季節變動調整方法에 關한 參考文獻集

第2卷　事前調整, ARIMA模型, 特異項

1986.　7.

調査統計局 統計分析課

季節變動調整方法에 關한 參考文獻集

第2卷　事前調整, ARIMA模型, 特異項

1986.　7.

調查統計局 統計分析課

일 러 두 기

當局에서는 景氣綜合指數, 產業生產指數 等의 各種 月別 經濟指標의 季節變動調整系列을 X-II-ARIMA 方法으로 作成하고 있으며, 個別 指標別 ARIMA 模型選定, 休日의 事前調整 等 보다 나은 季節變動調整을 위한 研究作業을 推進하고 있는바, 當課에서 그동안 參考資料로 活用하여 온 미국, 일본, 캐나다 등의 各種 研究論文 資料, Technical Paper 等을 보다 많은 利用·活用을 위하여 拔萃하여 發刊하는 것임.

分量關係로 2卷으로 나누었으며 收錄內容은 다음과 같음.

- **第1卷.** 季節變動調整方法에 관한 全般的인 內容
- **第2卷.** 事前調整, AR IMA 模型, 特異項 等에 관한 內容

統 計 分 析 課 長

目　　次

Ⅳ. 특이항 (EXTREME VALUES)

Ⅴ. SAS를 이용한 X-II-ARIMA 방법

O.D. Anderson, ed., . . SERIES

FESTIVAL AND WORKING DAYS PRIOR ADJUSTMENTS IN ECONOMIC TIME SERIES

D. Pfeffermann

Central Bureau of Statistics,
Jerusalem,
Israel.

J. Fisher

Bank of Israel,
Jerusalem,
Israel.

A new method of preadjusting time series for the effects of movements in festival dates and the variability in the number of working days is proposed and its properties compared to other commonly used methods are discussed. To examine the efficiency of the proposed method, several statistics are defined and detailed results of applying the method to three important time series are reported.

KEYWORDS: PRE-ADJUSTMENT, WORKING DAYS, FESTIVAL DATES, SEASONAL ADJUSTMENT.

1. INTRODUCTION

Many monthly economic time series behave in a multiplicative fashion

$$O_{ij} = T_{ij}\, S_{ij}\, I_{ij} \qquad (1)$$

where O_{ij} is the original datum for month j of year i, T_{ij} is the trend, S_{ij} is the seasonal factor and I_{ij} the irregular factor.

The irregular factor can be considered as representing, in general, that part of the original observation that cannot be explained by trend and seasonality. More specifically, we can think of this factor as breaking down into three major components.

a) Extreme one-time events such as wars, strikes, special weather conditions, implementation of extreme economic policy measures, etc. Such events often cause short term changes in trend, which are difficult to estimate, and could well affect the seasonal pattern. In particular, the occurence of such events can rarely be forecast.

b) Calendar changes, such as those associated with the moving dates of religious festivals, (which might just cause movements of economic activity between successive months, but could also affect the level of activity itself), changes in the number of working days between different months of the same year and between parallel months of different years, changes in the daily composition of months (the number of Sundays, the number of Mondays ...), relevant if activity alters with the day of the week etc. Calendar changes are characterized by the fact that their effect on time series can be estimated and forecast since the calendar structure is known in advance.

c) Residual irregularity.

The functional relation between the overall irregularity I_{ij} and the three components can take various forms. In case of a multiplicative relation, equation (1) can be expanded as

$$O_{ij} = T_{ij}\, S_{ij}\, P_{ij}\, \text{[illegible]}_{ij} \qquad (2)$$

where P_{ij} represents the identifiable part of the irregularity (described in (b)) and $[\ell_{ij}]$ the remainder, which, provided that no extreme events occur, can be considered as realizations of random variables with a common unit expectation and constant variance.

In this paper, we propose a new approach for preadjusting time series of the general structure (2), for the combined effect of movements in festival dates and/or the variability in the number of working days, termed hereafter "Calendar Effects". "Festival date movements" are typical of Jewish festivals, whose dates are fixed according to the lunar year, but vary, according to the Gregorian calendar, over a maximum of 29 days. The phenomenon is not confined, however, to Jewish festivals; the date of the Christian Easter, for example, exhibits similar variability. The term, 'number of working days', is used in a rather wide sense - it includes all the days in a month which are not officially declared as holidays. This definition does not take into account possible differences in daily levels of activity between days of the week; more refined measures, based on seven daily weights, can be constructed for this purpose.

The above two calendar effects are related. A festival falling in a certain month reduces the number of working days in that month. The two effects, however, do not necessarily overlap and, in many series, the occurence of festivals has a marginal effect beyond that of reducing the number of working days, like as increasing the demand for workers, increasing the extent of customs clearance of imported goods etc. Sometimes, the festival date is dominant and the number of working days itself is of small importance, like in certain series of retail trade, arrivals and departures of tourists etc.

The importance of preadjusting time series for calendar effects is obvious. Sometimes it is done solely in order to obtain smoother and more comparable data but in most cases it serves as a preliminary stage for decomposing series into their different components. As argued in section 3.2, when forecasts are concerned, preadjusting might be more efficient than incorporating the calendar variables into the forecast model as suggested by the so-called 'Intervention Analysis'.

In section 2 we review very briefly some of the methods currently used to adjust for this sort of irregularity and discuss their properties. In section 3 we suggest a different method which consists basically of estimating the separate calendar effects by postulating a regression relation (not necessarily linear) between them and the irregular components I_{ij}, allowing for different relations to hold in different months. In section 4 we test the efficiency of this procedure by considering the results of its application to three important empirical series which exhibit festival date and working days effects.

It should be noted that we concentrate on the multiplicative model (2) because we consider this type of model as a simple but rather good approximation to the behaviour of many series. The proposed procedure can, however, be easily generalized to the additive and other more complex models.

2. CURRENTLY USED METHODS

2.1 Adjusting for movements in festival dates.

Three regression type methods are currently used and a detailed description of them can be found in Baron [1].

The "percentage method", as suggested by Baron, is based on the assumption that movements in the festival date over its possible range imply a transfer of activity between the months affected by the festival date, without changing the annual totals. In order to stabilize the series, the original data for those months are adjusted so as to represent that amount of activity that would have been carried out, had the festival started at the median of the range of its possible dates.

To this end, the ratios $O_{ij} / \sum_{j=1}^{n} O_{ij}$ for any given month j affected by the festival date and where n denotes the number of months affected in a certain year, are regressed against the number of days between the actual starting date of the festival and the median of the dates range.

Basically, a similar idea is employed in the O.E.C.D. method, confined to the case n = 2. Let j and j+1 be these two months. By this method, the estimated irregular factors I_{ij}, i=1, 2 ... obtained by applying a program for seasonal adjustment to the raw data, are regressed against the festival date. Denoting by I^*_{ij} the fitted regression values, the adjusted data for month j are obtained as $O^*_{ij} = O_{ij}/I^*_{ij}$, i=1, 2 ... while for month j+1, $O^*_{i,j+1} = O_{i,j+1}/(1/I^*_{ij})$. Note that annual totals are not necessarily preserved.

The third method is the Burman method which, unlike the other two, does not assume any reciprocal relations between months affected by the festival date. By this method, the ratios $M_{ij} = O_{ij}/(O_{i,j-1} \; O_{i,j+1})^{\frac{1}{2}}$ for any given month with a festival effect are regressed against the festival date and the fitted values M^*_{ij} are divided into the original data in order to preadjust the series.

2.2 Adjusting for the number of working days

The following procedure is known to be widely used and is suggested also by Kendall [8, p.8]. The original data O_{ij} are divided by the ratios $D_{ij} = W_{ij} \Big/ \frac{1}{12} \sum_{j=1}^{12} W_{ij}$ where W_{ij} denotes the number of working days in month (i,j), so as to represent that amount of activity that would have been carried out, had the number of working days been the same for all months of the same year.

Öller [10], in forecasting Finnish foreign trade series preadjusted the original data by regressing them against the number of working days, using logarithmic scales and allowing for changes in the regression coefficients over the years, by applying an Adaptive Regression procedure.

2.3 Discussion

The methods described above have the advantage of being economical and simple to apply, a fact which cannot be ignored when considering that, for purposes of current analysis, many series have to be analysed simultaneously. Applying these methods may, however, lead to undesirable results. Notice first that the above methods are restricted in the sense that conceptually, adjustment can be made either for the effect of the festival date movements or for the effect of the variability in the number of working days but not for both effects simultaneously. As argued in the introduction, the festival date may affect the series in certain months beyond the implied change in the number of working days while in other months only one or neither one of the two effects might be observed.

The percentage method for festival adjustments imposes a symmetric transfer of activity between months affected by the festival date when the date moves from its median. This, however, should not necessarily be the case due, for example, to different seasonal effects in the corresponding months. The same, although to a less extent, can be argued about the O.E.C.D. method. The Burman method is free from any such restrictions. It assumes, however, that the ratios M_{ij} remain constant for fixed festival dates, an assumption which is violated when the number of working days has a marginal influence on the original figures and when the seasonal pattern or trend change along the years.

The methods for preadjusting for the number of working days described in section 2.2 can lead to serious consequences. The underlying assumption behind the first procedure is that the level of activity varies proportionately with the number of working days in every month. Theoretical considerations as well as empirical findings indicate, however, that in many series this assumption is violated and the number of working days affects the level of activity only in some of the months. Cf. Öller [10]. (For examples, see sections 4.2 and 4.3 below.). The proportionality assumption is relaxed by Öller's procedure which does assume, however, that every month is affected by the number of working days and in the same manner.

Another serious drawback of the first procedure is that it interferes in the seasonal part of the working days effect. The small number of working days found for example in February has to be considered as part of its seasonality and the procedure adjusts for this part as well and is thus incorrect. (See section 4.2 for an example.) By Öller's procedure the seasonal effects are incorporated into the model using 12 dummy variables and thus, conceptually, no such interference occurs.

3. THE PROPOSED PROCEDURE

3.1 Description and model justification

As implied by the discussion of the last section, the decomposition of the combined calendar effect P_{ij} into its separate components may have different forms in different months. The differences can be characterised by a single functional relation with different coefficients or by different functional relations. In the simple case, an additive relation with different coefficients in different months can be postulated, i.e.

$$P_{ij} = \alpha_j + \beta_j W_{ij} + \gamma_j F_i \qquad i=1,\ 2\ \ldots \tag{3}$$

where F_i denotes the festival date in year i, measured as the number of days between the actual starting date and the earliest possible starting date of the festival. For months j not affected by the festival date, γ_j is definitely zero.

From our own experience, the additive assumption although simple is appropriate for most empirical series. This, however, is not always the case as demonstrated in section 4.3. Notice also that sometimes two different festivals, not fluctuating in the same manner may affect the same months: an example, a tourist series affected both by the dates of Passover and of Easter, is analysed in section 4.4.

We can rewrite equation (2) in the following way:

$$O_{ij} = C_{ij}\,S_{ij}\,(P_{ij} + P_{ij}\,\ell'_{ij}) = C_{ij}\,S_{ij}\,(P_{ij} + \ell^*_{ij}) \tag{4}$$

where $\ell'_{ij} = \ell_{ij} - 1$ and $\ell^*_{ij} = \ell'_{ij}\,P_{ij}$.

It follows from (1) and (4) that

$$I_{ij} = P_{ij} + \ell^*_{ij} \tag{5}$$

Thus, the irregular factors I_{ij} can be decomposed as the sum of the calendar effects P_{ij} and error terms.

We stated in the introduction that unless extreme events occur, the $\{\ell_{ij}\}$ terms can be considered as random variables with a unit expectation and constant variance.

For fixed j, they can usually be assumed as uncorrelated. (This assumption is tested in the empirical investigation of section 4). Thus,

$$E\,\ell^*_{ij} = 0 \text{ for all } (i,j) \quad , \quad E\,\ell^*_{ij}\,\ell^*_{i'j} = \begin{cases} \delta^2 P^2_{ij} & i = i' \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

Having defined for each j the functional relation between the calendar effect and its components, (see comment 2 below) the preadjustment of the series is accomplished in the following stages:

(A) Estimate the irregular factors I_{ij} using an appropriate program for seasonal adjustments. Denote the estimates by $\hat{I}_{ij}$.

(B) Exclude extreme factors and estimate the calendar effects P_{ij} for each i and j by running the regression model (5), for each j separately, with the irregular factors I_{ij} replaced by their estimators $\hat{I}_{ij}$. (See comments 1 and 3 below.) Denote the estimators by $\hat{P}_{ij}$. By dividing the $\hat{P}_{ij}$ terms into the corresponding original data $\{O_{ij}\}$, the preadjusted series $\{\hat{O}_{ij}\}$ is obtained.

Comments:-

1) A vital stage in the above procedure is the preliminary estimation of the irregular factors I_{ij} and the question is whether these can be separated efficiently from the trend and the seasonal effects. In our empirical investigation, we estimated the irregular factors by applying the Bureau of the Census, X-11 program for seasonal adjustment described in Shiskin, Young and Musgrave [13]. Fase, Köning and Volgenant [7] examined the efficiency of several programs for seasonal adjustments by applying them to empirical and simulated data and found this program as well as the Burman (1965) procedure as generally appropriate. Cleveland and Tiao [4] proposed a stochastic model for which the census procedure is nearly optimal. This model seems to have a wide empirical basis. Pearce [11] noted the robustness of the X-11 program to various other models. It is clear, however, that in certain cases the program would perform very poorly and every series has therefore to be treated separately.

As a general criterion for the efficiency of the proposed preadjustment procedure, we argue that we would expect it to reduce the amount of irregularity presented in the series, but only insignificantly affect the trend and the seasonal pattern found before any prior adjustment has been made. In section 4.1 we define some simple statistics for appreciating the performance of the procedure which are based on this argument. Obviously, the 'goodness of fit' of the regression models is a further important indicator for the efficiency of the procedure in each case.

2) When describing the two stages of the above procedure, we assumed that the functional relations between the calendar effects and their components are known for every j. In practice, this knowledge can be based either on theoretical considerations or on studying the behaviour of the estimated irregular factors $\hat{I}_{ij}$. It should be emphasized on the other hand that treating each month separately reduces the number of degrees of freedom in each regression considerably and thus confines the amount of possible structures that can be selected.

3) As indicated in the description of stage B, the estimation of the calendar effects should be based only on non-extreme irregular factors. For most empirical series, the points of time in which extreme events occurred are known and in any case, these extreme values can also be identified empirically. (Most of the programs used for seasonal adjustments identify extreme values

as part of the process involved.) The X-11 program produces, as part of the output, irregular factors modified for extreme events. We have not yet investigated, however, the impact of using these modified factors. Alternatively, robust estimation procedures can be tried.

3.2 Comparison to currently used methods

This section completes the discussion of section 2.3, where we reviewed the main properties of the currently used methods. When comparing these methods to the above proposed method, the following points have to be noticed:

1) The proposed procedure allows for the adjustment of the combined calendar effect P_{ij} for each i and j. It is flexible in choosing the structural relation between the combined effect and its separate components, allowing for different relations to hold in different months.

2) As with the Burman method, the proposed procedure does not impose any sort of "reciprocal relations" between months affected by the movement in the festival date. On the other hand, when such relations exist, they should be obtained when applying this procedure. (See example in section 4.4.)

3) The proposed procedure does not interfere with that part of the working days effect which is seasonal. It stabilizes the seasonal effect of each month to that obtained at the "average" number of working days of that specific month over all years and not to the average over all months in a certain year. Likewise, the proposed procedure does not affect the seasonality stemming from the festival occurance itself.

4) A disadvantage of the proposed procedure compared to the currently used methods is the increase in the amount of skill and in the associated cost required for applying it. However, once the "correct" models have been identified, the supplementary stages are very simple and quick and can be built into any seasonal adjustment or forecast program.

In the discussion to the paper "Signal extraction - the last word in seasonal adjustments?" presented by P. Burman at the R.S.S. conference in Oxford (April '79) it was suggested that the irregularity caused by the movements in the date of Easter could be handled by incorporating Easter dates as an exogenous variable in the stochastic model fitted to the data, in an analogous way to what is known in the literature as "Intervention Analysis", Box and Tiao [3]. Obviously, the same can be argued about the other calendar effects but this might result in having to identify and estimate models which are too heavy and complex, especially when taking into consideration that the structural relation between the combined calendar effect and its components may differ between months. It seems therefore that preadjusting the data and fitting a model to the adjusted series might practically be more efficient and in any case, the proposed procedure can be considered as a useful tool in the identification process.

The above arguments are further supported by various empirical studies which indicated that fitting stochastic models to seasonal adjusted data can sometimes be more efficient than incorporating the seasonal effects into an overall model, (Makridakis and Hibon [9], Raveh and Tapiero [12]). Again, it is our feeling that this phenomenon stems from the identification stage i.e. it can sometimes be much simpler to identify a model for the seasonally adjusted data than from the raw data, particularly when the seasonal pattern is more complex, e.g. the seasonal effects are both additive and multiplicative. Such seasonal models were analysed by Durbin and Murphy [5] and Raveh and Tapiero [12].

4. EMPIRICAL RESULTS

In the following subsections we present and briefly discuss some empirical results as obtained for three selected time series. Besides being important economic indicators, each of the series exhibits different aspects of the problem under consideration.

4.1 Testing for the efficiency of the procedure

We argued before that the efficiency of the proposed procedure should be tested in three different levels:

i) The amount of reduction in the irregularity of the series

ii) Non-interference with the general seasonal pattern and the trend curve as obtained before any adjustment has been made

iii) The goodness of fit of the regression models on which the calendar effects estimation is based.

