect that the change could be modelled as in Figure B(a), so that immediately following the known step change in the input, an output step change of unknown magnitude would be produced according to

$$\mathcal{Y}_t = \omega B S_t^{(T)} \;.$$

Sometimes a step change would not be expected to produce an immediate response but rather a "first order" dynamic response like that in Figure B(b). The appropriate transfer function model is then

$$\mathcal{Y}_t = \{\omega B/(1 - \delta B)\}S_t^{(T)} \;,$$

($\delta < 1$). It is readily shown that the time constant of this system is estimated by $\{-\log_e\delta\}^{-1}$ and the steady state gain is $\omega/(1-\delta)$. When δ approaches the value unity, we have the transfer function model

$$\mathcal{Y}_t = \{\omega B/(1-B)\}S_t^{(T)}$$

in which a step change in the input produces a "ramp" response in the output (Figure B(c)).

Note that since

$$(1-B)S_t^{(T)} = P_t^{(T)} , \qquad (3.5)$$

any of these transfer functions could equally well be discussed in terms of the unit pulse $P_t^{(T)}$, and sometimes matters are best thought of directly in terms of $P_t^{(T)}$. Thus, suppose we have monthly sales data and wish to represent the effect of a promotion or advertising campaign lasting less than a month. The simple first order model

$$\mathcal{Y}_t = \{\omega_1 B/(1-\delta B)\}P_t^{(T)}$$

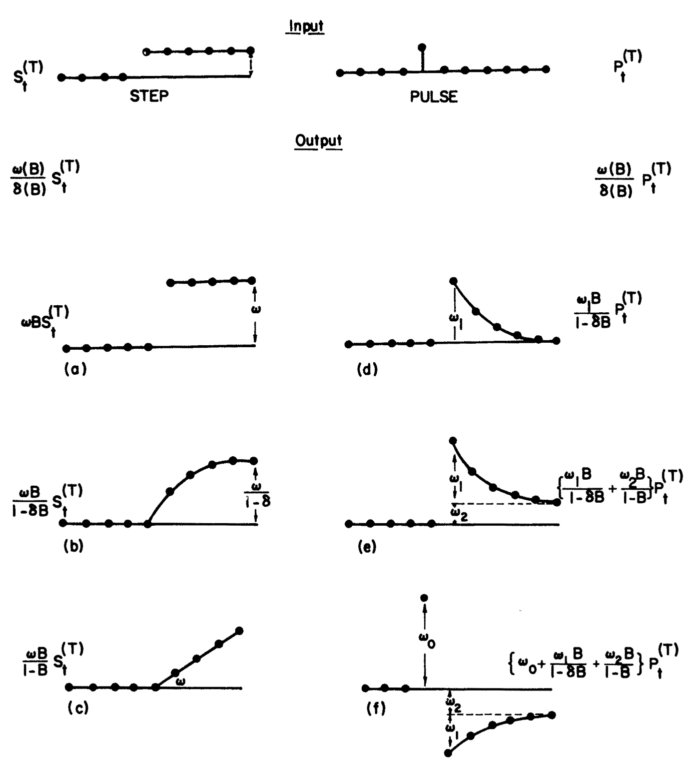
might do this (Figure B(d)) with ω_1 indicating the initial increase in sales immediately following the intervention and δ representing the rate of decay of this increase.

This particular model implies that no lasting effect will occur as a result of the intervention. When this might not be so, the model $B(e)$

$$\mathcal{Y}_t = \{(\omega_1 B/(1-\delta B)) + (\omega_2 B/(1-B))\}P_t^{(T)}$$

could be used in which the possibility is entertained that a residual gain (or loss) in sales ω_2 persists.

B. Responses to a Step and a Pulse Input [a]



[a] (a), (b), (c) show the response to a step input for various simple transfer function models; (d), (e), (f) show the response to a pulse for some models of interest.

If it were believed that the full impact of intervention might not be felt until the second month, after which there would be a decay and possibly a residual effect as in the previous case, the model

$$\mathcal{Y}_t = \{\omega_0 B + (\omega_1 B^2/(1-\delta B)) + (\omega_2 B^2/(1-B))\}P_t^{(T)}$$

might be appropriate. This would insert a preliminary value ω_0 into the output (which in the preceding context would usually be less than ω_1). The same form of model, shifted forward and with some sign changes in the parameters, could be useful to represent the effect of price changes. In the application shown in Figure B(f), ω_0 would be positive and would represent an immediate rush of buying when a prospective price change was announced. The reduction in buying immediately after the change occurred would be represented by $\omega_1 + \omega_2$ and the final effect of the change would be represented by ω_2 which is shown as negative but, of course, could have a zero or positive value.

Obviously, these difference equation models may be readily extended to represent many situations of potential interest.

The following points are worthy of note:

(i) The function $\mathcal{Y}_t$ represents the *additional* effect of the intervention over the noise. In particular, when N_t is nonstationary, large changes could occur in the output even with no intervention. Fitting the model can make it possible to distinguish between what can and what cannot be explained by the noise.

(ii) Intervention extending over several time intervals can be represented by a series of pulses. A three month advertising campaign might be represented, for example, by three pulses whose magnitude might represent expenditure in the three months.