In this subsection we describe the statistics used for this purpose. To simplify the notation, we use the single index t to denote the various months of observations rather than the double index (i,j) and let T denote the length of the series. No distinction is being made between the "true" factors C_t, S_t ... and their estimators, (obtained by applying the X-11 program).

Bongard [2, p.179] proved that approximately

$$\bar{O}^2 \cong \bar{C}^2 + \bar{S}^2 + \bar{I}^2 \qquad (7)$$

where $\bar{O}^2 = \frac{1}{T-1} \sum_{t=2}^{T} |O_t - O_{t-1}|/O_{t-1}$ and $\bar{C}^2$, $\bar{S}^2$ and $\bar{I}^2$ are defined accordingly.

Provided that (7) holds, $\bar{C}^2/\bar{O}^2$ = R.C.C. measures the relative contribution of the trend to the average monthly percentage change without regard to sign in the original data and similarly for the ratios R.C.S. = $\bar{S}^2/\bar{O}^2$ and R.C.I. = $\bar{I}^2/\bar{O}^2$. Denote by $\bar{C}_a^{\,2}$ and $\bar{S}_a^{\,2}$ the analogue expressions of $\bar{C}^2$ and $\bar{S}^2$ as obtained after preadjusting the series. Non-interference with the seasonal pattern and the trend implies that R.C.C_a $\cong$ R.C.C. and R.C.S_a $\cong$ R.C.S.

When approximately also

$$\bar{O}^2 \cong \bar{C}_a^{\,2} + \bar{S}_a^{\,2} + \bar{P}^2 + \bar{\ell}^2 \qquad (8)$$

where P_t and ℓ_t denote the preadjustment factor and the residual irregular factor for month t and $\bar{P}^2$ and $\bar{\ell}^2$ are defined similarly to $\bar{O}^2$, reduction in the irregularity presented in the series means that the ratio R.C.E. = $\bar{\ell}^2/\bar{O}^2$ is considerably smaller than the ratio R.C.I. Provided that (8) holds, the ratio R.C.P. = $\bar{P}^2/\bar{O}^2$ denotes the relative contribution of the calendar factors to the variability in the original data and can be considered as an indicator of the strength of the calendar effects.

Following the above arguments, we present for each series examined, the following statistics:

1) The ratios R.C.I. and R.C.E. as test statistics for the reduction in the irregularity (in percentages).

2) The ratios R.C.C, R.C.C_a, R.C.S, R.C.S_a as test statistics for non-interference with the seasonal pattern and the trend (in percentages).

3) The ratio R.C.P. as an indicator for the strength of the calendar effects (in percentages).

4) The ratios $O_{(7)} = \bar{O}^2/(\bar{C}^2 + \bar{S}^2 + \bar{I}^2)$ and $O_{(8)} = \bar{O}^2/(\bar{C}_a^{\ 2} + \bar{S}_a^{\ 2} + \bar{P}^2 + \bar{\ell}^2)$

as measures of the closeness of the approximations (7) and (8) and thus as validations for the use of the ratios presented in 1) to 3). Note that all the above ratios are computed by the X-11 program.

In addition to 1), the standard deviations of the irregular factors I_t and the residual irregular factors ℓ_t are presented. As a further test for non-interference with seasonality, we computed the following statistic for the year $i_o = 1977$.

$$N.I.S = \frac{1}{12} \sum_{j=1}^{12} |S_{i_o,j} - S^a_{i_o,j}| / \frac{1}{11} \sum_{j=1}^{11} |S_{i_o,j+1} - S_{i_o,j}|$$

Where $S^a_{i,j}$ denotes the seasonal factor for month (i,j) as obtained after preadjusting the series. Obviously, N.I.S. = 0 when no interference does occur and we would expect it to take in general very low values if the proposed procedure is to be proved efficient. Note, on the other hand, that, in general, the more irregular the original series is, the less can its components be separated efficiently. This implies that the N.I.S statistic would generally be larger, the more irregular the original series is. The N.I.S statistic is based on only one year because usually, the seasonal factors vary only slightly along the years.

To test for the goodness of fit of the regression models, we present the regression coefficients and their associated standard deviations, the squares of the multiple correlations and the Durbin-Watson test statistics for autocorrelations, with the corresponding 5% upper bounds. (Most of the fitted regressions do not contain an intercept. The upper bounds remain, however, the same as in the case of a non zero intercept, Farebother [6]. The above statistics are presented only for months with significant calendar effects.

4.2 The index of industrial production, Israel, 1969-1977

It is conventional to adjust data on industrial production for changes in the number of working days between months and build an index of production per "standard month". This is done by dividing the original data by the ratios D_{ij} defined in section 2.2. As discussed in section 2.3, the underlying assumption behind this procedure is that the monthly level of activity varies proportionately with the number of working days, for every given month. In fact, the index of industrial production is based on a series of indicators of which one of the most important is sales revenue, deflated by some price indices. It cannot be assumed a priori that sales revenue is influenced by the number of working days especially since revenue data include sales from stocks.

For several technical reasons, the data analysed were confined to the years 1969-1977 implying a maximum of only 9 observations for each regression. It became evident, however, that only four months are affected by the number of working days and that other two months are affected by festival dates movements, a fact not realized before. In the following Table 1 we present the empirical results obtained for this series. The regression models fitted are those obtained from (3). The coefficients were estimated by O.L.S.

TABLE 1: INDEX OF INDUSTRIAL PRODUCTION : REGRESSION RESULTS, TESTS FOR REDUCTION IN IRREGULARITY AND FOR NON-INTERFERENCE WITH SEASONALITY AND TREND(1).

Month	Coefficients (S.D)			Corr.	D.W.(5% U. bounds)
J	Const.	W.D.	F.D.	R^2	DW
April		4.22(.03)		.27	3.0 (.98)
May		3.93(.01)		.68	2.23(1.02)
June	39.49(6.19)	2.34(.24)		.94	1.50 (.98)
July		3.75(.02)		.28	2.57 (.98)
September	95.40(2.51)		.62(.17)	.64	2.39(1.02)
October	109.76(1.87)		-.64(.14)	.77	1.09 (.98)

Standard deviations of irregulars	S.D.I = 3.9	S.D.E = 3.1
Relative contributions to variation of original series	R.C.C = 1.7	$R.C.C_a$ = 1.5
	R.C.S = 57.36	$R.C.S_a$ = 61.2
	R.C.I = 30.23	R.C.E = 15.9
		R.C.P = 11.55
Closeness of approx.	O_7 = 1.12	O_8 = 1.10
Non-interference with seasonality		N.I.S = .07

1) The statistics presented in the table are defined in section 4.1.

The results seem to be very significant. The closeness of O_7 and O_8 to unity implies that about 12% of the variation in the original series is explained by calendar movements. By applying the proposed procedure, only 16% of the variance of the original data remains unexplained, compared to 30% when no preadjustment is carried out. Table 1 indicates also that the general seasonal pattern and the trend curve are in fact unaffected by preadjusting the series in this way.

It is interesting to note that the absolute values of the festival date coefficients in September and October are virtually identical, implying an accordingly reciprocal relation between the festival effects in the two months, (but not transfer of activity under the multiplicative model (2)).

For comparison, we preadjusted the series also by the conventional procedure described before i.e. by dividing the original data O_{ij} by the ratios $\bar{P}_{ij}$. The following results obtained:

$R.C.C_a$ = 1.3, $R.C.S_a$ = 26.1, R.C.E = 19, R.C.P = 38,
S.D.E = 3.6, N.I.S. = .71, O_8 = .74

The low value of O_8 implies that the first four ratios have in fact no meaning in this case. The results however are very striking. Adjusting the series in this way completely changes the seasonal pattern and makes the calendar 'effects' dominate all the other components. For example, the seasonal factor obtained for

February 1977 after applying this procedure is 109.7 compared to 97.3 obtained when no adjustment is made and 97.2 when the series is preadjusted as proposed. This issue was discussed in 2.3. Note that the standard deviation of the residual irregular factors, S.D.E = 3.6 is still higher than that obtained by the proposed procedure.

4.3 Workseekers at labor exchanges, Israel, 1962-1977

A workseeker is any person who registers at the labour exchange at least once in the month for employment. Thus, we do not generally expect to find working days effects in this series. On the other hand, since aggregate demand increases prior to festivals, so does the demand for labor, implying in turn a temporary decrease in the number of workseekers. Consequently, the series is expected to be affected by the movements in the festivals dates. In fact, festival effects were found both for passover (whose starting date range is from March 27 to April 25) and for the group of festivals falling in September and October (with starting date range from September 6 to October 5 for the first festival and from September 20 to October 19 for the latest).

Figure 1.b indicates very clearly that the passover effect is linear but confined to about 10 days before the actual starting day of the festival. For years in which passover starts after April 10 (15 days after the earliest possible starting day), all the extra demand for workers is concentrated in April and so the monthly data are independent of the exact festival date. (The point denoted by d corresponds to April 76 in which labor exchanges were on strike.)

Assuming that the point denoted by c (March 64) is an outlier, a similar pattern can be observed in Figure 1.a for March where in addition, a working days effect is observed for the period starting on April 11, while dominated by the festival occurance in the period before. (The phenomenon is not so evident as in April and, in fact, several other models could be identified in this case. As any such model should be in concordance with the model holding for April, the model we adopted seems to be at least a good approximation.)

Actually, the following models proved to fit the data

$$I_{i3} = \alpha_3 + \beta_3 W_{1i_3} + \gamma_3 F_{1i} + \varepsilon_{i3} \qquad (10)$$

$$I_{i4} = \alpha_4 + \beta_4 W_{2i_4} + \gamma_4 F_{1i} + \varepsilon_{i4} \qquad (11)$$

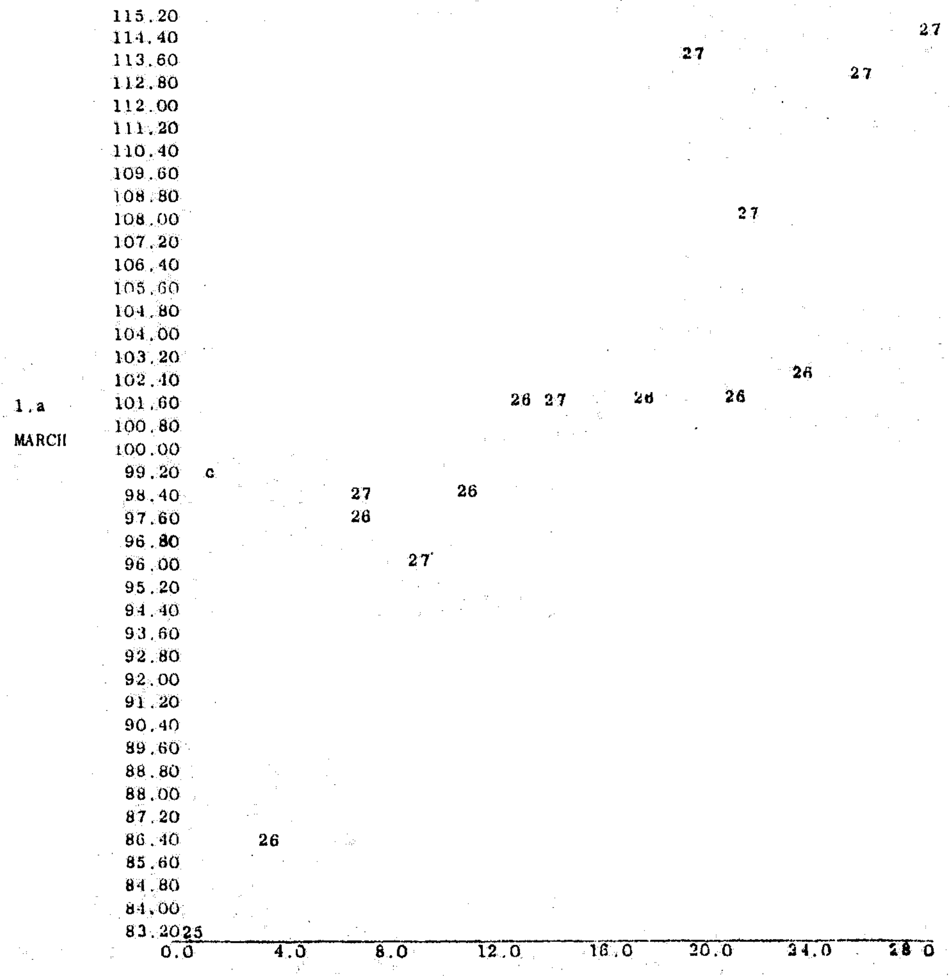
where

$$F_{1i} = \begin{cases} F_i & F_i \le 15 \\ 0 & \text{otherwise} \end{cases} \qquad W_{1i_3} = \begin{cases} W_{i3} & F_i > 15 \\ 0 & \text{otherwise} \end{cases} \qquad W_{2i_4} = \begin{cases} 1 & F_i > 15 \\ 0 & \text{otherwise} \end{cases}$$

The estimators $\hat{\beta}_3$ and $\hat{\beta}_4$ appear in the following table 2 in the column headed W.D. (Working Days). The estimators $\hat{\gamma}_3$ and $\hat{\gamma}_4$ appear in the column headed F.D. (Festival Dates). The estimators were obtained by O.L.S.

As mentioned before, festival effects were found also in September and October. Anyhow the special pattern observed in March and April does not repeat for these months and the relations seem to be approximately linear, as easily explained by the range of dates of the various festivals occuring in these two months and affecting the data.

Figure 1: Workseekers at labor exchanges, scattergrams of the irregular factors observed in March and April (on the vertical axes) against passover dates (on the horizontal axes)[1].



(1) The numbers indicating the points are the number of working days in the corresponding months. The symbols c and d are explained above.

Figure 1 (continued)

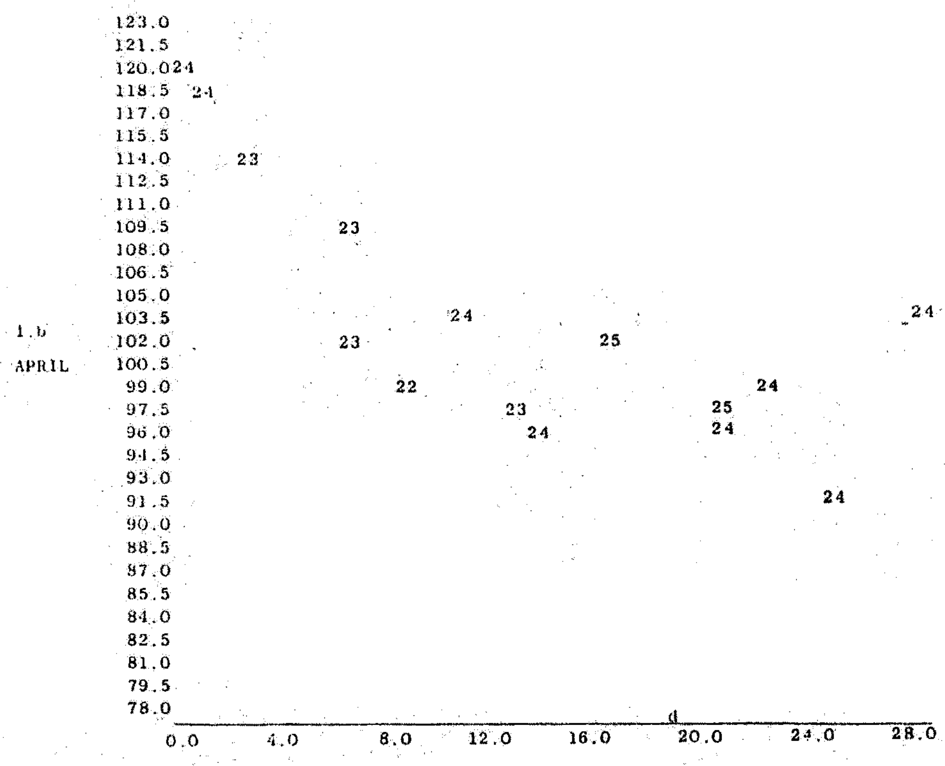


TABLE 2: WORKSEEKERS AT LABOR EXCHANGES: REGRESSION RESULTS, TESTS FOR REDUCTION IN IRREGULARITY AND FOR NON-INTERFERENCE WITH SEASONALITY AND TREND(1).

Month	Coefficients (S.D.)			Corr.D.W.(5% U.bounds)	
J	Constant	W.D.(2)	F.D.(2)	R^2	D
March	84.42(4.43)	.89(.13)	1.36(.35)	.79	2.02(1.36)
April	118.68(6.79)	-16.87(5.75)	-1.65(.56)	.45	2.24(1.36)
June		3.96(.15)		.73	1.47(1.20)
August		3.77(.016)		.51	1.83(1.23)
September	94.50 (.79)		.33(.05)	.82	2.12(1.18)
October	112.08(2.91)		-.94(.19)	.68	1.68(1.18)

(1) The statistics presented in the table are defined in section 4.1

(2) For March and April, these variables are defined by (12)

TABLE 2 (Continued)

Standard deviations of irregulars	S.D.I = 6.2	S.D.E = 4.1
Relative Contributions to variation of original series	R.C.C = 1.67	$R.C.C_a$ = 1.65
	R.C.S = 55.4	$R.C.S_a$ = 54.12
	R.C.I = 34.5	R.C.E = 15.37
		R.C.P = 13.85
Closeness of approx	O_7 = 1.09	O_8 = 1.18
Non interference with seasonality	N.I.S = .14	

Although the value of O_8 is relatively high, the results seem again to be very significant. Note that for this series, the absolute size of the festival date coefficients varies considerably within the two months of each pair affected by the movements in the festivals dates.

4.4 Tourists arrivals by air, Israel, 1959-1977

Series of airline passengers have been widely analysed in the time series literature. The special interest in the present series stems from the fact that the data of March and April are affected by the moving dates of both Passover and Easter which, while for most years are close to each other, are widely separated for the other years. (In 1959, 1967 and 1970 Passover starts towards the end of its dates range while Easter starts towards the beginning of its range.)

The festival effect in this series is expected to be negative in March and positive in April i.e. the earlier any one of the festivals starts, the more tourists are expected to arrive in March and less in April. To take into account the combined effect of the moving dates of Passover and Easter, we expand (3) to

$$P_{ij} = \alpha_j + \gamma_j F_i + \theta_j F_i^* \qquad j = 3,4 \ , \ i = 1,2 \ldots \qquad (13)$$

where F_i^* denotes the date of Easter in the i^{th} year.

TABLE 3: TOURISTS ARRIVALS BY AIR: REGRESSION RESULTS, TESTS FOR REDUCTION IN IRREGULARITY AND FOR NON-INTERFERENCE WITH SEASONALITY AND TREND(1).

Month	Coefficients (S.D.)			Corr.	D.W. (5% U. bounds)
J	Constant	Passover D.	Easter D.	R^2	D
March	119.60(3.73)	-.65(.17)	-.86(.17)	.73	2.06(1.38)
April	80.48(2.57)	.65(.13)	.75(.14)	.85	2.39(1.38)

(1) The statistics presented in the table are defined in section 4.1

TABLE 3 (Continued)

Standard deviations of irregulars	S.D.I = 9.8	S.D.E = 9
Relative contributions to variation of original series	R.C.C = .35	$R.C.C_a$ = .35
	R.C.S = 89	$R.C.S_a$ = 91
	R.C.I = 7.3	R.C.E = 5.3
		R.C.P = .51
Closeness of approx.	O_7 = 1.03	O_8 = 1.03
Non interference with seasonality	N.I.S = .025	

The present series is very seasonal and the calendar effects explain only .51% of the variability in the original data. The preadjustment decreases, however, considerably the irregularity presented in the series without affecting the seasonal pattern and the trend curve. It is interesting to note that for this series, the preadjustment changed only slightly the annual totals. The largest change was of .25%. Since people determine their holiday period either on fixed dates or accordingly to events like the occurance of festivals, this is a straightforward result.

ACKNOWLEDGEMENT

This paper was accomplished while D. Pfeffermann was visiting at the University of Southampton, supported by a training grant from the Friends of the Hebrew University of Jerusalem, the Michael and Anna Wix Trust. The authors are grateful to Miss E. Salameh for her valuable assistance in the empirical investigation.

REFERENCES

[1] Baron, R., Analysis of trend and seasonality in statistical series, methodology and applications in Israel, Israel Central Bureau of Statistics, Technical paper no. 39, (1973).

[2] Bongard, J., Notes on the characteristics of seasonal variation in Europe (1960), seasonal adjustment on electronic computers. Organization for Economic Cooperation and Development.

[3] Box, G.E.P. & Tiao, G.C., Intervention analysis with applications to economic and environmental problems, Journal of the American Statistical Association 70 (1975) 70-79.

[4] Cleveland, W.P. & Tiao, G.C., Decomposition of seasonal time series: A model for the Census X-11 program, Journal of the American Statistical Association 71 (1976) 581-587.

[5] Durbin, J. & Murphy M.J., Seasonal adjustment based on a mixed additive multiplicative model, Journal of the Royal Statistical Society A 138 (1975) 385-410.

[6] Farebrother, R.W., The Durbin Watson test for serial correlation when there is no intercept in the regression, Discussion paper, Department of Econometrics and Social Statistics, University of Manchester (1978).

[7] Fase, M.M.G., Koning, J. & Volgenant, A.F., An experimental look at seasonal adjustment, De Economist 121 (1973) 441-480.

[8] Kendall, M.G., Time Series (2nd ed. Griffin, London, 1975).

[9] Pierce, D.A., comment to the paper: A comparison and assessment of seasonal adjustment methods for employment and unemployment statistics by E.B. Dagum, Background paper No. 5 of the National Commission on Employment and Unemployment Statistics (1979).

[10] Öller, L.E., Time series analysis of Finnish foreign trade, (The Finnish Statistical Society, 1978).

[11] Raveh & Tapiero, Forecasting with a mixed additive multiplicative seasonality model. A nonmetric approach, submitted to the Journal of the Royal Statistical Society A (1979).

[12] Shiskin, J., Young, A.H. & Musgrave, J.C., The X-11 variant of census method II seasonal adjustment program. Bureau of the Census, U.S. Dept. of Commerce, Technical Paper No. 15 (1967).

Instructions for Using the Holiday Version of X-11 (X-11 AYE)

All options and procedures contained in Technical Paper No. 15 remain in force. The holiday version of X-11, called X-11 AYE, requires an additional holiday option card for each series. This card is needed whether or not one desires to perform holiday adjustment.

The card is of the following form:

Column 1: 'A'

Column 2: Easter holiday options:

- '1' = linear regression adjustment
- '2' = cubic regression adjustment
- '3' = the so-called "ideal" adjustment. This procedure is described in the enclosure entitled "Holiday Adjustment of Retail Sales."

Blank = no adjustment for Easter holiday.

Column 3: Labor Day options:

- '1' = linear regression adjustment

Blank = no adjustment for Labor Day.

Column 4: Thanksgiving - Christmas options:

Options are the same as for Labor Day.

This card is placed immediately behind the standard option card.

Holiday Adjustment of Retail Sales

Monthly retail sales series for many types of business contain variation which is related to the date of a particular holiday. The seasonal adjustment process removes the major part of this variation, that which recurs in the same month each year. An example is the high level of sales every year associated with Christmas buying. In making holiday adjustment the concern is not with the variation removed by the seasonal pattern, but with that variation "in relationship" between adjoining months caused by the variable date of the holiday. Such variation appears in the irregular component just as trading day variation does.

Holiday adjustment is justified if holiday variation is actually present in the series and its removal will result in (a) the month to month variation in the seasonally adjusted data being reduced so that the trend cycle will be more clearly revealed and (b) the estimates of trading day variation being more clearly revealed. Note table 1 for summary measures of certain retail sales kinds of business with and without holiday adjustment.

The importance of holiday variation relative to other types of month-to-month fluctuations is shown in table II for selected kinds of business. In each series the holiday variation is several times as large as the monthly variation in the trend cycle component. Holiday variation is especially important when attempting to assess the underlying cyclical movement over short time spans (1 or 2 months). Over longer spans it is of less importance, since it does not cumulate as do seasonal and cyclical movements.

In the standard ratio to moving average methods of seasonal adjustment, the unadjusted data is sequentially separated into three components, trend cycle, seasonal, and irregular. When holiday variation is present in the unadjusted data, it is primarily included in the estimate of the irregular component, rather than in the seasonal or trend cycle, since it closely resembles random movement. To estimate holiday variation from internal evidence, it is therefore, necessary to examine the estimate of the irregular component provided by the ratio to moving average seasonal adjustment.

The holiday adjustments currently being applied to the irregular component of most nondurables kind-of-business series are:

I. Easter

The shifting of Easter between March 22 and April 25 causes retail sales to be high in April and low in March when Easter is late and to be high in March and low in April when Easter is early. The method used to determine the effect of the date of Easter upon March and April sales is a modification of the "ideal" procedure described by Marris in "The Measurement of Calendar Variation" Seasonal Adjustment on Electronic Computers, (OECD, 1960, pp 356-59). It can be briefly described as follows: For each year the March irregular value and the "reciprocal" of the April irregular value is assigned a number (t) corresponding to the date of Easter for that year (e.g., if Easter is March 22 t=1, March 23 t=2, etc.). If Easter variation is present, both the March irregular and the "reciprocal" of the April irregular for a given year will be high when Easter is early and low when Easter is late. The theory behind the "reciprocal" values is that if Easter occurs in March, April sales will be affected exactly counter to March sales. For example, if Easter is March 29 for some year and the April irregular is 97.2, the "reciprocal" is 102.8. This process gives two data points for each year for which data is available. The relationship is described by line segments as follows:

1. An initial horizontal segment is determined by calculating the simple average of data points for years when Easter occurs between March 22 and

April 1 inclusive.

2. An ending horizontal segment determined by calculating the simple average of data points for April 16-25 inclusive.

3. For the period April 2-8 a simple average is calculated and charted at April 5. A line segment is passed through this point and the right end point (April 1) of the initial segment. A similar average is calculated for April 9-15 and charted at April 12. A line segment is passed through this point and the end point (April 16) of the ending segment.

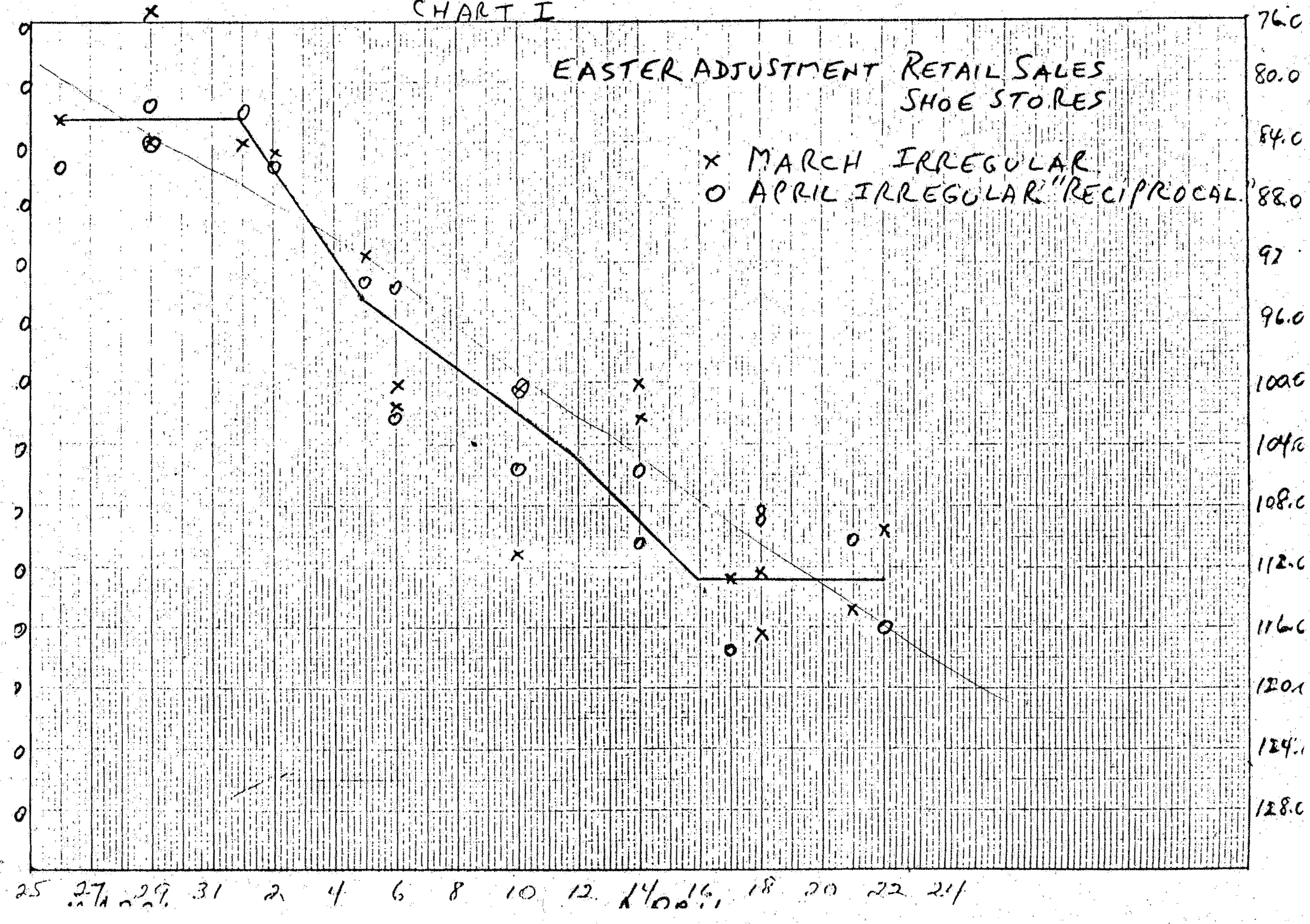
4. A line segment is passed through the April 5 average and the April 12 average to describe this interval.

In fitting the lines extreme irregular values are identified and excluded. For each year the value of the fitted curve serves as the March adjustment and the "reciprocal" as the April adjustment. These factors are used to prior adjust the data before proceeding to trading day and seasonal adjustment.

II. Labor Day and Thanksgiving

For both holidays a simple linear regression line is fitted to irregulars of one month and the "reciprocal" of the other month involved. The values of the fitted curves are then used as prior adjustment factors similar to the Easter factors. The Labor Day adjustment is based on the theory that September and August sales are influenced by whether Labor Day occurs early or late in September. When Labor Day is early (on or near September 1), certain retail sales are concentrated in August rather than September. Chart II shows the regression fit of the Labor Day adjustment for Shoe Stores. The Thanksgiving adjustment is based on the theory that people begin their Christmas shopping after Thanksgiving. If Thanksgiving is early, the sales for November will be relatively high and the sales for December will be lower; if the date is late, the sales for November will be relatively low and the sales for December will be high. Chart III illustrates the Thanksgiving Christmas variation for Department Stores.

CHART I


EASTER ADJUSTMENT RETAIL SALES
SHOE STORES
× MARCH IRREGULAR
O APRIL IRREGULAR "RECIPROCAL"
76.0
80.0
84.0
88.0
92
96.0
100.0
104.0
108.0
112.0
116.0
120.0
124.0
128.0
25
27
29
31
2
4
6
8
10
12
14
16
18
20
22
24

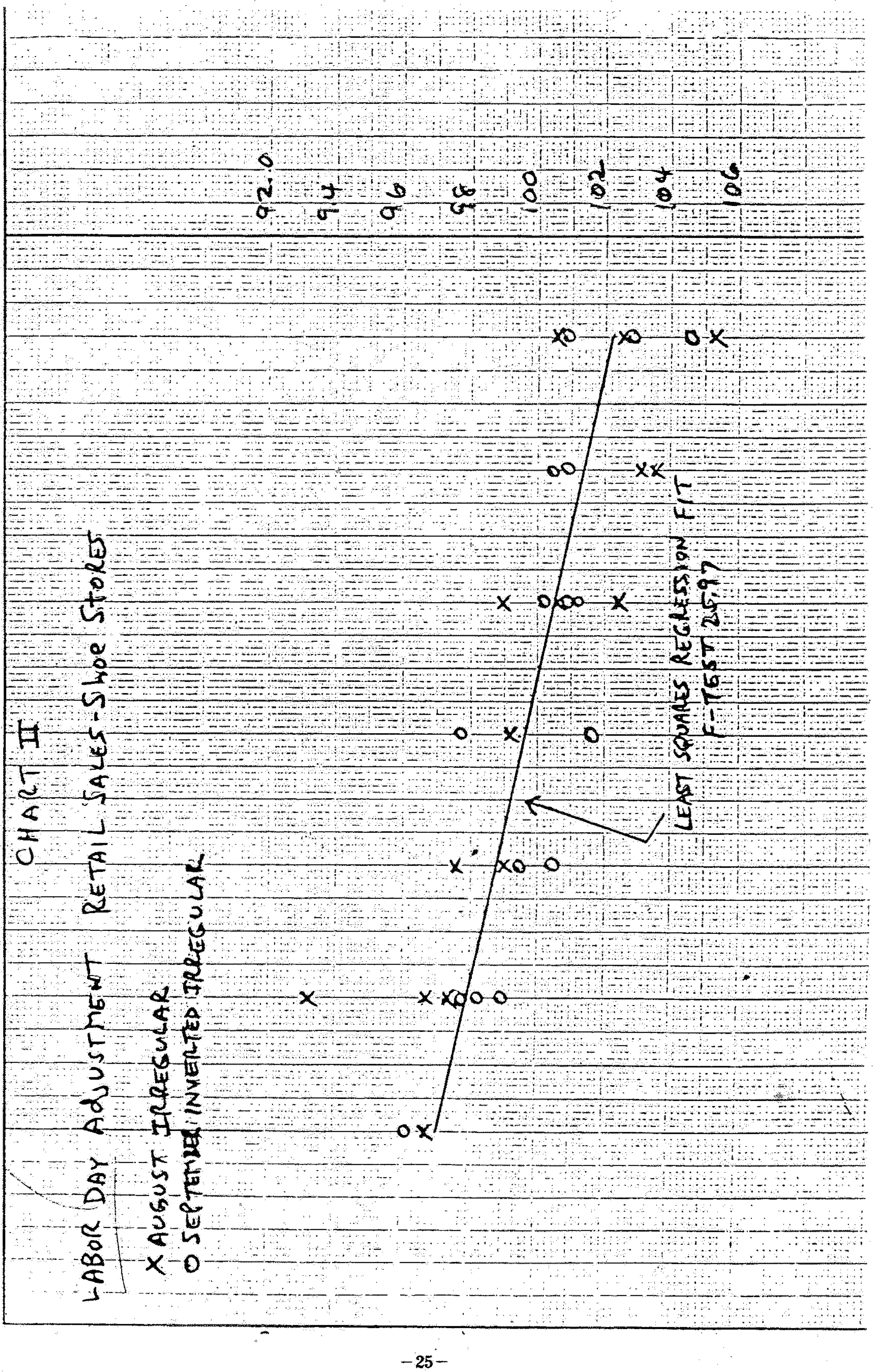


CHART II
LABOR DAY ADJUSTMENT RETAIL SALES-SHOE STORES
X AUGUST IRREGULAR
O SEPTEMBER INVERTED IRREGULAR
LEAST SQUARES REGRESSION FIT
F-TEST 25.97
92.0
94
96
98
100
102
104
106