4. CALCULATIONS BASED ON THE LIKELIHOOD

Suppose we entertain a model of the form

$$y_t = \sum_{j=1}^{k} \mathcal{Y}_{tj} + N_t \qquad (4.1)$$

where $\sum_{j=1}^{k} \mathcal{Y}_{tj}$ is the transfer function given in (3.1) associated with known interventions, N_t assumes the form in (2.3), and a time series is available of length $n + d + sD$. Then the likelihood may be obtained in terms of an n dimensional vector $\mathbf{w}$ whose tth element is $w_t = (1-B)^d(1-B^s)^D(y_t - \sum_{j=1}^{k} \mathcal{Y}_{tj})$. The corresponding model for w_t,

$$w_t = \{\theta_1(B)\theta_2(B^s)/\phi_1(B)\phi_2(B^s)\}a_t , \qquad (4.2)$$

is stationary. Thus, following the argument given, e.g., in [1, p. 273], and with the vector $\boldsymbol{\beta}$ having for its g elements the stochastic and dynamic parameters in the model, the likelihood function may be written

$$L(\boldsymbol{\beta}, (\sigma_a)^2|\mathbf{y}) = (2\pi(\sigma_a)^2)^{-(n/2)}|\mathbf{M}|^{\frac{1}{2}} \cdot \exp\{-S(\boldsymbol{\beta})/2(\sigma_a)^2\} \qquad (4.3)$$

where $\mathbf{M}^{-1}(\sigma_a)^2$ is the covariance matrix of the vector w and

$$S(\beta) = \mathbf{w}'\mathbf{M}\mathbf{w} = \sum_{t=-\infty}^{n} [a_t|\mathbf{y}, \beta]^2 \qquad (4.4)$$

with $[a_t|\mathbf{y}, \beta]$ as the expected value of a_t conditional on β and $\mathbf{y}$.

If none of the roots in (4.2) is close to the unit circle, then for moderate and large n, the likelihood is dominated by the exponent. The values of the elements of β minimizing (4.4), which we shall call the *least squares* values, are to a close approximation also the maximum likelihood values. Alternatively, if we introduce a prior distribution such that in the neighborhood where the likelihood is nonnegligible $p(\beta, \sigma_a) \propto p(\beta)(\sigma_a)^{-1}$, we obtain the posterior distribution

$$p(\beta|\mathbf{y}) \propto p(\beta)|\mathbf{M}|^{\frac{1}{2}}\{S(\beta)\}^{-(n/2)} . \qquad (4.5)$$

Again for moderate or large samples and for a noninformative distribution $p(\beta)$, the term involving $S(\beta)$ dominates and approximately

$$p(\beta|\mathbf{y}) \mathrel{\dot{\propto}} \{S(\beta)\}^{-(n/2)} \qquad (4.6)$$

so that the least square estimates correspond with the point of maximum posterior density.

Now if, over the region where the density is appreciable, $S(\beta)$ is approximately quadratic (and in any given case it is easy to check this numerically), then the posterior distribution is approximately a multivariate t. Then,

$$p(\beta|\mathbf{y}) \mathrel{\dot{\propto}} \{1 + (\sum_{ij} S_{ij}(\beta_i - \hat{\beta}_i)(\beta_j - \hat{\beta}_j)/ (n-g)(s_a)^2)\}^{-(n/2)} \qquad (4.7)$$

where

$$S_{ij} = \tfrac{1}{2}\partial^2\{S(\beta)\}/\partial\beta_i\partial\beta_j|_{\beta=\hat{\beta}}$$

and $(s_a)^2 = S(\hat{\beta})/(n-g)$. Thus, for moderate or large n, β is approximately distributed as multivariate normal with mean $\hat{\beta}$ and covariance matrix

$$V(\beta) = (s_a)^2\{S_{ij}\}^{-1} .$$

The square roots of the diagonal elements of $V(\beta)$ will be referred to as standard errors (S.E.).

In practice we may obtain $\hat{\beta}$, $V(\beta)$ and $(s_a)^2$ using a standard nonlinear least squares computer program for the numerical minimization of $S(\beta)$. To do this we need only to be able to compute the quantities $[a_t|\mathbf{y}, \beta]$ for any β and we may proceed as follows. Since the model for w_t is stationary, $[a_t|\mathbf{y}, \beta]$ will be negligible for values $t \leq -Q$ where Q is some suitably chosen positive number. We, therefore, replace $S(\beta)$ by the finite sum $\sum_{t=-Q}^{n} [a_t|\mathbf{y}, \beta]^2$. It is shown in [1] that the initial values $[a_0]$, $[a_{-1}]$, $\cdots$, $[a_{-Q}]$ may often be obtained conveniently by a process of "back forecasting" which also indicates an appropriate value for Q.

5. TWO ILLUSTRATIVE EXAMPLES

The theory developed here is illustrated in this section by two examples, one employing the Los Angeles oxidant data and the other, the rate of change in the United States consumer price index, to determine the effect of known interventions.

5.1 Example 1: The Los Angeles Oxidant Data

Monthly averages of the oxidant (O_3) level in Downtown Los Angeles from January 1955 to December 1972 are shown in Figure A.

Identification (Specification) of the Model. The periods 1955–60 and 1960–65 were regarded as containing no major intervention which would affect the O_3 level. The series themselves and the sample autocorrelation functions within these periods suggest nonstationary and highly seasonal behavior. The autocorrelation functions of such differences $(1 - B^{12})y_t$ taken twelve months apart show significant correlations only at lags 1 and 12. This suggests the following model for the noise N_t:

$$(1 - B^{12})N_t = (1 - \theta_1 B)(1 - \theta_2 B^{12})a_t . \qquad (5.1)$$

Interventions I_1 and I_2 of potential major importance are:

I_1: In 1960 the opening of the Golden State Freeway and the coming into effect of a new law (Rule 63) reducing the allowable proportion of reactive hydrocarbons in locally sold gasoline.

I_2: From 1966 onwards regulations required engine design changes in new cars which would be expected to reduce the production of O_3.

As already argued, I_1 might be expected to produce a step change in the O_3 level at the beginning of 1960. The effect of I_2 might be most accurately represented if we knew, for example, the proportion of new cars having specified engine changes which were in the pool of all cars driven at any point in time. Unfortunately, such data are not available to us presently. We have, therefore, represented the possible effect of intervention as a constant intervention change from year to year reflecting the increased proportion of "new design vehicles" in the car population. As explained more fully in [6], the engine changes would be expected to slow down the photochemical reactions which produce O_3 and, because of the summer-winter atmospheric temperature inversion differential and the difference in the intensity of sunlight, the net effect would be different in winter when oxidant pollution is low from that in summer when it is high.