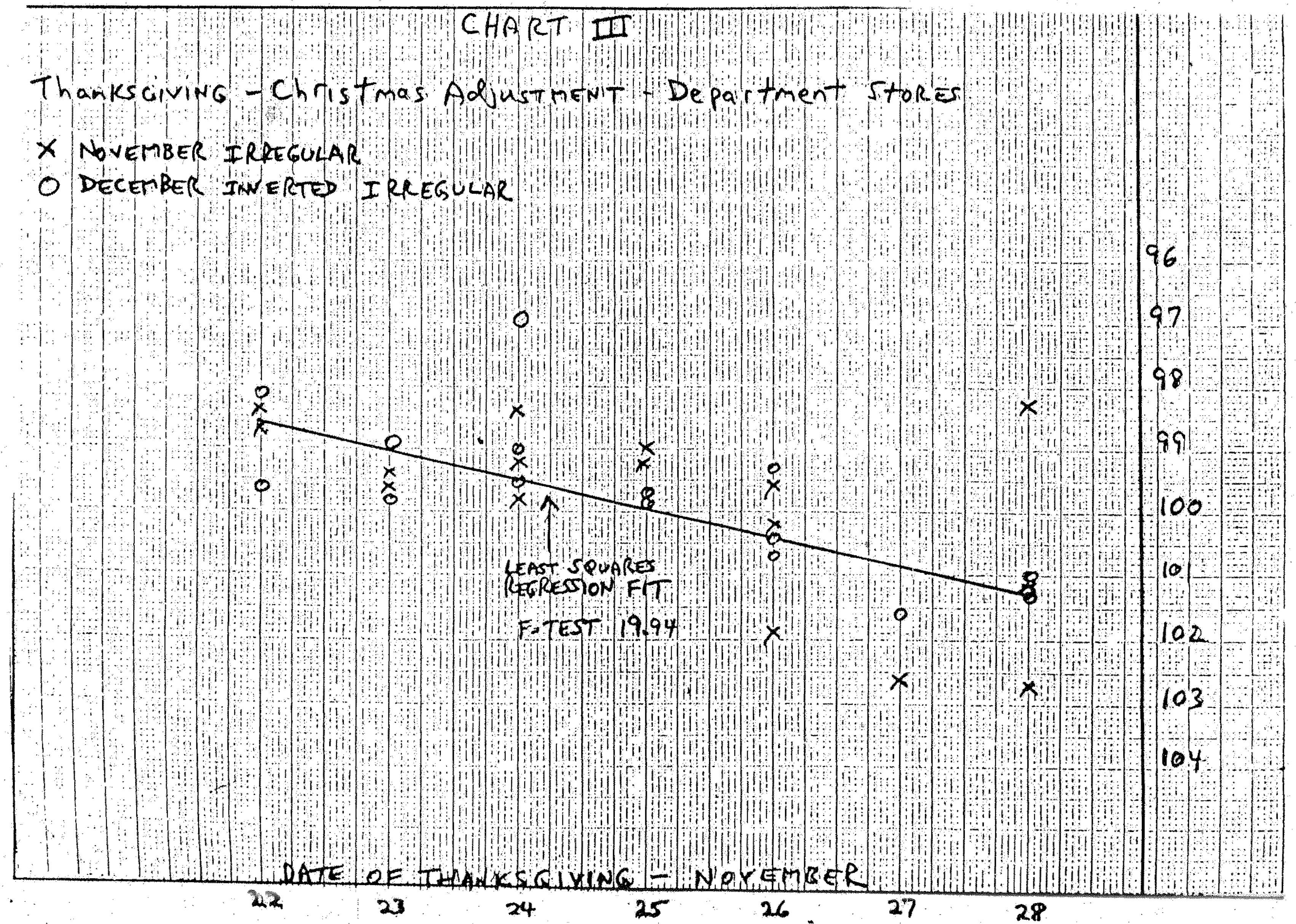


CHART III
Thanksgiving - Christmas Adjustment - Department Stores
X NOVEMBER IRREGULAR
O DECEMBER INVERTED IRREGULAR
96
97
98
99
100
101
102
103
104
LEAST SQUARES REGRESSION FIT
F-TEST 19.94
DATE OF THANKSGIVING - NOVEMBER
22
23
24
25
26
27
28

TABLE I. SUMMARY MEASURES OF RETAIL SALES KINDS OF BUSINESS WITH AND WITHOUT HOLIDAY ADJUSTMENTS

(JANUARY 1953 - DECEMBER 1962)

	$\overline{CI}$	$\overline{C}$	$\overline{I}$	MCD
Mens and Boys Apparel Wear				
(a) With adjustment	2.63	.70	2.43	4
(b) Without adjustment	3.87	.65	3.66	7
Womens Apparel and Accessories				
(a) With adjustment	2.13	.56	2.02	4
(b) Without adjustment	3.63	.56	3.50	7
Shoe Stores				
(a) With adjustment	2.47	.54	2.41	5
(b) Without adjustment	5.58	.56	5.47	11

$\overline{CI}$ Average month - month percent change w/o regard to sign in seasonally adjusted series.

$\overline{C}$ " " " " " " " " " " trend cycle component.

$\overline{I}$ " " " " " " " " " " irregular component.

MCD Months for cyclical dominance (i.e. that span of month-month percent change for which $\overline{I}/\overline{C}$ is less than 1).

TABLE 11. IMPORTANCE OF HOLIDAY VARIATION RELATIVE TO OTHER TYPES OF MONTH-TO-MONTH FLUCTUATIONS

Percent distribution of variation attributable to the various components of month-to-month variation for selected retail sales kinds of business

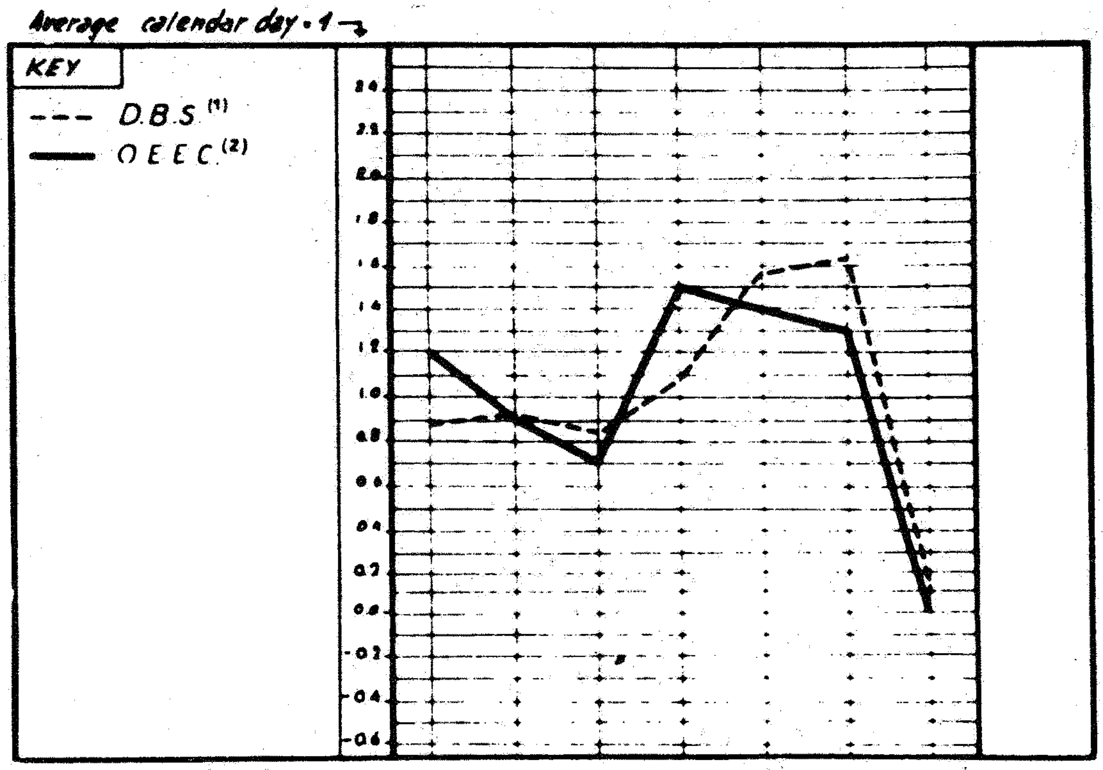
(January 1953 - December 1968)

Kind of Business	Holiday variation	Trading-Day variation	Seasonal variation	Irregular variation	Trend-cycle variation	Total variation
	Components of Mo. - Mo. Variation					
Department Stores	.67	6.20	92.11	.87	.16	100.0
Mail Order Houses	.57	12.63	84.34	2.31	.15	100.0
Men's and Boys Apparel Wear	.79	4.40	93.26	1.43	.12	100.0
Women's Apparel Accessory Stores....	1.53	6.24	90.71	1.42	.11	100.0
Shoe Stores	5.08	6.54	86.56	1.74	.09	100.0

THE METHODS USED BY THE O.E.E.C.
TO ESTIMATE HOLIDAY VARIATION

14. With the method used by the O.E.E.C. for day-of-the-week variation, it is not possible to attempt to estimate holiday variation simultaneously, as for example in the methods used by Eisenpress and the Netherlands Central Bureau of Statistics. In practice, the two kinds of variation can only be estimated simultaneously if it is possible to assimilate public holidays with Sundays. For some series this may be appropriate, but in many cases it is not. Thus for example, the influence of a holiday may vary according to whether it falls in the middle of the week, or near the beginning or end. (There is evidence that this is the case for industrial production at Christmas in several European countries.) And for

3 Comparison of "day of week" factors
calculated on basis of external and internal evidence
Canada, retail trade, total excluding motor vehicle dealers



(1) Based on sample survey of daily retail sales conducted by the Dominion Bureau of Statistics.

(2) Based on internal evidence derived from the monthly series of retail sales.

some series, such as retail sales, the date or day of a public holiday may influence the figures in a quite unpredictable way.

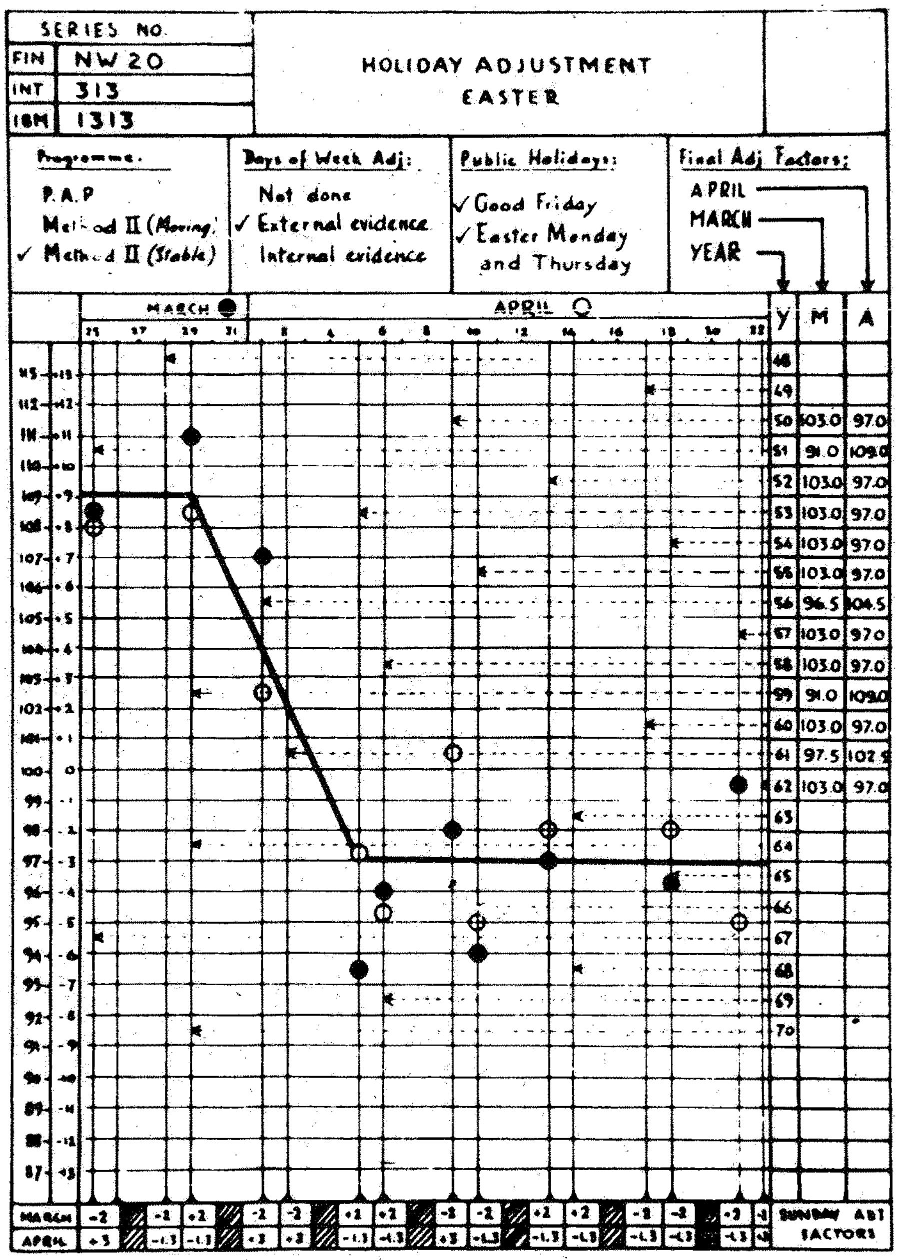
15. For the present, therefore, the O.E.E.C. is using an iterative and largely *ad hoc* approach. For the first run, the series are prior-adjusted crudely for both day-of-the-week and holiday variation on the basis of external evidence. Then, after a rough check on the holiday adjustments, the day-of-the-week adjustment factors are checked using the procedure described above. Then after a revision of the day-of-the-week factors where necessary, a rather more refined check is made for holiday variation using specially designed work sheets.

16. Diagram 4 shows a reproduction of the work-sheet used for Easter variation as filled in for the Norwegian index of industrial production. The irregular ratios for March and April, defined with respect to a stable seasonal (vertical scale), is entered against the date of Easter (horizontal scale). It can be shown that Easter should have an almost symmetrical but opposite effect on March and April, i.e. if March production is low because Easter is in March, the April irregular will be about as positive as March is negative. For this reason, the March irregular is inverted and the reciprocal plotted, thus facilitating the estimation of the appropriate adjustment factors. As can be seen from Diagram 4, this series does not seem to have been adjusted for Easter variation. Adjustment factors have thus been estimated, as shown by the solid line, and entered in the two columns on the right-hand side.* The fitting of this line requires some care, since sharp movements are likely as Easter moves just in March to just in April, and it is generally not appropriate to fit smooth curves. It may be noted that for some foreign trade series, however, the influence of Easter seems to be progressive, the amount by which April is depressed and March raised steadily increasing as Easter moves further into April.

17. Similar methods are used to estimate the influence of other moving date holidays (Whitsun) and changing day-of-the-week holidays (Christmas). With the latter kind of holiday, there are often not enough observations to justify basing the adjustment solely, or even mainly, on internal evidence. Over the ten years from 1950 to 1959, for example, Christmas fell only twice on a Tuesday, Thursday or Friday, and only once on Monday, Wednesday, Saturday and Sunday. Nevertheless, it is useful to make a quick check on the validity of external-based prior adjustments.

* If the series has not been prior-adjusted for day-of-the-week variation, the two lines at the bottom indicate rough adjustments which can be made in cases where it is reasonable to assume that activity on Sunday is zero.

4 Example of easter adjustment worksheet (Norway industrial production)



○ Irregular ratios for the month of April

● Reciprocal of the irregular ratios for the month of March.

▬ Final adjustment factors

Box No. 8

March 1973

An Easter Adjustment Procedure

Pierre Cholette
Research and Development Unit
General Time Series Section
Statistics Canada

Introduction

Several time series are affected by the arrival of Easter. The purpose of this document is to determine which series need to be adjusted for Easter, and how this Easter effect should be computed.

The Easter Adjustment Procedure proposed here is, with one modification, that contained in Mr. J. Cudmore's pamphlet (for internal use) titled "A New Easter Adjustment Method".* We do not criticize this method in terms of mathematical theory; we set it out as it stands, reserving the right to propose a new method when we can. We assume that Mr. Cudmore's procedure has had some practical success, since it is used by Statistics Canada. It also has the advantage of using the same approach as X-11 ("ratio-to-average").