A model form was, therefore, tentatively entertained for all the available monthly O_3 data from January 1955 to December 1972, which may be conveniently written as:

$$y_t = \omega_{01}\xi_{t1} + \omega_{02}\frac{\xi_{t2}}{1 - B^{12}} + \omega_{03}\frac{\xi_{t3}}{1 - B^{12}} + \frac{(1 - \theta_1 B)(1 - \theta_2 B^{12})}{(1 - B^{12})} a_t \qquad (5.2)$$

where

$$\xi_{t1} = \begin{cases} 0, & t < \text{January, 1960} \\ 1, & t \geq \text{January, 1960} \end{cases}$$

$$\xi_{t2} = \begin{cases} 1, & \text{"summer" months June–October beginning 1966} \\ 0, & \text{otherwise} \end{cases}$$

$$\xi_{t3} = \begin{cases} 1, & \text{"winter" months November–May beginning 1966} \\ 0, & \text{otherwise.} \end{cases}$$

This allows for a step change in the level of O_3 beginning in 1960 of size ω_{01} associated with I_1 and for progressive yearly increments in the O_3 level beginning 1966 of ω_{02} and ω_{03} units, respectively, for the summer and the winter months. This representation is admittedly somewhat crude, and we hope to improve on it as more data become available.

Estimation Results. The maximum likelihood estimates and the associated standard errors are as follows:

Parameter	*MLE*	*S.E.*
ω_{01}	−1.09	.13
ω_{02}	−0.25	.07
ω_{03}	−0.07	.06
θ_1	−0.24	.03
θ_2	0.55	.04

Since examination of residuals $\hat{a}_t$ fails to show any obvious inadequacies in the model, we interpret the results as follows. The marginal distributions *a posteriori* of ω_{01}, ω_{02} and ω_{03} are very nearly normal and centered at the maximum likelihood estimate values with the approximate standard deviations shown.

Thus, there is evidence that

(i) associated with I_1 is a step change of approximately $\hat{\omega}_0 = -1.09$ units in the level of O_3;

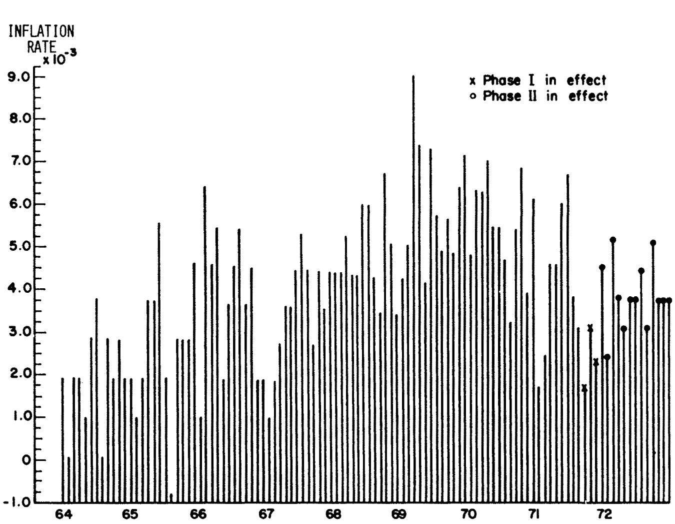
(ii) associated with I_2 there is a progressive reduction in O_3. Over the period studied, there is a yearly increment of approximately $\hat{\omega}_{02} = -.25$ in the summer months, but the increment (if any) in the winter is slight.

5.2 Example 2: The Rate of Change in the U.S. Consumer Price Index

A second example supplies further intuitive appreciation for the kind of calculations being performed.

Figure C shows the latter part of a record of the monthly rate of change in the consumer price index (CPI) given more completely in [4]. The complete (July 1953 to December 1972) data include 234 successive values, 218 of which occurred prior to the institution of controls in August 1971. As indicated in the figure, in the three months beginning September 1971, Phase I control was applied; and after that to the end of the recorded period, Phase II was in effect.

C. Monthly Rate of Inflation of the U.S. Consumer Price Index: January 1964–December 1972



Inspection of the autocorrelation functions of the first 218 observations and their differences prior to Phase I suggests a noise model of the form

$$(1 - B)N_t = (1 - \theta B)a_t \ . \qquad (5.3)$$

The maximum likelihood values for the parameters are:

Parameter	*MLE*	*S.E.*
θ	0.84	.04
σ_a	0.0019	

Inspection of the residuals and their autocorrelations reveals no obvious inadequacies of this model, so we adopt it.

We now ask the question, "What are the possible effects of Phases I and II?" To answer, we suppose:

(i) that Phases I and II can be expected to produce changes in level of the rate of change of the CPI,

(ii) that the form of the noise model remains essentially the same.

On these assumptions, the approximate model (ignoring estimation errors in the noise structure) is

$$y_t = \omega_{01}\xi_{t1} + \omega_{02}\xi_{t2} + \{(1 - .84B)/(1 - B)\}a_t \quad (5.4)$$

where

$$\xi_{t1} = \begin{cases} 1, & t = \text{September, October and November 1971} \\ 0, & \text{otherwise} \end{cases}$$

$$\xi_{t2} = \begin{cases} 1, & t \geq \text{December 1971} \\ 0, & \text{otherwise} \end{cases}$$

which may be written

$$z_t = \omega_{01}x_{t1} + \omega_{02}x_{t2} + a_t \ . \qquad (5.5)$$

The sequences $\{z_t\}$, $\{x_{t1}\}$, $\{x_{t2}\}$ may be readily calculated from the equations

$$(1 - .84B)z_t = (1 - B)y_t$$
$$(1 - .84B)x_{t1} = (1 - B)\xi_{t1}$$
$$(1 - .84B)x_{t2} = (1 - B)\xi_{t2}$$

using, e.g., the initial approximation $z_1 = x_{11} = x_{12} = 0$.