We present first a descriptive analysis of the establishment of Easter factors and their use. A corresponding mathematical statement will then be found in Appendix A. Arithmetical computations are to be found in Appendix B in tabular form, to be filled in by individuals responsible for making the computations.

*Not supplied to translator.

I. Series affected by the arrival of Easter

We know that Easter affects certain retail trade series, "Women's Clothing" for example. The number of marriages and chocolate sales may also be expected to respond to the arrival of Easter. In the absence of any direct experimentation we do not know exactly which series need to be adjusted for Easter. It will be necessary for officers in units producing series to use the method described here to examine how their series respond to Easter. In most cases, however, it will not be difficult for them to determine *a priori* whether their series require such adjustment, on the basis of economic theory and common sense.

In testing whether an Easter effect is present in a series, there are roughly speaking three stages:

(1) adjusting the series for the number of working days, using X-11;

(2) computing the Easter factors and their level of statistical significance from the series corrected for working days produced by the preceding stage (Table C19 of X-11);

(3) making a final X-11 run using as input the raw original series and the Easter factors computed in the preceding stage. This final stage will only be undertaken if the Easter factors have proven to be statistically significant in (2).

II. Stages in the establishment of Easter factors

(1) Correcting the series for the number of working days

We believe *a priori* that series affected by Easter will probably be affected by the number of working days. An eligible series will therefore have to be adjusted for the number of working days. We think the adjustment should be made first for working days, then for Easter, as in any series there are more degrees of freedom to measure the effect of working days than to measure the Easter effect.

Correction for working days will be by a preliminary X-11 run using the "Prior Daily Weights" option, and/or the "Computed Daily Weights" option. The input needed in computing the Easter factors will be found in Table C19, "Original Series adjusted by Trading Day adjustment factors".

(2) Computing Easter factors as such

Succeeding stages to be followed in computing the Easter factors are as follows:

(a) Ratios for March to the totals for March and April are first computed for the whole series (adjusted for the number of working days). The numbers obtained will be in the vicinity of 0.50.

(b) The next stage consists of transforming the date of Easter in terms of natural integers. Since Easter can fall between March 25 and April 22, we suggest assigning the number 1 to March 25, 2 to March 26, and so on. This produces a mathematically useable variable which we designate the "Easter date variable".

(c) We now proceed to linear regression (and it is here that we differ from Mr. Cudmore's method, which suggests visual adjustment) of the March ratios to the totals over the Easter date variable X. This stage attempts to weed out irregularities in the ratios and generate the "expected" ratios for March to the totals of

March and April, by the following expression:

$$\hat{R}_{iM} = a + b\ (X_i - \bar{X}),$$

in which the $\hat{R}_i$s are the required expected ratios for the whole series; X_i the values of the Easter date variable; $\bar{X}$ the average for these dates; and a and b the coefficients of regression.[1]

The expected ratios for April are obtained by additive complementarity β relative to 1.0.

If the value of t[2] is inadequate in this regression, then there is no statistically significant Easter effect for this series. Computation is discontinued in such cases.

(d) The next stage consists of eliminating average seasonal and cyclo-tendential effects from these expected ratios. This is done by dividing the expected ratios for March by their average (= a), and the expected ratios for April by their

---

1 $a = \Sigma R_i / n$

$b = (\Sigma R_i\ (X_i - \bar{X})) / \Sigma (X_i - \bar{X})^2$

2 $t = b / \sqrt{\Sigma (R_{iM} - \hat{R}_{iM})^2 / \Sigma (X_i - \bar{X})^2}$

average (= (1. -a)). From this we get Easter factors which are not forced to the sum of 2.0.

(e) Finally, the products of the preceding stage must be forced to the sum of 2.0. This is done by dividing the corrected expected ratios by their current average. From this we get the Easter factors, which are centred on 1.0.

III. Use of Easter factors

The Easter factors are then used in the final adjustment of the series. They are stored in the prior adjustment factor, P. The factors for number of working days will be recomputed from the computed daily weights in the second X-11 run. The factors for working days will thus have been repeated twice more than if a single X-11 run had taken place.

IV. Generating Easter factors for the year following the survey period

We can now generate Easter factors for the next year. This is done in several stages, comparable to those for establishing the Easter factors for the ~~abbreviation~~ period.

(1) A natural number must first be assigned to the date of Easter according to the same rule of correspondence established in II(b). We thus get X_{n+1}.

(2) We generate the expected ratios for March to the totals for March and April by substituting the value of the Easter date established in (1) in the equations estimated in II(c): $\hat{R}_{(n+1)M} = a \div b\ (X_{n+1} - \bar{X})$. The expected ratio for April is obtained by additive complementarity relative to 1.0 with that for March.

(3) These expected ratios must then be cleared of the average seasonal and cyclo-tendential effect by dividing them by their respective averages, a and (1-a) of II(c).

We designate the result the "corrected expected ratios".

(4) Finally, these results must be forced to the sum of 2.0 by dividing by their current average, thus giving the Easter factors for the next year: $P_{(n+1)M}$, $P_{(n+1)A}$.

APPENDIX A. Mathematical statement

I. Setting Easter factors for the survey period

(100) Let $0 = C\ S\ I\ D\ P$, the original series, and

(101) $0' = C\ S\ I\ P$, the series adjusted for working days.

$0'$ is obtained by a preliminary run of X-11 (Table C19).

Let R_{iM} be the ratios for March to the March and April totals:

(102) $$R_{iM} = 0'_{iM} / (0'_{iM} + 0'_{iA}),\ (i = 1,n)$$

where the index i denotes the year; the index M the month of March; the index A, April; and n is the number of years in the survey period.

Let $\hat{R}_{iM}$ be the expected ratios for March to the March and April totals:

(103) $\hat{R}_{iM} = a + b\ (X_i - \bar{X})$, where X_i is a linear transformation of the date of Easter into natural numbers (positive integers);

$\bar{X}$ the averages of the dates transformed; and a and b are the regression coefficients of the model associated with this equation.

(104) $$R_{iM} = A + B\ (X_i - \bar{X}) + E_i.$$

Let $\hat{R}_{iA}$ be the expected ratios for April to the March and April totals. These ratios are given by the identity.

(105) $$\hat{R}_{iA} = 1 - \hat{R}_{iM}.$$

Let $\bar{R}_M$ be the average of the expected ratios for March to the totals:

(106) $$\bar{R}_M = \sum_{i1}^{n} \hat{R}_{iM} = a.$$

(107) Similarly $\bar{R}_A$: $\bar{R}_A = 1 - \bar{R}_M = 1 - a$

Let $\hat{R}_{iM}^{c}$ and $\hat{R}_{iA}^{c}$ be the expected ratios for March and April to the totals corrected for average seasonal and cyclo-tendential effects:

(108) $$\hat{R}_{iM}^{c} = \hat{R}_{iM} \ /\ (\bar{R}_M = a),\ \hat{R}_{iA}^{c} = \hat{R}_{iA}^{c} /\ (\bar{R}_A = 1 - a)$$

The factors P_{iM} and P_{iA}, Easter factors forced to the sum of 2.0, are:

(109) $$P_{iM} = \hat{R}_{iM}^{c} /\ (\hat{R}_{iM}^{c} + \hat{R}_{iM}^{c})\ /2$$

(110) $$P_{iA} = 2.0 - P_{iM}$$

II. Using the results

The final adjusted series is obtained by a final X-11 run, using the Easter factors and the raw series as input.

III. Determining Easter factors for the year following the survey period

(1) Let $\hat{R}_{(n+1)M}$ and $\hat{R}_{(n+1)A}$ be the expected ratios for the year following the survey period, defined by:

$$(112) \qquad \hat{R}_{(n+1)M} = a + b\,(X_{n+1} - \bar{X}),$$

$$(113) \qquad \hat{R}_{(n+1)A} = 1 - \hat{R}_{(n+1)M},$$

where a, b, and $\bar{X}$ are the same values previously computed; and where X_{n+1} is the transformation of the date of Easter into natural numbers applying the same rule used above.

(2) Let $\hat{R}^c_{(n+1)M}$ and $\hat{R}^c_{(n+1)A}$ be the corrected expected ratios defined by:

$$(114) \qquad \hat{R}^c_{(n+1)M} = \hat{R}_{(n+1)M}/a \text{ and}$$

$$(115) \qquad \hat{R}^c_{(n+1)A} = \hat{R}_{(n+1)A}/(1-a)$$

(3) The Easter factors for the following year will be:

$$P_{(n+1)M} = \hat{R}^c_{(n+1)M} \Big/ (\hat{R}^c_{(n+1)M} + R^c_{(n+1)A})/2$$

$$P_{(n+1)A} = 2.0 - P_{(n+1)M}.$$

CENSUS TRADING-DAY ADJUSTMENT METHOD

Allan Young
Economic Research and Analysis Division
Bureau of the Census

This paper is reprinted separately as an aid to the understanding of business indicators and the methods and concepts followed in the Business Cycle Developments report. It was printed initially in the May 1964 issue of that report.

U. S. DEPARTMENT OF COMMERCE, Luther H. Hodges, Secretary
BUREAU OF THE CENSUS, Richard M. Scammon, Director

Technical Papers and Background Materials

To aid users of Business Cycle Developments, technical papers dealing with the statistical adjustments and series used in BCD are included in this report from time to time. The following papers have been included as part of this program:

No. 1.—Summary Description of the X-9 and X-10 Versions of the Census Method II Seasonal Adjustment Program (published as appendix E in the September 1963 issue). A new version of this program is scheduled to be released in the fall. Announcement will be made at that time.

No. 2.—Business Cycle Indicators—The Known and the Unknown (published as appendix H in the September 1963 issue). This paper explains what is known about business cycle indicators, the problems of using them, and the research needed to improve their usefulness. It was presented at the 34th session of the International Statistical Institute in Ottawa, Canada, on August 24, 1963.

No. 3.—Census Trading-Day Adjustment Method (published in May 1964 issue).

A limited number of copies of these articles are available, free of charge, from the Chief Economic Statistician, Bureau of the Census, Washington, D.C., 20233.

Census Trading-Day Adjustment Method

Allan Young, Economic Research and Analysis Division, Bureau of the Census

INTRODUCTION

An important source of month-to-month variation in many monthly economic time series is trading-day variation.[1] In activities such as production, sales, and shipments in domestic and foreign trade, the monthly rate of activity is related to the number of working or trading days in the month. A familiar example is retail sales where more sales are made on Fridays and Saturdays than on other days; therefore, months that contain five Fridays and/or Saturdays have higher sales than months with four.

By making a trading-day adjustment, the month-to-month variation in seasonally adjusted data can be reduced and the trend-cycle component revealed more clearly.[2]

The importance of trading-day variation relative to other types of month-to-month variation is shown in table 1 for seven series adjusted for trading-day variation by the Bureau of the Census. In each series, trading-day variation is several times as large as the monthly variation in the trend-cycle component. For imports, trading-day variation is considerably more important than seasonal variation, and for wholesale sales, those two components are almost equal. In both cases, about half the total variation is accounted for by trading-day variation. Trading-day variation is of prime concern when attempting to assess the underlying cyclical movement over short spans of 1 or 2 months. Over longer spans, trading-day variation is of less importance since it frequently reverses direction over time and does not cumulate as do the seasonal and cyclical movements.

Techniques for making trading-day adjustments in the past have often relied heavily upon observation or a priori information concerning the daily activity, referred to as external evidence. The customary practice has been to establish a rate of activity for each day of the week. Such a practice is usually limited by the available information and the cost of obtaining it. Usually, some fairly simple pattern of daily activity is assumed after examining the available information which often consists only of the weekly schedule of hours of work. For example, in a manufacturing activity the assumption might be a 5-day week with the same rate of activity for each weekday and zero for Saturday and Sunday.

The variation actually found in monthly economic series usually does not arise solely from one simple pattern of daily activity, but from a mixture of many factors related to the numbers of each day of the week in the month (i.e., to the calendar composition of the month). These sources of variation include the following: Different rates of daily activity in various processes, some of which are in continuous operation seven days a week; practices

[1] The term "trading-day variation" can be considered interchangeable with the terms "working-day variation" and "calendar variation." Earlier studies that have considered estimating trading-day variation from the monthly time series are discussed by Marris in reference 2 (see end of paper). The Eisenpress method (reference 1) is compared with the Census method in reference 5.

[2] A familiarity with seasonal-adjustment techniques is assumed. At times reference is made to specific measures provided by the Census Method II ratio-to-moving-average method of seasonal adjustment. Details concerning Census methods can be found in references 3 and 4 (see end of paper).

NOTE: Several people have made substantial contributions to the recent development of trading-day adjustment techniques. The work at the Bureau of the Census was carried on under the supervision and encouragement of Julius Shiskin. John Musgrave developed much of the mathematical formulation and made many other valuable contributions. Gerald Donahoe advised and assisted in the application of the method to Census series. Morton Somer, Norman Bakka, Richard Bartlett, and Barry Beckman provided programing and other assistance. Marie Wann and Geraldine Censky provided editorial review.

Much of the work draws upon Stephen Marris' earlier work at the Organization for Economic Co-operation and Development. James Nettles and David Staiger of the Federal Reserve Board made helpful suggestions.

Table 1.--PERCENT DISTRIBUTION OF COMPONENTS OF MONTH-TO-MONTH VARIATION FOR INDICATED TIME PERIODS AND SELECTED BUSINESS ACTIVITY

(Percent)

Business activity	Time period	Components of variation				
		Total	Trading day	Seasonal	Irregular	Trend-cycle
Retail sales..	1953-62	100.0	7.0	[1]91.8	0.8	0.4
Wholesale sales.......	1956-62	100.0	47.2	49.4	2.5	0.9
Mfrs.' shipments.......	1953-62	100.0	20.1	71.9	5.6	2.5
Mfrs.' new orders......	1953-62	100.0	19.3	62.6	14.3	3.8
U.S. exports..	1953-63	100.0	18.3	63.1	15.8	2.8
U.S. imports..	1953-63	100.0	50.3	38.2	10.0	1.5
Building permits........	1954-62	100.0	21.2	69.1	7.8	1.9

[1] Includes a slight contribution from variation in sales of selected kinds of business in March and April because of shifting date of Easter.

Reference 5, appendix C (see end of paper for citation) will show the derivation of these measures from Census Method II summary measures of monthly change, $\overline{O}$, $\overline{S}$, $\overline{I}$, $\overline{C}$.

concerning overtime work; contracts and schedules specifying a fixed amount of activity each month regardless of the calendar composition; and bookkeeping practices that modify actual variations. In many instances the effect of these factors is to reduce the variation that might be expected after cursory examination of external evidence of the daily rate of activity. Although the variations arising from this mixture of factors often cannot be estimated from external evidence, their net effect can be estimated from the monthly data using techniques that have been developed and tested at the Bureau of the Census and elsewhere.[3] Such techniques, in general, yield better trading-day and seasonal adjustments and are preferable to earlier techniques.

A trading-day routine, described briefly below, is being added to the Census Method II seasonal adjustment program. This routine tests for the existence of significant trading-day variation in the monthly data and adjusts the original observations when trading-day variation is present. Then seasonal factors and other Census Method II measures are developed. More complete information will be available in forthcoming specifications for new seasonal programs.

The trading-day routine provides estimates of seven daily weights. In the absence of a complex of various factors, these estimated daily weights correspond to the actual daily rates of activity. When there are various factors at work, the estimates cannot be interpreted as representing actual daily rates of activity but only as statistical weights that represent the net effect of several variables.

[3] Notably at the Organization for Economic Cooperation and Development. See reference 2 at end of paper.

DEFINITION OF TRADING-DAY VARIATION

In adjusting for trading-day variation, two types of variation in the calendar are often considered. The first is differences in the length of the month; i.e., differences between 28-, 29-, 30- and 31-day months. On average, the longer months tend to have a higher volume of activity than the shorter months. Variations in the volume of activity arising from months of different length, however, cannot be statistically separated from seasonal influences that also cause differences between months. Length-of-month variation, therefore, is defined and estimated as part of the seasonal component.

The second type of variation is referred to as calendar composition variation. The number of Mondays, Tuesdays, etc., in a given month varies from year to year. For example, a 31-day month in 1 year may contain five Fridays, Saturdays, and Sundays and four of each of the other four days of the week, while in another year it may contain five Mondays, Tuesdays, and Wednesdays and four of the other days. If some days are more important in the economic activity than others, this variation gives rise to variation in the monthly volume of activity. Such variation, which is not included in the definition or estimation of the seasonal variation, can be estimated statistically by relating the monthly economic series to the calendar. Our definition of trading-day variation is, therefore, that it is the variation in the monthly series related to the calendar composition variation.

METHOD OF ESTIMATING TRADING-DAY VARIATION

If no allowance is made for trading-day variation prior to seasonal adjustment, the trading-day variation is left as a residual in the irregular component. Therefore, the sequence of steps in estimating trading-day variation in the Census Method II seasonal adjustment program is (1) to separate the trading-day and other irregular variations from the seasonal and trend-cycle, seasonally adjust the original series; (2) from the combined irregular and trading-day variations, estimate the trading-day variation in terms of seven daily weights (see below); (3) from the seven daily weights, derive monthly trading-day adjustment factors and adjust the original series for trading-day variation; and (4) using the trading-day adjusted data, make a second seasonal adjustment.[4]

The seven daily weights are estimated by regressing the irregular component upon seven independent variables, representing the number of times each day of the week occurs in a particular month, as follows:

$$I_i = \frac{X_{1i} B_1 + X_{2i} B_2 + \dots + X_{7i} B_7 + E_i}{N_i},$$

[4] There are several possible variations upon this sequence and upon the regression form used to estimate the trading-day variation. (See reference 5 at end of paper.) For example, it is possible to incorporate the trading-day estimation with a simultaneous rather than sequential Method II solution for the seasonal and trend-cycle components.

where

I_i is the estimate of the irregular component for month i that includes both the "true" irregular variation and the trading-day variations. The mean of I_i is 1.