Also, since

$$(1 - B)/(1 - \theta B) = 1 - B(1 - \theta)(1 + \theta B + \theta^2 B^2 + \cdots) \ ,$$

we have

$$z_t = y_t - \bar{y}_{t-1} \ , \quad x_{t1} = \xi_{t1} - \bar{\xi}_{t-1,1} \ , \quad x_{t2} = \xi_{t2} - \bar{\xi}_{t-1,2}$$

where $\bar{y}_{t-1}$, $\bar{\xi}_{t-1,1}$ and $\bar{\xi}_{t-1,2}$ are exponentially weighted moving averages of values prior to time t, e.g.,

$$\bar{y}_{t-1} = (1 - \theta)(y_{t-1} + \theta y_{t-2} + \theta^2 y_{t-3} + \cdots) \ .$$

We see that (5.5) is very much like the regression equations we are all familiar with in which the deviation of y_t from its average is related to the deviations of ξ_{t1} and ξ_{t2} from their averages. Notice, however, that the model copes with nonstationarity by using not the usual arithmetic averages, but local exponentially weighted averages which change as the series progresses.

Using (5.5), the constants ω_{01} and ω_{02} may now be estimated by ordinary linear least squares as

Parameter	*MLE*	*SE*
ω_{01}	−0.0022	0.0010
ω_{02}	−0.0007	0.0009

Alternatively, a nonlinear least squares program may be employed to estimate ω_{01}, ω_{02} and θ simultaneously from the complete set of 234 data values yielding the estimates (essentially as before):

Parameter	*MLE*	*SE*
θ	0.85	.05
ω_{01}	−0.0022	0.0010
ω_{02}	−0.0008	0.0009

The analysis suggests that a real drop in the rate of increase of the CPI is associated with Phase I, but the effect of Phase II is less certain.

6. NATURE OF THE MAXIMUM LIKELIHOOD ESTIMATORS FOR SOME LEVEL CHANGE PARAMETERS

The maximum likelihood estimators of parameters such as ω_{01}, ω_{02} and ω_{03} in (5.2) and (5.4) which measure level changes are functions of the data. It is instructive to consider the nature of these functions. Several results in the summation of series useful in the following discussion are given in the appendix.

6.1 One Parameter "Linear" Dynamic Model

Consider first the dynamic model in (3.2). Formally, it can be written

$$Q(B)y_t = (\varphi(B)/\theta(B))(\omega(B)/\delta(B))\xi_t + a_t \quad (6.1)$$

where $Q(B) = \varphi(B)/\theta(B)$, even though in practice the y_t are only available for $t = 1, \cdots, n$. Since the roots of $\theta(B)$ all lie outside the unit circle, $Q(B)$ can be expressed as a power series in B which converges for $|B| = 1$.

Here we discuss the situation where

$$(\varphi(B)/\theta(B))(\omega(B)/\delta(B)) = \beta R(B) \quad (6.2)$$

and investigate the nature of the maximum likelihood estimator of β, assuming that (i) the coefficients in $Q(B)$ and $R(B)$ are known and (ii) the power series $R(B)$ converges for $|B| = 1$.

Letting

$$z_t = Q(B)y_t \quad \text{and} \quad x_t = R(B)\xi_t ,$$

we can write (6.1) in the form of the usual linear model

$$z_t = \beta x_t + a_t \quad (6.3)$$

so that the maximum likelihood estimator of β is

$$\hat\beta = \sum_{t=1}^{n} z_t x_t / \sum_{t=1}^{n} (x_t)^2$$

with (6.4)

$$\mathrm{Var}\,(\hat\beta) = (\sigma_a)^2 \left(\sum_{t=1}^{n} (x_t)^2\right)^{-1}.$$

For large n, we apply the results (A.6) and (A.7) in the appendix to obtain

$$\sum_{t=1}^{\infty} z_t x_t = \sum_{t=1}^{\infty} Q(B)y_t R(B)\xi_t = \sum_{t=1}^{\infty} \xi_t R(F)Q(B)y_t = R(F)Q(B)C_{\xi y}(0)$$

where $F = B^{-1}$ and

$$\sum_{t=1}^{n} (x_t)^2 = \sum_{t=1}^{\infty} R(B)\xi_t R(B)\xi_t = R(F)R(B)C_{\xi\xi}(0) ,$$

where

$$C_{\alpha\beta}(k) = \sum_{t=1}^{\infty} \beta_t \alpha_{t-k} , \quad k = 0, \pm 1, \pm 2, \cdots ,$$

and for a given k

$$B^l C_{\alpha\beta}(k) = C_{\alpha\beta}(k - l) , \quad l = 0, \pm 1, \pm 2, \cdots .$$

Thus,

$$\hat\beta = R(F)Q(B)C_{\xi y}(0)/R(F)R(B)C_{\xi\xi}(0) \quad (6.5)$$

and

$$\mathrm{Var}\,(\hat\beta) = (\sigma_a)^2/R(F)R(B)C_{\xi\xi}(0) .$$

Making use of (A.10) in the appendix, we can write $R(B)R(F)$ as

$$R(B)R(F) = r_0 + \sum_{l=1}^{\infty} r_l(B^l + F^l) . \quad (6.6)$$

Suppose that $\xi_t = P_t^{(T)}$ is a pulse at time T, and a large number of observations are available before and after T. In this case

$$C_{\xi\xi}(k) = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad \text{and} \quad C_{\xi y}(k) = y_{T-k} , \quad (6.7)$$

so that

$$\hat\beta = (r_0)^{-1}R(F)Q(B)y_T \quad \text{and} \quad \mathrm{Var}\,(\hat\beta) = (\sigma_a)^2(r_0)^{-1} \quad (6.8)$$

where it is understood that B is operating on T.

Now, nonstationarity in time series data can often be removed by differencing. In what follows we suppose that the polynomial $\varphi(B)$ in (6.1) is divisible by $(1 - B)$. We consider two special cases of interest.

Case (i). $\qquad \omega(B)/\delta(B) = \beta B ,$ (6.9)

that is, the pulse input $P_t^{(T)}$ gives rise to a response at time $(T + 1)$ measured by β which dissipates completely after the $(T + 1)$th period. It should be noted that with any number of periods of pure delay, the response will follow the same pattern but be appropriately shifted. In this case, $Q(B) = R(B)F$ so that, from (6.6) and (6.8),

$$\hat\beta = y_{T+1} - \tfrac{1}{2} \sum_{l=1}^{\infty} \lambda_l (y_{T+1+l} + y_{T+1-l}) , \quad (6.10)$$

where $\lambda_l = -2r_l/r_0$. Also, since $\varphi(B)$ is assumed divisible by $(1-B)$, $r_0 + 2\sum_{l=1}^{\infty} r_l = 0$, and hence $\sum_{l=1}^{\infty} \lambda_l = 1$.

As an example, consider the integrated moving average model of order one for the noise term N_t for which

$$\varphi(B) = 1 - B \quad \text{and} \quad \theta(B) = 1 - \theta B \ . \qquad (6.11)$$

Since

$$R(B)R(F) = \frac{(1-B)(1-F)}{(1-\theta B)(1-\theta F)}$$

$$= (1+\theta)^{-1}\cdot[2 - (1-\theta)\sum_{l=1}^{\infty} \theta^{l-1}(B^l + F^l)] \ ,$$

we find that

$$\lambda_l = (1-\theta)\theta^{l-1} \ . \qquad (6.12)$$

Thus, $\hat{\beta}$ represents a comparison between y_{T+1} and the mean of two exponentially weighted averages, one of the observations before time $(T+1)$ and the other after, with the magnitude of the weights $(1-\theta)\theta^{l-1}$ monotonically decreasing as l increases.

This formulation is applicable to situations where the response to the pulse input is expected to be short-lived, e.g., the effect on the demand for electricity during a sudden heat wave in the summer or the sale of beer in Wisconsin should the Packers win the Super Bowl. Essentially, we are comparing the observation y_{T+1} with the neighboring ones to determine if y_{T+1} is an "aberrant" or "outlying" observation. The results in (6.10) and (6.12) are appealing since, in forming the comparison, more weight is given to observations close to the intervening event and less and less weight to observations remote from the time of the event.

Case (ii). $\quad \omega(B)/\delta(B) = \beta B/(1-B) \ . \qquad (6.13)$

Here, the response to the pulse $P_t^{(T)}$ is a step change in the level of the observations measured by β. Thus

$$Q(B) = (1-B)R(B)F \qquad (6.14)$$

and, from (6.6), (6.8) and (A.11), we have that

$$\hat{\beta} = (r_0)^{-1}R(B)R(F)(1-B)y_{T+1}$$

$$= \sum_{l=0}^{\infty} \alpha_l y_{T+1+l} - \sum_{l=0}^{\infty} \alpha_l y_{T-l} \qquad (6.15)$$

where $\alpha_l = (r_0)^{-1}(r_l - r_{l+1})$ so that $\sum_{l=0}^{\infty} \alpha_l = 1$.

The quantity $\hat{\beta}$ is, therefore, a contrast between two weighted averages, one of observations before the intervening pulse $P_t^{(T)}$ and the other afterward, where the weights are symmetrical.

As a first example, consider again the integrated moving average model in (6.11). We find

$$\hat{\beta} = (1-\theta)\sum_{l=0}^{\infty} \theta^l y_{T+1+l} - (1-\theta)\sum_{l=0}^{\infty} \theta^l y_{T-l} \qquad (6.16)$$

as obtained in [2].

As a second example, we return to the model in (5.2) for the monthly averages of ozone in downtown Los Angeles. For illustration, we shall ignore the effect of interventions after 1966 and discuss the step change

$$(\beta B/(1-B))P_t^{(T)} = \omega_{01}\xi_{t1} \ , \quad T = \text{December } 1959$$

in the level of the series due to the intervening events around that time. In this case, the noise model is such that

$$\varphi(B) = (1 - B^{12})$$

and

$$\theta(B) = (1-\theta_1 B)(1-\theta_2 B^{12}) \ .$$

Thus,

$$R(B)R(F) = \frac{(\sum_{j=0}^{11} B^j)(\sum_{j=0}^{11} F^j)}{(1-\theta_1 B)(1-\theta_2 B^{12})(1-\theta_1 F)(1-\theta_2 F^{12})}$$

$$= (\sum_{j=0}^{\infty} \pi_j B^j)(\sum_{j=0}^{\infty} \pi_j F^j) \qquad (6.17)$$

so that from (A.10),

$$r_l = \sum_{j=0}^{\infty} \pi_j \pi_{j+l} \ .$$

The π_j can be obtained from the relationship

$$(1-\theta_1 B)(1-\theta_2 B^{12}) \sum_{j=0}^{\infty} \pi_j B^j = \sum_{j=0}^{11} B^j \ .$$

By writing $\pi_j = 12n + m$, we find

$$\pi_{12n+m} = (1-\theta_1)^{-1}(\phi - \theta_2)^{-1}[(\theta_1)^{m+1}\{(1-\phi)\phi^n - (1-\theta_2)(\theta_2)^n\} + (\phi-\theta_2)(\theta_2)^n] \ ,$$
$$m = 0, \cdots, 11; n = 0, \cdots, \infty \qquad (6.18)$$

where $\phi = (\theta_1)^{12}$.