N_i is 31, 30, or 28.25 depending upon whether month i is a 31- or 30-day month or February;

X_{ji} is the number of times the day-of-the-week j occurs in month i;

B_j's are seven daily weights, totaling 7;

E_i is the "true" irregular for month i.

The method yields estimates, b_j, of seven daily weights, B_j, and estimates of the standard errors of the b_j. Thus, one may perform a standard t-test to determine whether a weight is significantly different from any specified value and an F-test to determine the significance of the regression (i.e., the existence of significant trading-day variation in the irregular).

INTERPRETATION OF DAILY WEIGHTS

The regression routine is designed to provide seven daily weights that total "7". A 5-day week where the weekdays are of equal importance and Saturday and Sunday are "0" is expressed as Mon., ..., Fri. = 1.4; Sat., Sun. = 0.0. A series with no trading-day variation would yield equal weights for all seven days, Mon., ..., Sun. = 1.0. Alternative formulations, such as expressing each day of the week as a percent of the total for all days of the week can also be used.

There are two general types of variation which modify the actual daily rates and create a complex of variations. The first type arises from economic transactions that are independent of calendar composition.[5]

When a portion of a series represents activity independent of calendar composition, the estimated daily weights can be considered as composed of two parts, one part consisting of equal weights for each day and the other having differential weights. Consider the following hypothetical set of daily weights:

	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.	Total of weights	Percent
(1)	1.00	1.10	1.20	1.30	1.40	0.50	0.50	7.00	100

These weights (1) can be separated into two parts:

(1a)	.50	.60	.70	.80	.90	0.0	0.0	3.5	50
(1b)	.50	.50	.50	.50	.50	.50	.50	3.5	50

[5]Some transactions that are independent of calendar composition are also independent of the length of the month. For example, monthly rental payments are independent of both calendar composition and the length of the month while the consumption of heating fuel is more or less independent of calendar composition but not length of the month.

Part (1a) reflects the part of the series that varies with the number of different type days and accounts for 50 percent of the series, and (1b) is independent of calendar composition and accounts for 50 percent of the total series.[6]

The second general type of variation that modified the actual daily rates are the variations induced by bookkeeping practices. Suppose a firm that operates on a 6-day week (Monday, ..., Saturday = 1.17, Sunday = 0.0) follows the practice of closing its books for the month on Friday whenever the month ends on Saturday (which occurs only about twice a year) and includes the Saturday activity in the following month. This practice would apply to 31-day months beginning on Thursday and 30-day months beginning on Friday. For such months, reported sales would be decreased by almost 4 percent. Conversely, reported sales are increased by almost 4 percent for the following months which always begin on Sunday. Reporting sales in this manner results in a monthly series which yields a weight pattern of Monday, ..., Friday = 1.17, Saturday = 0.0 and Sunday = 1.17 rather than the actual daily rate which had a zero Sunday.[7]

Another bookkeeping practice, followed by some firms, that affects trading-day variation is the plan known as the 4-4-5 plan where the first and second months of each quarter always contain exactly four weeks and the third month five weeks. This practice eliminates trading-day variation since each period has exactly the same number of each type of day of the week as the corresponding periods in earlier years. When some activity is reported under such plans and other is reported on a strict calendar month basis, the combined variation can be represented with the above example where the (1b) weights represent the portion of the series that contains no trading-day variation.

There are three types of possible variation related to the calendar which are not included in the formulation of the daily weights. They are the variation related to holidays, changes in the trading-day pattern over time and changes in the trading-day pattern from one season to another. Experience at Census suggests that these three possible varia-

[6]Although it is useful for illustration to separate the hypothetical set of daily weights into two parts, (1a) and (1b), such a technique has limited usefulness. From information contained in monthly data no reliable conclusions can be made concerning the actual daily rates of activity. For example, the activity represented by the (1b) Sunday weight may not actually take place on Sunday, but be distributed over the other days. What does correspond is the variation in the monthly series and the trading-day adjustment factors derived from the daily weights.

[7]The effect upon the Sunday weight may appear extreme since the Saturday is only 1/30 or 1/31 of the month. However, 4 full weeks, 28 days, are found in all months, and only 2 or 3 days are unique to each month. The shift of one day's activity is 1/2 or 1/3 of the monthly variation attributable to trading-day differences.

Table 2.--COMP[illegible]ON OF ALTERNATIVE TRADING-DAY ADJUSTMENTS FOR SE[illegible]CTED BUSINESS ACTIVITY

Item	Daily weights								$\bar{I}$ Avg. monthly change, without regard to sign, in irregular component		Rank (lowest $\bar{I}$ = "1")	
	Total	Mon.	Tues.	Wed.	Thur.	Fri.	Sat.	Sun.	Historical	Current	Historical	Current
									1953-61	1962-63	1953-61	1962-63
RETAIL SALES												
Furniture stores:												
Regression, 1953-61...	7.00	1.24	0.91	0.84	1.23	1.10	1.11	0.57	1.25	1.61	1	1
Regression, 1957-61...	7.00	1.23	0.89	0.97	1.14	1.15	1.07	0.55	1.27	1.66	2	3
1962 daily sales......	7.00	1.30	1.04	0.95	1.13	1.18	1.39	0.01	1.45	1.62	3	2
No adjustment[1]........	...	...	...	...	...	...	...	...	1.75	1.86	4	4
Lumber and building materials dealers:												
Regression, 1953-61...	7.00	1.13	1.11	1.16	1.22	1.15	0.67	0.56	1.69	0.87	1	1
Regression, 1957-61...	7.00	1.00	0.99	1.39	1.11	1.24	0.64	0.63	1.77	1.05	2	2
1962 daily sales......	7.00	1.29	1.26	1.16	1.22	1.19	0.87	0.01	1.81	1.43	3	3
No adjustment[1]........	...	...	...	...	...	...	...	...	2.86	2.30	4	4
Hardware stores:												
Regression, 1953-61...	7.00	1.08	0.96	0.78	1.29	1.23	1.15	0.51	1.58	1.92	1	1
Regression, 1957-61...	7.00	0.96	0.84	0.96	1.19	1.44	0.90	0.71	1.62	2.22	2	3
1962 daily sales......	7.00	1.11	1.04	1.10	1.05	1.22	1.48	0.00	2.22	2.18	3	2
No adjustment[1]........	...	...	...	...	...	...	...	...	2.79	2.67	4	4
									1954-61	1962-63	1954-61	1962-63
BUILDING PERMITS												
Regression, 1954-61.....	7.00	1.20	1.11	1.48	1.25	1.60	-0.32	0.69	3.17	2.92	2	1
A priori................	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	3.13	3.72	1	2
No adjustment[1]..........	...	...	...	...	...	...	...	...	5.19	5.23	3	3
									1953-61	1962-63	1953-61	1962-63
MANUFACTURERS' SHIPMENTS												
Tobacco:												
Regression, 1953-61...	7.00	1.89	0.96	1.42	1.18	1.24	0.00	0.31	1.54	1.58	1	1
A priori..............	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	2.68	2.31	2	2
A priori..............	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	2.92	3.13	3	3
No adjustment[1]........	...	...	...	...	...	...	...	...	4.35	4.96	4	4
Kitchen articles and pottery:												
Regression, 1953-61...	7.00	0.39	1.30	1.43	1.07	1.01	0.91	0.89	4.96	5.56	1	3
A priori..............	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	5.78	5.21	3	2
A priori..............	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	5.75	5.57	2	4
No adjustment[1]........	...	...	...	...	...	...	...	...	6.23	4.71	4	1
Radio and TV:												
Regression, 1953-61...	7.00	0.46	0.93	1.41	0.94	1.15	0.89	1.22	5.00	3.96	1	2
A priori..............	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	6.79	4.56	4	4
A priori..............	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	5.77	3.61	3	1
No adjustment[1]........	...	...	...	...	...	...	...	...	5.40	4.20	2	3
Engines and turbines:												
Regression, 1953-61...	7.00	1.19	0.36	1.00	2.41	0.34	0.73	0.97	8.50	23.61	1	1
A priori..............	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	9.35	25.78	4	4
A priori..............	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	9.20	23.65	3	2
No adjustment[1]........	...	...	...	...	...	...	...	...	9.06	23.79	2	3

[1]Shown without weights. An alternative would be to show a weight of 1.0 for each day of the week.

tions are usually insignificant and rarely impair the usefulness of the trading-day adjustment technique.[8]

TESTS OF ALTERNATIVE TRADING-DAY ADJUSTMENTS

The regression method has been tested with real economic series and artificial series that contain a known trading-day pattern. Table 2 presents the results of a few of the tests upon economic series.

The criterion used to evaluate alternative trading-day adjustments of economic series is that the best method is the one resulting in the smallest month-to-month change, without regard to sign, in the irregular or unexplained variation (referred to as $\bar{I}$). It is not sufficient, however, to make this determination only for the "historical" period from which the estimates were made. A "historical" comparison is biased if the estimates of one or the other method, by fitting the data too closely, explain not only the trading-day variation, but part of the irregular variation. A more sufficient test is to apply the estimates made for the "historical" period to the "current period," where methods that are too sensitive to the historical irregular fluctuations and those which inadequately allow for the characteristics of trading-day variation will both yield large fluctuations that will be included in the computed $\bar{I}$.

Evaluation, therefore, consists of the following steps: (1) estimate trading-day variation with each method from an historical period; (2) make the trading-day adjustment to the historical and current data with the historical estimate; (3) obtain an irregular component by seasonally adjusting the combined historical and current data and compute $\bar{I}$ for each period; and (4) compare the $\bar{I}$'s from the various methods giving particular attention to the current period.

Results of the following tests are shown in table 2.

Retail sales.—For the period 1953-63, sales of three retail kinds of business were adjusted for trading-day variation by regression estimates computed from the period 1953-61 and also 1957-61. These regression adjustments are compared with trading-day adjustments based upon average rates of sales on each day of the week computed from unpublished daily retail sales for 1962 that are available at the Census Bureau. They also are compared with series not adjusted for trading days.

The regression estimates for 1953-61 yield the smallest $\bar{I}$'s, even for the current period of 1962-63 where we might expect the results to be biased in favor of the 1962 daily sales rates. The 1957-61 estimates are next best, followed by the daily sales rates and no adjustment.

Even though the unpublished estimates of the 1962 daily sales are not up to the usual Census publication standards,[9] the seven average daily rates are based on more evidence than is often available for an external adjustment and they appear to be reasonably close to what our experience would suggest as the daily sales pattern (see table 2). This comparison strongly suggests that external observation of the daily pattern of activity does not provide an adequate basis for a trading-day adjustment.

Building permits.—For U.S. building permits, regression estimates computed from the period 1954-61 are compared with a 5-day week which might be selected a priori and with the series not adjusted for trading days. The regression estimates and the 5-day-week adjustment yield approximately the same results for the historical period and for the current period of 1962-63. They substantially reduce the irregular fluctuations found in the series not adjusted for trading days.

Manufacturers' shipments.—For manufacturers' shipments in four selected industries, regression estimates computed from the period 1953-61 are compared with two sets of weights that might be selected a priori: (1) Weights for a 5-day week, and (2) weights where Saturday and Sunday receive partial weights. The regression estimates are also compared with the series not adjusted for trading days. The a priori weights where Saturday and Sunday receive partial weights have been found to be appropriate for the aggregate shipment series and they might be expected to be appropriate for each component.

The regression estimates yield the smallest $\bar{I}$ in the historical period for each of the four series. In the current period 1962-63, they are best for two series while the a priori weights where Saturday and Sunday receive partial weight are best for one series and no adjustment for trading days is best for one series.

These four selected manufacturer's shipments series contain larger irregular variations than do the retail sales data. The differences between alternative adjustments are small relative to the magnitude of $\bar{I}$ and the results are less conclusive. This test illustrates the fact that for highly irregular series the possible gain in trading-day adjustment of a series is small.

Reference 5 (see end of paper) will give a further discussion of the value of adjusting highly irregular series for trading-day variation.

These tests (and other tests on real and artificial series shown in reference 5) suggest that the regression method performs quite well in comparison to other alternatives.

---

[8] The only Census series in which significant holiday variation has been found are retail sales where for sales of some kinds of business, the date of Easter affects March and April and the dates of Labor Day and Thanksgiving slightly affect August-September and November-December. Series based upon a survey covering 1 week of the month, although not containing trading-day variation may contain holiday variation. For example, the series on the average workweek can be affected by holidays that fall in the survey week. A change, over time, in the trading-day pattern of Canadian retail sales is discussed in reference 2.

[9] The 1962 daily sales were voluntarily reported by about a fourth of the 1,600 respondents in the weekly survey of stores with 10 or less outlets. The number of reports by kind of business is quite small, 27 for furniture stores, 36 for lumber and building materials dealers, and 17 for hardware stores and the daily data, therefore, contains some inaccuracies.

REFERENCES

1. Eisenpress, Harry, "Regression Techniques Applied to Seasonal Corrections and Adjustments for Calendar Shifts," Journal of the American Statistical Association, vol. 51, No. 276, December 1956, pp. 615-620.

2. Marris, Stephen N., "The Measurement of Calendar Variation," Seasonal Adjustment on Electronic Computers, Organization for Economic Cooperation and Development, 1960, pp. 345-360.

3. Shiskin, Julius, Electronic Computers and Business Indicators, Occasional Paper 57, National Bureau of Economic Research, Inc., 1957, (reprinted from Journal of Business, October 1957).

4. Specifications for the X-9 version of Census Method II seasonal adjustment program are available from the Chief Economic Statistician, Bureau of the Census, Washington, D.C. 20244.

5. Young, Allan, "The Estimation of Trading-Day Variation in Monthly Economic Time Series," (in preparation).

Technical Paper No. 12

ESTIMATING TRADING-DAY VARIATION IN MONTHLY ECONOMIC TIME SERIES

U S. DEPARTMENT OF COMMERCE
Bureau of the Census

BUREAU OF THE CENSUS TECHNICAL PAPER NO. 12

ESTIMATING TRADING-DAY VARIATION IN MONTHLY ECONOMIC TIME SERIES

by
Allan Young
Economic Research and Analysis Division

U.S. DEPARTMENT OF COMMERCE
John T. Connor, Secretary

BUREAU OF THE CENSUS
A. Ross Eckler, Acting Director

BUREAU OF THE CENSUS
A. Ross Eckler, Acting Director

MORRIS H. HANSEN, Assistant Director for Research and Development

Economic Research and Analysis Division
JULIUS SHISKIN, Chief

Library of Congress Card No. A65-7383

SUGGESTED CITATION

U.S. Bureau of the Census. *Estimating Trading-Day Variation in Monthly Economic Time Series. Technical Paper No. 12.*
U.S. Government Printing Office, Washington, D.C., 1965

For sale by the Superintendent of Documents, Government Printing Office
Washington, D.C., 20402 - Price 30 cents

Preface

Many types of monthly economic time series contain variations which are related to the number of times a particular day or days of the week occur in the calendar month. These variations are usually referred to as trading-day variations. Recently, substantial use and development of a technique to estimate and remove these variations have occurred at the Bureau of the Census. This new technique estimates trading-day variation by relating the monthly volume of activity to the number of times each day of the week occurs in the month. By and large, this approach yields better trading-day and seasonal adjustments than the often used techniques that rely upon independent, external evidence of the percent of the week's activity that occurs on each day of the week.

This paper attempts to examine fully the subject of trading-day variation. Briefer treatments of the subject are found in "Census Trading-Day Adjustment Method," Business Cycle Developments, May 1964, and in specifications that will be made available for a new version of the Census Method II seasonal adjustment program, designated the X-11 variant, which includes a routine to estimate trading-day variation.

Several people have made substantial contributions to the recent development of trading-day adjustment techniques. The work at the Bureau of the Census was carried on under the supervision and encouragement of Julius Shiskin. John Musgrave made many valuable contributions and prepared the mathematical presentation in appendixes A and B. Gerald Donahoe assisted in much of the development and application of the method. Morton Somer prepared the computer programs. Harry Rosenblatt and Edward Melnick made helpful suggestions. Norman Bakka, Richard Bartlett, and Barry Beckman provided substantial assistance. Geraldine Censky and Marie Wann provided editorial review.

Much of the work draws upon Stephen Marris' earlier work at the Organization for Economic Cooperation and Development. James Nettles and David Staiger of the Federal Reserve Board made helpful suggestions.

Allan Young
Bureau of the Census
April 1965

Contents

ESTIMATING TRADING-DAY VARIATION IN MONTHLY ECONOMIC TIME SERIES

I. INTRODUCTION

An important source of the month-to-month variation in many monthly economic time series is trading-day variation.[1] In activities such as production, sales, and shipments in domestic and foreign trade, the monthly rate of activity is related to the number of working or trading days in the month. A familiar example is retail sales, where more sales are made on Fridays and Saturdays than on other days and higher sales are made in months containing five Fridays and/or Saturdays than those with four.

Trading-day variation is systematic and its characteristics relatively stable over several years. Therefore, it is possible to estimate trading-day variation from information contained in the monthly data.