From (6.18) and after some algebraic reduction, we obtain, on setting $l = 12k + s$,

$$r_{12k+s} = (1-\theta_1)^{-2}(1-(\theta_2)^2)^{-1} \cdot\left[12 - s(1-\theta_2) + \frac{\theta_1(1-\theta_2)^2}{1-(\theta_1)^2} \cdot\left(\frac{\phi(\theta_1)^{-s}}{1-\phi\theta_2} - \frac{(\theta_1)^s}{\phi-\theta_2}\right)\right](\theta_2)^k + (1-\theta_1)^{-2} \cdot(\phi-\theta_2)^{-1}(1-\phi\theta_2)^{-1}(1-(\theta_1)^2)^{-1} \cdot(1-\phi)^2(\theta_1)^{s+1}\phi^k \ , \qquad (6.19)$$

$$s = 0, \cdots, 11; \quad k = 0, \cdots, \infty \ .$$

The resulting weight function for the Los Angeles data is shown in Figure A above the observations.

6.2 The General "Linear" Dynamic Model

The result in (6.5) can be readily extended to the case of more than one parameter. In the general dynamic model with k inputs in (4.1), letting

$$(\varphi(B)/\theta(B))(\omega_j(B)/\delta_j(B)) = \beta_j R_j(B) \qquad (6.20)$$

we can write

$$Q(B)y_t = \sum_{j=1}^{k} \beta_j R_j(B)\xi_{tj} + a_t, \quad t = 1, \cdots, n \qquad (6.21)$$

where, as before in (6.1), $Q(B) = \varphi(B)/\theta(B)$. Assuming that all the coefficients in $Q(B)$ and $R_j(B)$ are known and these $k+1$ power series converge for $|B| = 1$, the model is then linear in the k parameters $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k)'$. It readily follows that, for large n, the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ satisfies the normal equations

$$\mathbf{A}\hat{\boldsymbol{\beta}} = \mathbf{b} \tag{6.22}$$

where $\mathbf{A}$ is a $k \times k$ matrix and $\mathbf{b}$ a $k \times 1$ vector such that

$$\mathbf{A} = [a_{ij}] \ , \quad a_{ij} = R_i(F)R_j(B)C_{\xi_i\xi_j}(0)$$
$$\mathbf{b} = (b_1, \cdots, b_k)'$$

with

$$b_j = R_j(F)Q(B)C_{\xi_j y}(0) \ ; \quad i, j = 1, \cdots, k \ .$$

In what follows, we investigate the special case having two parameters,

$$y_t = \{\beta_1\eta(B)B + \beta_2(1-B)^{-1}B\}P_t^{(T)} + (\theta(B)/\varphi(B))a_t \ . \tag{6.23}$$

In this model, $\beta_1\eta(B)BP_t^{(T)}$, where $\eta(B)$ is assumed to converge for $|B| = 1$, measures the transient effect, and β_2 represents the eventual change in the level of the observations induced by the pulse input $P_t^{(T)}$ (see Figure B(e) for the special case $\eta(B) = (1-\delta B)^{-1}$). When $\beta_1 = 0$, the model reduces to that considered in (6.13). It is, therefore, of particular interest to know to what extent the nature and precision of the estimator of β_2 is affected by the presence of β_1. We again suppose that the noise term is nonstationary so that $\varphi(B)$ is divisible by $(1-B)$.

To facilitate comparison with the model (6.13) we again define a quantity $R(B)$ such that

$$Q(B) = (1-B)R(B)F \ ,$$

so that in (6.22)

$$R_1(B) = Q(B)\eta(B)B = R(B)\eta(B)(1-B)$$

and

$$R_2(B) = R(B) \ .$$

It follows that, provided $|\mathbf{A}| \neq 0$,

$$\begin{aligned} \hat{\beta}_1 &= |\mathbf{A}|^{-1}\{a_{22}b_1 - a_{12}b_2\} \ , \\ \hat{\beta}_2 &= |\mathbf{A}|^{-1}\{a_{11}b_2 - a_{12}b_1\} \end{aligned} \tag{6.24}$$

where

$$|\mathbf{A}| = a_{11}a_{22} - (a_{12})^2 \ ,$$
$$b_1 = R(B)R(F)(1-F)\eta(F)(1-B)y_{T+1} \ ,$$
$$b_2 = R(B)R(F)(1-B)y_{T+1} \ ,$$

and a_{11}, a_{12} and a_{22} are, respectively, the coefficients of B^0 in the power series

$$R(B)R(F)\eta(B)\eta(F)(1-B)(1-F) \ ,$$
$$R(B)R(F)\eta(B)(1-B) \ ,$$
$$R(B)R(F) \ .$$

Some Properties of $\hat{\beta}_1$ and $\hat{\beta}_2$.

(i) Both b_1 and b_2 are linear functions of the observations y_t. By setting $B = F = 1$, the sum of the coefficients associated with y_t is zero for both of these functions. Thus, $\hat{\beta}_1$ and $\hat{\beta}_2$ are linear contrasts in y_t.

(ii) The estimator $\hat{\beta}_2$ can be expressed in the form

$$\hat{\beta}_2 = \sum_{l=0}^{\infty} \alpha_{1l}y_{T+1+l} - \sum_{l=0}^{\infty} \alpha_{2l}y_{T-l}$$

where

$$\sum_{l=0}^{\infty} \alpha_{1l} = \sum_{l=0}^{\infty} \alpha_{2l} = 1 \ , \tag{6.25}$$

i.e., a contrast between two weighted averages, one of observations on or before the pulse input and the other afterward. To see this, since $\hat{\beta}_2$ is a linear contrast, it suffices to show that $\sum_{l=0}^{\infty} \alpha_{1l} = 1$.

From the expression for b_2 in (6.24), letting

$$G(B) = R(B)R(F) \ , \quad H(B) = 1-B$$

and

$$b_2 = \sum_{l=-\infty}^{\infty} d_l y_{T+1-l}$$

it follows from (A.11) that $\sum_{l=-\infty}^{0} d_l = a_{22}$.

Further, making use of (A.12) and (A.13), we see that a_{12} in (6.24) is also the coefficient of B^0 in $R(B)R(F) \cdot (1-F)\eta(F)$. If we now set

$$G_1(B) = R(B)R(F)(1-F)\eta(F) \ , \quad H_1(B) = 1-B$$

and

$$b_1 = \sum_{l=-\infty}^{\infty} d_l^* y_{T+1-l} \ ,$$

we then have $\sum_{l=-\infty}^{0} d_l^* = a_{12}$. The desired result follows since

$$\sum_{l=0}^{\infty} \alpha_{1l} = |\mathbf{A}|^{-1}\{a_{11}\sum_{l=-\infty}^{0} d_l - a_{12}\sum_{l=-\infty}^{0} d_l^*\} = 1 \ .$$

This property is similar to that of $\hat{\beta}$ in (6.15) for the model (6.13), except that the weight functions are no longer symmetrical. From least squares theory, we have

$$\hat{\beta}_2 = \hat{\beta} - (a_{12}/|\mathbf{A}|)(b_1 - a_{12}\hat{\beta}) \ , \tag{6.26}$$

and the second term on the right side measures the effect of the presence of the term $\beta_1\eta(B)BP_t^{(T)}$ in the model.

(iii) One would expect that addition of the parameter β_1 to the model would reduce the precision with which β_2 could be estimated. A useful measure of the loss of information is the variance ratio $\text{Var}(\hat{\beta}_2)/\text{Var}(\hat{\beta})$ where it is understood that the denominator corresponds to the model in (6.13). Now

$$\text{Var}(\hat{\beta}_2)/\text{Var}(\hat{\beta}) = (1-\rho^2)^{-1}$$

where

$$\rho = a_{12}/(a_{11}a_{22})^{\frac{1}{2}} \ . \tag{6.27}$$

We illustrate these results in terms of a specific example. Consider the case of (6.23) in which

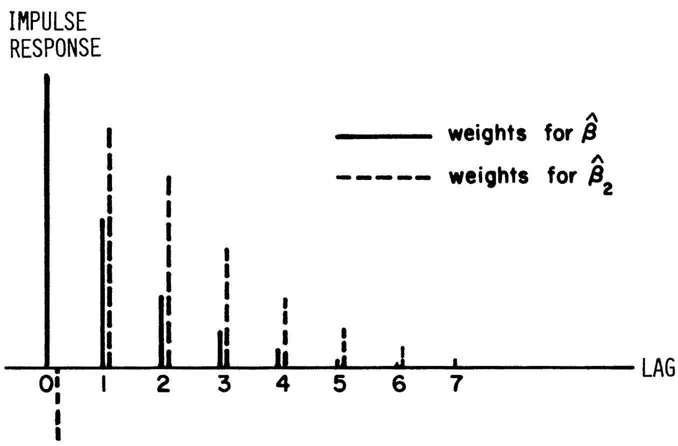
$\eta(B) = (1-\delta B)^{-1}$, $\varphi(B) = 1-B$ and $\theta(B) = 1-\theta B$.

We find

$$\hat{\beta}_2 = \hat{\beta} - \frac{(1-\theta)(1+\delta)}{(\theta-\delta)} \sum_{l=0}^{\infty} [(1-\delta)\delta^l - (1-\theta)\theta^l] y_{T+1+l} \;, \quad (6.28)$$

where $\hat{\beta}$ is given in (6.16). In this case only the weights associated with the observations after the intervening pulse $P_t^{(T)}$ are affected by the presence of $\beta_1(1-\delta B)^{-1}BP_t^{(T)}$ in the model. The weight function is shown in Figure D for $\theta = .5$ and $\delta = .25$.

D. Comparison of Weights Associated with y_{T+1+l} $\hat{\beta}_2$ and for $\hat{\beta}$ ($\theta = .05$, $\delta = .25$, $l = 0, 1, 2, \ldots$)



Also, for this model the variance ratio is

$$V = \mathrm{Var}(\hat{\beta}_2)/\mathrm{Var}(\hat{\beta}) = 1 + ((1-\theta)(1+\delta)/(1+\theta)(1-\delta)) \;. \quad (6.29)$$

The value of this ratio for various values of θ and δ is shown in the following tabulation:

	δ				
θ	$-.5$	$-.25$	0	$.25$	$.5$
$-.5$	2.00	2.80	4.00	6.00	10.00
$-.25$	1.56	2.00	2.67	3.78	6.00
0	1.33	1.60	2.00	2.67	4.00
.25	1.20	1.36	1.60	2.00	2.80
.5	1.11	1.20	1.33	1.56	2.00

Thus, the presence of β_1 in the model can cause large increases in the variance of $\hat{\beta}_2$, compared with $\hat{\beta}$, when θ is negative and δ is positive.

7. CONCLUDING REMARKS

In the past, much attention has been given to statistical analysis linking phenomena which are coincidental in time. In practice, it is perhaps more often the case that a response at a given point of time depends on events, both known and unknown, which have occurred not necessarily coincidentally but over the recent past. Statistical methods have, in a word, "lacked memory." The dynamic characteristics of both the transfer function and the noise parts of the model have tended to be ignored. The application of time series methods can amend this situation. This is illustrated in this article in the particular case where the object is to study the possible effect of interventions in the presence of dependent noise structure.

APPENDIX

We here state some useful results in the summation of series.