The justification for a trading-day adjustment is that it reduces the month-to-month variation in seasonally adjusted data so that the trend-cycle component is revealed more clearly.[2] This reduction in month-to-month variation is illustrated in the chart, which shows seasonally adjusted U.S. imports, nonagricultural job placements, and sales of department stores with and without an adjustment for trading-day variation.

The importance of trading-day variation relative to other types of month-to-month fluctuations is shown in table 1 for seven series adjusted by the Bureau of the Census. In each series, trading-day variation is several times as large as the monthly variation in the trend-cycle component. For imports, trading-day variation is considerably more important than seasonal variation, and for wholesale sales, these two components are almost equal. In both cases, about half the total variation is accounted for by trading-day variation. Trading-day variation is of prime concern when attempting to assess the underlying cyclical movement over short spans (1 or 2 months). Over longer spans trading-day variation is of less importance, since it frequently reverses direction and does not cumulate as do the seasonal and cyclical movements.

Trading-day variation cannot be estimated in series where the irregular variation is large. There is an upper limit, in economic series, to the amount of variation arising from trading days. For example, the monthly change induced by a 5-day week[3] averages 5.3 percent and the change induced by a 6-day week averages 2.8 percent. When $\bar{I}$ is between 5 and 10 percent, the monthly change of 3 to 5 percent or less in the trading-day variation ceases to be important and its estimation becomes difficult. Likewise, there is less need for a trading-day adjustment since there are large irregular variations which will continue to obscure the trend-cycle even after the trading-day variation has been removed.

Census Bureau series covering a broad spectrum of economic activity show less variation attributable to the calendar than would be expected. For example, assume that a manufacturing plant operates on a 5-day week. We would expect, then, the volume of activity in a January with 23 workdays (4 Saturdays and 4 Sundays) to be 9 percent higher than in a January with 21 workdays (5 Saturdays and 5 Sundays). This often does not appear to be the case. Instead, the January with 23 workdays is only about 4- or 5-percent higher. Such a phenomenon can arise either from factors inherent in the economy or from various practices of recording and reporting data. It appears that a substantial proportion of economic activity occurs on the basis of monthly plans and schedules that are drawn up with little or no attention to the number of trading days within the calendar months and/or is recorded and reported on a basis that takes little account of the number of trading days in calendar months.

The implication of this moderation in trading-day variation is that the narrow definition implicit in many trading-day adjustments does not allow for the types of variation that may actually exist in the monthly data. Using such a definition of trading-day variation, one might, after noting that an economic activity operates on a 5-day week, proceed

[1] The term "trading-day variation" can be considered interchangeable with the terms "working-day variation" and "calendar variation."

[2] A familiarity with seasonal-adjustment techniques is assumed. At times, reference is made to specific measures provided by the Census Method II ratio-to-moving-average method of seasonal adjustment. Details concerning Census Bureau methods can be found in references 7 and 9 (see end of paper).

[3] A set of daily weights where Mon., . . . , Fri. = 1.4; Sat., Sun., = 0.0 is commonly referred to as a 5-day week (also 6-day, 5 1/2 day, etc.).

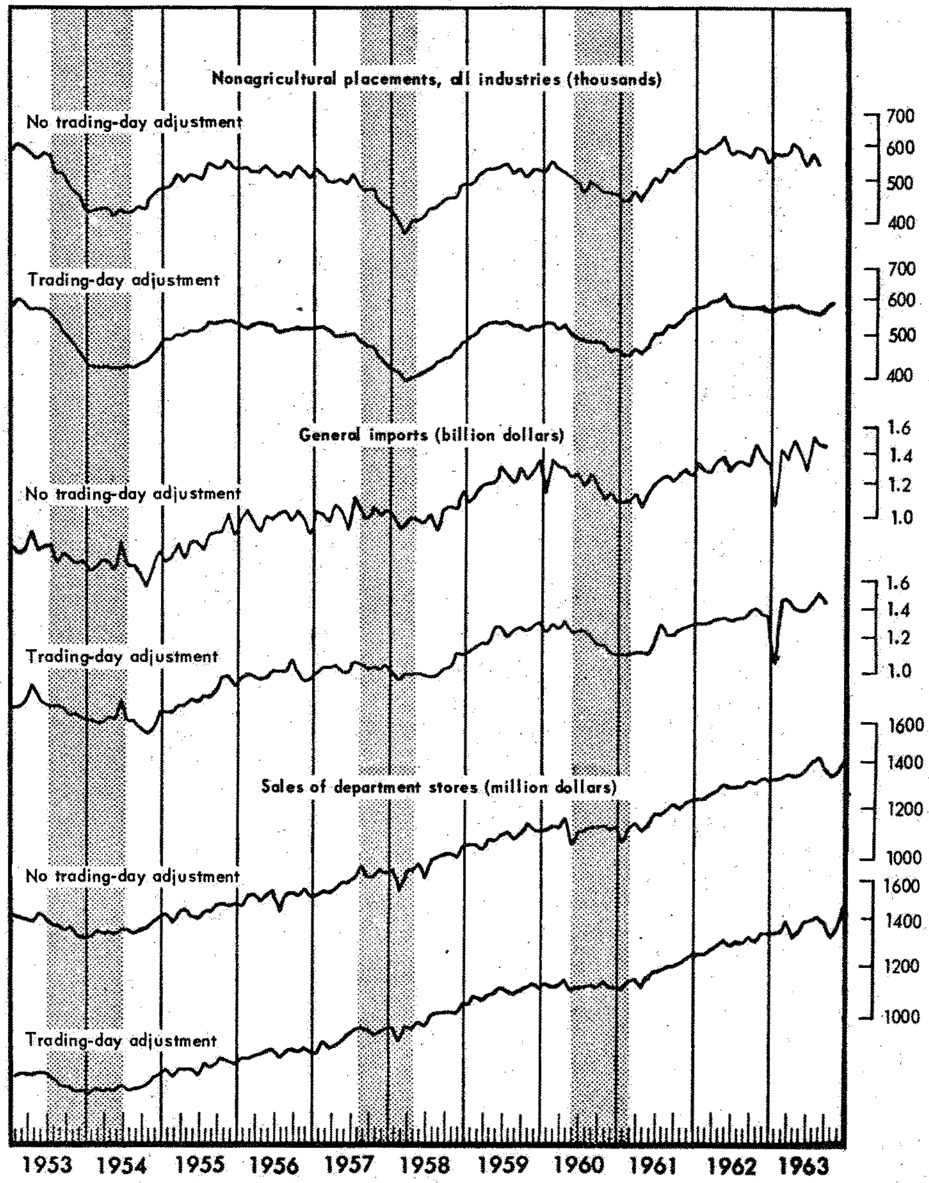
Table 1.--PERCENT DISTRIBUTION OF VARIATION ATTRIBUTABLE TO THE VARIOUS COMPONENTS OF MONTH-TO-MONTH VARIATION, FOR SELECTED BUSINESS ACTIVITIES DURING INDICATED TIME PERIODS

Business activity	Time period	Components of month-to-month variation				
		Trading-Day variation	Seasonal variation	Irregular variation	Trend-cycle variation	Total variation
Retail sales	1953-63	7.7	[1]91.1	0.8	0.4	100.0
Wholesale sales	1960-63	47.8	48.7	2.9	0.6	100.0
Manufacturers' shipments	1953-62	20.1	71.9	5.6	2.5	100.0
Manufacturers' new orders	1953-62	19.3	62.6	14.3	3.8	100.0
U.S. exports	1953-63	18.3	63.1	15.8	2.8	100.0
U.S. imports	1953-63	50.3	38.2	10.0	1.5	100.0
Building permits	1954-62	21.2	69.1	7.8	1.9	100.0

[1] Includes a slight contribution from variation in sales of selected kinds of business in March and April because of shifting date of Easter.

Note: See appendix C for the derivation of these measures from Census Method II summary measures of monthly change, $\bar{O}$, $\bar{S}$, $\bar{I}$, $\bar{C}$.

COMPARISON OF SELECTED ECONOMIC SERIES BEFORE AND AFTER TRADING-DAY ADJUSTMENT



The shaded areas represent contraction periods in general business activity.

Sources of data: Nonagricultural placements, all industries--Department of Labor, Bureau of Employment Security (seasonal and trading-day adjustment by Bureau of the Census); general imports, total, and sales of department stores--Bureau of the Census.

Estimated daily weights used in making the trading-day adjustments:

	Mon.	Tues.	Wed.	Thur.	Fri.	Sat.	Sun.
Nonagricultural placements.	1.00	1.38	1.26	.97	1.23	.45	.71
General imports.	1.61	1.47	1.33	1.33	1.26	0.00	0.00
Department store sales.	.94	1.00	.99	1.14	1.17	1.35	.42

to adjust for a 5-day week with no activity on Saturday and Sunday. That is, the series would be placed on a daily-rate basis by dividing by the number of weekdays in each month. Results presented below show that such a procedure often yields much poorer results than do the techniques described in this paper which relate the variations in the monthly series to the number of times that each day of the week occurs in the month.

This paper describes the concept of trading-day variation. It reports the techniques for estimating trading-day variation and the results of applying these methods to artificial and real economic series. It also presents tests for determining if trading-day variation is present. On the basis of the work reported in this paper, a routine to estimate trading-day variation is being added to the Census Method II seasonal adjustment program.

II. DEFINITIONS

A. Characteristics of the Calendar

Before defining trading-day variation in economic series, an understanding of two features of the calendar is necessary. We are familiar with the fact that months vary in length—31 days, 30 days, and Februaries of 28 and 29 days. Such variation between 31-day months, 30-day months, and Februaries is termed "length-of-month" or "between-month" variation.

In addition to this variation, the composition of each calendar month varies from year to year. Twenty-two types of calendar months occur—seven 31-day months, seven 30-day months, and eight Februaries.[4] Let us examine them by the day on which they begin. Thirty-one-day months beginning on Monday contain five Mondays, Tuesdays, and Wednesdays and four of the other days of the week. Thirty-one-day months beginning on Tuesday contain five Tuesdays, Wednesdays, and Thursdays and four of the other days of the week, and so on for the other types of 31-day months. (See table 2). In a similar fashion, 30-day months beginning on Monday contain a fifth Monday and Tuesday and so on. Februaries contain 28 days except in leap years. There are seven types of leap-year Februaries distinguished by the beginning day of the month. This variation between months of equal length in the number of times a particular day or days of the week occur in a calendar month is termed "calendar composition" or "within-month variation."

A 28-year cycle occurs before the calendar begins to repeat the same pattern of days and months. (The 28-year cycle is broken at the beginning of each century which is not divisible by 400; i.e., at 1900 and 2100, but not at 2000.) Dates of movable holidays such as Easter do not conform to a regular pattern during the 28-year cycle. A calendar of these 22 types of months for the 28-year period, 1944-71, is shown as table 3. By subtracting or adding 28 years to the period shown, the calendar can be adapted for the periods 1916-43 and 1972-99.

Let us look at these properties of the calendar in more detail by examining a tabulation of the number of weekdays, Saturdays, and Sundays for 5 years, 1960-64, and the average number over the 28-year cycle shown in table 4, page 4. Note the magnitude of the within-month variation. May, a 31-day month, varies from 21 to 23 weekdays a month, a 9-percent range; while June, a 30-day month varies from 20 to 22 week-

[4]If month-to-month changes are considered there are 49 combinations in the calendar. There are 7 combinations between each of the sequences of 31- to 30-day, 30- to 31-day and 31- to 31-day months and 14 between 31-day months to Februaries and Februaries to 31-day months. This paper considers the pattern for 22-month types appropriate for series such as production and sales. For series such as changes in inventories it may be useful to consider the 49 combinations of monthly changes.

Table 2.--NUMBER OF TIMES EACH DAY OF THE WEEK OCCURS IN EACH OF 22 TYPES OF MONTHS

Type-of-month code	First day of month	Number of specified days per month						
		Sundays	Mondays	Tuesdays	Wednesdays	Thursdays	Fridays	Saturdays
		31-day months						
1	Monday.......	4	5	5	5	4	4	4
2	Tuesday......	4	4	5	5	5	4	4
3	Wednesday....	4	4	4	5	5	5	4
4	Thursday.....	4	4	4	4	5	5	5
5	Friday.......	5	4	4	4	4	5	5
6	Saturday.....	5	5	4	4	4	4	5
7	Sunday.......	5	5	5	4	4	4	4
		30-day months						
8	Monday.......	4	5	5	4	4	4	4
9	Tuesday......	4	4	5	5	4	4	4
10	Wednesday....	4	4	4	5	5	4	4
11	Thursday.....	4	4	4	4	5	5	4
12	Friday.......	4	4	4	4	4	5	5
13	Saturday.....	5	4	4	4	4	4	5
14	Sunday.......	5	5	4	4	4	4	4
		Leap-year Februaries						
15	Monday.......	4	5	4	4	4	4	4
16	Tuesday......	4	4	5	4	4	4	4
17	Wednesday....	4	4	4	5	4	4	4
18	Thursday.....	4	4	4	4	5	4	4
19	Friday.......	4	4	4	4	4	5	4
20	Saturday.....	4	4	4	4	4	4	5
21	Sunday.......	5	4	4	4	4	4	4
		Non-leap-year Februaries						
22	Any day......	4	4	4	4	4	4	4

Table 3.--CALENDAR, BY TYPE-OF-MONTH CODE, 1944 to 1971

(Figures are type-of-month codes. See table 2 for the number of times each day of the week occurs in each of the months)

Year	January	February	March	April	May	June	July	August	September	October	November	December
1944......	6	16	3	13	1	11	6	2	12	7	10	5
1945......	1	22	4	14	2	12	7	3	13	1	11	6
1946......	2	22	5	8	3	13	1	4	14	2	12	7
1947......	3	22	6	9	4	14	2	5	8	3	13	1
1948......	4	21	1	11	6	9	4	7	10	5	8	3
1949......	6	22	2	12	7	10	5	1	11	6	9	4
1950......	7	22	3	13	1	11	6	2	12	7	10	5
1951......	1	22	4	14	2	12	7	3	13	1	11	6
1952......	2	19	6	9	4	14	2	5	8	3	13	1
1953......	4	22	7	10	5	8	3	6	9	4	14	2
1954......	5	22	1	11	6	9	4	7	10	5	8	3
1955......	6	22	2	12	7	10	5	1	11	6	9	4
1956......	7	17	4	14	2	12	7	3	13	1	11	6
1957......	2	22	5	8	3	13	1	4	14	2	12	7
1958......	3	22	6	9	4	14	2	5	8	3	13	1
1959......	4	22	7	10	5	8	3	6	9	4	14	2
1960......	5	15	2	12	7	10	5	1	11	6	9	4
1961......	7	22	3	13	1	11	6	2	12	7	10	5
1962......	1	22	4	14	2	12	7	3	13	1	11	6
1963......	2	22	5	8	3	13	1	4	14	2	12	7
1964......	3	20	7	10	5	8	3	6	9	4	14	2
1965......	5	22	1	11	6	9	4	7	10	5	8	3
1966......	6	22	2	12	7	10	5	1	11	6	9	4
1967......	7	22	3	13	1	11	6	2	12	7	10	5
1968......	1	18	5	8	3	13	1	4	14	2	12	7
1969......	3	22	6	9	4	14	2	5	8	3	13	1
1970......	4	22	7	10	5	8	3	6	9	4	14	2
1971......	5	22	1	11	6	9	4	7	10	5	8	3

NOTE: By subtracting or adding multiples of 28 years, the calendar can be used for the period 1901 to 2099. The 28-year cycle is broken at 1900 and 2100 because these years are not leap years. For 1900 and earlier years, see an almanac or a perpetual calendar.

days a month, a 10-percent range. Such variation arises only when we distinguish between particular days of the week.

Also note that the 3 percent difference in length of month between May with 31 days and June with 30 days appears in the average number of weekdays over the 28-year period, 22.14 weekdays for May and 21.43 for June. Unlike within-month, length-of-month variation arises whether we consider the total number of days or distinguish between particular days of the week in the month.

B. Definition of Trading-Day Variation

Trading-day variation could be described as the variation in a monthly economic time series related to the calendar composition and length-of-month variation in the calendar. However, variations in the volume of activity arising from months of different length (except for leap-year and non-leap-year Februaries) cannot be statistically separated from other seasonal influences also causing differences between months. Therefore, length-of-month variation is defined and estimated as part of the seasonal component.[5] On the other hand, within-month variation which varies abruptly from year to year within a month, is not included in the definition or estimation of the seasonal. Our definition, then, becomes simply this: Trading-day variation is the monthly variation in a series related to the within-month variation or calendar composition.

[5] The seasonal can be defined as a constant or gradually increasing or decreasing function within a month over several years.

Table 4.--NUMBER OF WEEKDAYS, SATURDAYS, AND SUNDAYS IN MAY AND JUNE 1960 to 1964

Year	May			June		
	Week-days	Satur-days	Sun-days	Week-days	Satur-days	Sun-days
1960.........	22	4	5	22	4	4
1961.........	23	4	4	22	4	4
1962.........	23	4	4	21	5	4
1963.........	23	4	4	20	5	5
1964.........	21	5	5	22	4	4
28-year average.....	22.14	4.43	4.43	21.43	4.29	4.29

C. Internal vs. External Evidence

Note that the definition of trading-day variation, above, is in terms of the monthly variation in the series. The trading-day adjustment implied by this definition is derived from the internal evidence contained in the monthly data. This is a more general approach than that of the frequently used techniques which rely upon external or a priori evidence of the daily rates of activity in making a trading-day adjustment. The differences between the use of internal and external evidence are subtle and are considered in some detail in this section.