Lemma 1: Let $\{v_k\}_0^\infty$ be a sequence of numbers and let $\{x_t\}_{-\infty}^\infty$ and $\{y_t\}_{-\infty}^\infty$ be two sequences of numbers such that $x_t \equiv y_t \equiv 0$ for $t \leq 0$. If one of the following three double sums is absolutely convergent,

$$S_1 = \sum_{t=1}^{\infty}\sum_{k=0}^{\infty} x_t v_k y_{t-k} \;, \quad S_2 = \sum_{u=1}^{\infty}\sum_{k=0}^{\infty} y_u v_k x_{u+k} \;, \quad (A.1)$$

$$S_3 = \sum_{k=0}^{\infty}\sum_{u=1}^{\infty} v_k y_u x_{u+k} ,$$

the other two are absolutely convergent and

$$S_1 = S_2 = S_3 \;.$$

Proof of the lemma can be found in any standard text on infinite series.

It is convenient to express S_1, S_2 and S_3 in terms of the backshift operator B and its reciprocal, the forward shift operator $F = B^{-1}$. Letting

$$V(B) = \sum_{k=0}^{\infty} v_k B^k \quad \text{and} \quad V(F) = \sum_{k=0}^{\infty} v_k F^k \quad (A.2)$$

we can then write

$$S_1 = \sum_{t=1}^{\infty} x_t V(B) y_t \quad \text{and} \quad S_2 = \sum_{t=1}^{\infty} y_t V(F) x_t \;. \quad (A.3)$$

Further, suppose we define

$$C_{xy}(k) = \sum_{t=1}^{\infty} y_t x_{t-k} \;, \quad C_{yx}(k) = \sum_{t=1}^{\infty} x_t y_{t-k} \;, \quad k = 0, \pm 1, \pm 2, \cdots$$

so that

$$C_{xy}(k) = C_{yx}(-k) \;. \quad (A.4)$$

The quantity S_3 in (6.1) can be expressed as

$$S_3 = \sum_{k=0}^{\infty} v_k C_{xy}(-k) \;,$$

and, by letting $C_{xy}(-k) = B^k C_{xy}(0)$, we have

$$S_3 = V(B) C_{xy}(0) \;. \quad (A.5)$$

It follows that when the conditions of Lemma 1 hold,

$$\sum_{t=1}^{\infty} x_t V(B) y_t = \sum_{t=1}^{\infty} y_t V(F) x_t = V(B) C_{xy}(0) \;. \quad (A.6)$$

This result can be readily extended to the following:

Lemma 2: Suppose $W(B) = V_1(B) + V_2(F)$ where $V_1(B)$ and $V_2(F)$ are two power series in B and F, respectively, such that the sum $\sum_{t=1}^{\infty} x_t W(B) y_t$ is absolutely convergent. Then

$$\sum_{t=1}^{\infty} x_t W(B) y_t = W(B) C_{xy}(0) \;. \quad (A.7)$$

Lemma 3: Let $G(B) = \sum_{j=-\infty}^{\infty} g_j B^j$ and $H(B) = \sum_{k=-\infty}^{\infty} h_k B^k$ be two power series in B and converge for $|B| = 1$, and let $D(B)$

$= G(B)H(B)$. Then

$$D(B) = \sum_{l=-\infty}^{\infty} d_l B^l \tag{A.8}$$

where
$$d_l = \sum_{j=-\infty}^{\infty} g_j h_{l-j} .$$

In particular

(i) if $g_j = g_{-j}$ and $h_k = h_{-k}$, then

$$d_l = d_{-l} = \sum_{u=0}^{\infty} h_u g_{u+l} + \sum_{u=1-l}^{\infty} g_u h_{u+l} \;, l = 0, \cdots, \infty \;; \tag{A.9}$$

(ii) if $g_j = 0, j \leq -1$ and $H(B) = G(F)$, then

$$d_l = d_{-l} = \sum_{j=0}^{\infty} g_j g_{j+l} \;, \quad l = 0, \cdots, \infty \;; \tag{A.10}$$

(iii) if $H(B) = 1 - B$, then

$$d_l = g_l - g_{l-1} \qquad l = 0, \pm 1, \cdots, \pm \infty \;, \tag{A.11}$$

so that $\sum_{l=1}^{\infty} d_l = -g_0$ and $\sum_{l=-\infty}^{0} d_l = g_0$;

(iv) if $g_j = g_{-j}$ and $h_j = 0 \quad j \leq -1$, then

$$d_0 = \sum_{j=0}^{\infty} h_j g_j \;; \tag{A.12}$$

(v) if $g_j = g_{-j}$ and $h_j = 0 \quad j \geq 1$, then

$$d_0 = \sum_{j=-\infty}^{0} h_j g_j \;. \tag{A.13}$$

[Received October 1973. Revised August 1974.]

REFERENCES

[1] Box, G.E.P. and Jenkins, G.M., *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day, Inc., 1970.

[2] ——— and Tiao, G.C., "A Change in Level of a Non-stationary Time Series," *Biometrika*, 52 (June 1965), 181–92.

[3] ——— and Tiao, G.C., *Bayesian Inference in Statistical Analysis*, Reading, Mass.: Addison-Wesley Publishing Co., 1973.

[4] Feige, E. and Pearce, D.K., "Inflation and Income Policy: An Application of Time Series Models," Report #7318, Social Systems Research Institute, University of Wisconsin, Madison, 1973.

[5] Glass, G.V., "Estimating the Effects of Intervention into a Nonstationary Time Series," *American Educational Research Journal*, 9, No. 3 (1972), 463–77.

[6] Tiao, G.C., Box, G.E.P. and Hamming, W.J., "Analysis of Los Angeles Photochemical Smog Data: A Statistical Overview," *Journal of Air Pollution Control Association*, 25 (March 1975).