The customary use of external evidence is to estimate the proportion of the week's activity occurring, on average, on each day of the week. In the remainder of this paper these estimates are referred to as the seven actual daily rates of activity. The available external evidence often consists of the daily schedule of hours of work, direct personal observation of the activity, or records of the daily volume of activity. Although such evidence is usually limited, it is often considered sufficient to provide estimates of seven actual daily rates of activity.

To make the trading-day adjustment, each monthly figure is divided by a factor constructed by aggregating the daily rates to a monthly sum. The construction of the factors can be represented as follows:

(Equation 1)

$$M_i = \frac{X_{1i}r_1 + X_{2i}r_2 + \ldots + X_{7i}r_7}{N_i} \quad (i = 1, \ldots, n),$$

where M_i is the trading-day adjustment factor for month i;

X_{ji} is the number of times day-of-the-week j occurs in month i;

r_j is the daily rate, i.e., the proportion of the week's activity that occurs on the day-of-the-week j ($j = 1, \ldots, 7$), where $\sum_1^7 r_j = 7$;

N_i is either 31, 30, or 28.25 depending upon whether month i is a 31-day month, 30-day month, or February;

n is the number of months of data available.

Use of such factors implies that the seven actual daily rates (the r's) can be aggregated to weekly and then to monthly levels corresponding to the actual monthly variation.

This assumption implied in the use of seven actual daily rates of activity is incorrect when part or all of the economic process operates under schedules that do not take into account the calendar composition of the month. Allowance must be made for the relation of the activity on each day of the week to the monthly volume of activity rather than for the relation of the daily activity to the weekly volume of activity. The relation of the activity on each day of the week to the monthly volume can vary with the calendar composition of the month, while the actual daily rates that are proportions of the weekly volume remain fixed.

The following simplified example illustrates the relationship that sometimes tends to exist between the daily, weekly, and monthly activity. Suppose that all consumers receive one-twelfth of their annual income each month and that they dispose of all their month's income before the next payday by shopping at retail stores. If the retail stores are open Monday through Saturday, the seven daily rates of activity include a zero Sunday, and we can assume the other days to be equal; i.e., Mon., . . ., Sat. = 1.17; Sun. = 0.0; Total = 7.00. On the basis of the daily rates of activity, we are led to conclude that the monthly volume of sales in a January with 23 shopping days is 9 percent higher than in a January with 21 shopping days, other things being equal. However, this is not the case. Since the consumers dispose of an equal amount of income in both Januaries, the monthly volume of sales in the two Januaries is the same. Although the daily rates expressed as a percent of the week remain the same in the two months, the daily activity for Monday through Saturday expressed as a percent of the total monthly activity is higher by 9 percent when the monthly total is spread over 21 rather than 23 shopping days. In this example, the occurrence of Sunday or any other day five rather than four times has no effect upon the monthly variation and knowledge of the daily activity as a percent of the week's activity is extraneous.

The same relationship could be illustrated if, instead of assuming a fixed amount of income each month, we assume a continuous daily or weekly use that affects the timing of the purchases. This possibility is most obvious for grocery sales. The amount of groceries purchased by the consumer in January might reflect the fact that approximately the same amount of food is placed upon his dining table regardless of the calendar composition of the month.

Note that in one respect the two examples are different. In the first, the fixed amount of activity that occurs each January, regardless of calendar composition, is the same as the fixed amount that would occur in a 30-day month or a February. In the second example the amount of activity that is fixed with respect to calendar composition is not fixed with respect to the length of the month. More food is used in 31-day months than in 30-day months or in Februaries. At times, recognition of these two possibilities will be useful.

There may also be other types of relationships between daily, weekly, and monthly activity that are more complex than that implied in the use of seven actual daily rates. For example, total activity this month might reflect next month's demand which varies with next month's calendar composition. To allow for such complex variations, the relations of the daily activity to the monthly activity must be examined for each type of calendar composition. This examination can be done either by compiling much more external evidence of daily activity than is customary or practical, or by statistically relating calendar composition to the variation in the monthly activity. Although the above examples are oversimplified and extreme, it appears reasonable to assume that there are, in many areas of the economy, tendencies towards more complex relationships than those implied by the use of seven daily rates of activity. The empirical evidence presented in section IV, B supports such a supposition. Even though there may be other unknown explanations for the empirical evidence, it appears worthwhile for the present to accept the above line of reasoning rather than the rigid approach imposed by the use of seven actual daily rates of activity.

A second problem with external evidence is the availability of information. Often the cost of obtaining the information forces the estimation of the actual daily rates to be based upon insufficient information, particularly with respect to such factors as overtime practices and continuous process activities that operate around the clock. The availability of information is a problem, whether or not the assumption implied by the use of seven actual daily rates of activity holds true for a particular series.

Finally, there is a third problem which is an extension of the second. Often there are unobserved bookkeeping, reporting, and data-processing practices related to calendar composition that are at work modifying the actual variations. These modifications cannot readily be observed from external evidence.

The first and third shortcomings of external evidence discussed above are not present when an adjustment is derived from internal evidence. An adjustment based upon internal evidence allows for the total variation that can be related to calendar composition; that is, to the occurrence of a particular day of the week five times rather than four times in a

Table 5.—IRREGULAR COMPONENT FOR DEPARTMENT STORE SALES, ARRANGED BY CALENDAR COMPOSITION, 1953 to 1962

(I = Irregular component; date = month and year of basic data)

31-day months beginning on--													
Sunday		Monday		Tuesday		Wednesday		Thursday		Friday		Saturday	
I	Date	I	Date	I	Date	I	Date	I	Date	I	Date	I	Date
99.2	Mar. 1953	100.4	Mar. 1954	100.8	Dec. 1953	101.0	July 1953	101.0	Jan. 1953	100.5	May 1953	98.5	Aug. 1953
97.9	Aug. 1954	99.3	Aug. 1955	101.2	Mar. 1955	101.1	Dec. 1954	101.1	Oct. 1953	98.1	Jan. 1954	98.0	May 1954
98.2	May 1955	99.4	Oct. 1956	101.3	May 1956	101.6	Aug. 1956	101.2	July 1954	98.6	Oct. 1954	100.7	Jan. 1955
98.1	Jan. 1956	100.1	July 1957	99.8	Jan. 1957	101.2	May 1957	99.6	Dec. 1955	100.6	July 1955	100.5	Oct. 1955
95.3	July 1956	101.4	Dec. 1958	98.4	Oct. 1957	103.0	Jan. 1958	100.9	Mar. 1956	98.7	Mar. 1957	97.8	Dec. 1956
100.1	Dec. 1957	100.3	Aug. 1960	101.5	July 1958	100.4	Oct. 1958	103.9	Aug. 1957	101.1	Aug. 1958	99.9	Mar. 1958
95.8	Mar. 1959	100.2	May 1961	99.2	Dec. 1959	101.2	July 1959	103.3	May 1958	100.4	May 1959	97.3	Aug. 1959
96.1	May 1960	100.4	Jan. 1962	99.1	Mar. 1960	101.9	Mar. 1961	102.1	Jan. 1959	99.9	Jan. 1960	101.2	Oct. 1960
96.5	Jan. 1961	98.3	Oct. 1962	100.9	Aug. 1961	100.8	Aug. 1962	102.8	Oct. 1959	99.9	July 1960	99.5	July 1961
97.7	Oct. 1961	.....	....	102.5	May 1962	.....	....	101.7	Dec. 1960	101.3	Dec. 1961	96.1	Dec. 1962
99.8	July 1962	.....	....	.....	....	.....	....	101.2	Mar. 1962	.....	....	.....	....
Mean													
97.7		100.0		100.5		101.4		101.7		99.9		99.0	
30-day months beginning on--													
Sunday		Monday		Tuesday		Wednesday		Thursday		Friday		Saturday	
I	Date	I	Date	I	Date	I	Date	I	Date	I	Date	I	Date
98.9	Nov. 1953	100.6	June 1953	99.8	Sept.1953	100.2	Apr. 1953	101.4	Apr. 1954	102.8	Apr. 1955	100.5	Sept. 1956
98.3	Apr. 1956	101.1	Nov. 1954	101.4	June 1954	100.2	Sept.1954	101.2	Sept.1955	102.3	June 1956	98.9	June 1957
97.9	Sept.1957	99.2	Apr. 1957	100.1	Nov. 1955	98.3	June 1955	102.7	Nov. 1956	101.2	Nov. 1957	97.1	Nov. 1958
96.0	June 1958	99.8	Sept.1958	100.3	Apr. 1958	101.3	Apr. 1959	100.1	Sept.1960	104.3	Apr. 1960	97.4	Apr. 1961
98.6	Nov. 1959	101.1	June 1959	99.7	Sept.1959	99.8	June 1960	101.3	June 1961	100.6	Sept.1961	99.6	Sept. 1962
98.2	Apr. 1962	.....	....	99.5	Nov. 1960	100.6	Nov. 1961	102.3	Nov. 1962	99.2	June 1962	.....	....
Mean													
98.0		100.4		100.1		100.1		101.5		101.7		98.7	

Februaries (Mean 99.0[1])

I	Year	I	Year	I	Year	I	Year	I	Year	I	Year	I	Year	I	Year	I	Year	I	Year
99.7	1953	100.6	1954	97.1	1955	[2]101.5	1956	100.8	1957	95.1	1958	100.5	1959	[2]100.7	1960	100.9	1961	97.2	1962

[1]Does not include leap-year Februaries. [2]Leap year.

month.[6] The second shortcoming, lack of sufficient information, is sometimes a problem with internal evidence in the sense that statistically reliable estimates cannot be obtained when there are too few observations or there is too large an unexplained residual.

III. ESTIMATION

In the standard ratio-to-moving-average methods of seasonal adjustment such as Census Method II, the unadjusted data is sequentially separated into three components, trend-cycle, seasonal, and irregular (i.e., the residual). When trading-day variation is present in the unadjusted data, it is primarily included in the estimate of the irregular component, rather than the seasonal or trend-cycle, since in many ways it more closely resembles the random fluctuations in the irregular component. (Over a period of months trading-day variation frequently reverses direction while the trend-cycle is a smooth curve; over a period of years within a month it also frequently reverses direction while the seasonal is a constant or a gradually increasing or decreasing curve.) To estimate trading-day variation from internal evidence, it is necessary, therefore, to examine the estimate of the irregular component (more accurately termed the combined irregular and trading-day-variation component) provided by the ratio-to-moving-average seasonal adjustment.

In simultaneous solutions for the seasonal, trend-cycle, and irregular components an allowance for trading-day variation can be included in a manner analogous to those described below for the ratio-to-moving average method.

A. Grouping Months by Calendar Composition

The simplest approach to estimating trading-day variation is to sort the values of the combined irregular and trading-

6Some earlier studies have considered trading-day variations to be more complex than merely counting the number of days the store is open each month: (a) Kuznets (4) states that "To establish the number of working days in an industry is often impossible, the number being at best an estimated. In many economic processes it is difficult to assume that volume of activity is directly proportional to the number of working days (for example bank clearings or retail sales)"; (b) Eisenpress (3) presents a method of adjusting bank debits for trading-day variation where Sunday, a nontrading day (banks are closed), does not receive a zero weight, but one determined by the variation in the monthly series; (c) Marris (5) widens the concept of trading-day variation and introduces the effect of reporting practices upon trading-day variation. Marris also presents a summary of other work in the field. (See references at end of paper.) The present paper attempts to further modify the concept of trading-day variation suggested by Marris by considering in more detail and emphasizing the factors that modify the actual daily rates of activity. These factors virtually require that trading-day variation be considered at the level of monthly activity, not daily activity.

day component on the basis of calendar composition, thus placing them in 22 different groups. Significant variation between these groups is evidence of trading-day variation. For example, the irregular values for months that contain five Saturdays might tend to be higher or lower than the irregular values for months with four Saturdays. To make a trading-day adjustment, the means of the irregular values for each group are computed and then divided into the series.

Table 5 illustrates this technique for making a trading-day adjustment. It shows the irregular component of department store sales arranged by calendar composition. The means of the groups of irregulars, shown at the bottom of the table, vary between 101.7 (for 31-day months beginning on Thursday) and 97.7 (for 31-day months beginning on Sunday). Because of the length of the series, no means are shown for leap-year Februaries.

This technique for adjusting for trading-day variation is similar to the technique of computing stable seasonals where all Januaries are placed together, all Februaries, and each of the other months, and the mean for each group is taken as an estimate of the seasonal. The only difference is the framework in which the data is ordered. To compute the seasonal, we group months of like name because the rate of activity is similar in the same month each year. To compute the trading-day factor, we group together months of like calendar composition, since it is the number of each type of day of the week in the month that gives rise to trading-day variation.

B. Estimating Daily Weights

It is possible to refine the above approach to trading-day variation. The means of the irregular values for each of the 22 groups are only estimates of the effect of particular days of the week in combination with adjacent days. Also, in a series of less than 28 years, there is only one observation for some types of leap-year Februaries and none for the other types. The refinement is to estimate seven weights, one for each day of the week, using the data from all months. By reducing the system to seven estimates, the separate effect of each day of the week is determined and the reliability of the estimates is increased.

These seven weights, at the risk of confusion with the seven actual daily rates of activity discussed before, will be referred to as daily weights. Before proceeding, it is necessary to stress the relation between these seven daily weights derived from internal evidence and the seven daily rates derived from external evidence. The daily weights will correspond to the actual daily rates if the variation in the monthly volume of activity corresponds to the variation in the monthly factors constructed from the daily rates, r_j, in equation 1. Otherwise the daily weights will not correspond with the actual daily rates. When there is no correspondence, either the economic process contains complex relationships between the daily, weekly, and monthly volumes of activity or there are bookkeeping and reporting practices that are affecting the data. Experience at the Census Bureau suggests that often there is sufficient lack of correspondence between the daily rates and the actual daily weights to frustrate attempts to interpret the estimated daily weights in light of a known pattern of daily activity. (See, for example, the estimated daily weights in table 8.) The daily weights must usually be viewed only as statistical weights representing the effect of several variables. In the next section we shall return to the problem of interpreting the daily weights.

The Bureau of the Census has used two methods of estimating these daily weights. The first method was developed at the Organization for Economic Cooperation and Development (OECD)(5). It also was developed and used by the Federal Reserve Board.[7] Its chief merit is that the computations are simple and can be done by hand. The method now used at the Census Bureau and being introduced in the Census Method II seasonal-adjustment program is a regression method which makes more complete use of the data and yields more reliable estimates then the OECD method. Since multiple regression techniques are used, it requires an electronic computer for efficient application. Also, it can be combined more readily with holiday or seasonal adjustments in simultaneous solutions rather than the sequential solution used in Census Method II.

The trading-day routine that includes the regression estimates of the daily weights (developed more fully in appendix A) is being added to Census Method II in the following sequence:

(1) A seasonal adjustment is first made with no trading-day adjustment or with a given set of daily weights (referred to as a prior adjustment). This adjustment produces an irregular component which is composed of (a) the "true" irregular and (b) the trading-day variation (or the residual trading-day variation if a prior adjustment was made).

(2) This irregular is then modified for extreme values which would tend to distort the estimated daily weights and regressed upon seven independent variables representing the number of times each day of the week occurred in the month in the following manner:

(Equation 2)

$$I_i = \frac{X_{1i}B_1 + X_{2i}B_2 + \cdots + X_{7i}B_7 + E_i}{N_i},$$

where I_i is the modified irregular for month i and $E[I_i] = I_i$;

X_{ji} is the number of times day-of-the-week j occurs in month i;

B_j's are the seven daily weights, where $\sum_1^7 B_j = 7$;

N_i is either 31, 30, or 28.25 depending upon whether month i is a 31-day, 30-day month, or February;

E_i is the "true" irregular for month i.

(3) The trading-day adjustment factors are the estimated values $\hat{I}_i$ of the I_i, that is,

(Equation 3)

$$\hat{I}_i = \frac{X_{1i}b_1 + X_{2i}b_2 + \ldots + X_{7i}b_7}{N_i},$$

where b_j is the least-squares estimate of B_j (j = 1, . . ., 7).

In the Method II routine, they are divided into the unadjusted data and then a second seasonal adjustment is made, based upon trading-day adjusted data. Details of how the trading-day routine is inserted in the Method II sequence will be included in forthcoming specifications for a new version of the Method II program.

[7]The OECD article (5) presents equations for 31-day months only. Federal Reserve Board and Census Bureau have extended the approach by also making estimates of daily weights from the irregulars for 30-day months and combining the 31- and 30-day estimates by weighting them by the ratios 7/11 and 4/11, respectively. The combined estimates were found to have smaller σ's than the 31-day month estimates.

If a prior trading-day adjustment was made, the estimated set of residual weights (b_j) may be combined with the set of prior weights (p_j) to obtain total weights (D_j) by the formula $D_j = b_j + P_j - 1$.

(4) One may perform a standard t-test to determine whether an estimated weight b_j is significantly different from any specified value and an F-test to determine the significance of the regression (i.e., the existence of significant trading-day variation in the irregular).

C. Interpreting Daily Weights

As stated above, if the assumption implied by the use of seven daily rates of activity holds true for a paticular series, the estimated daily weights will correspond with the actual daily rates of activity. (This can be seen by comparing equations 1 and 2.) For example, the negligible weekend activity in many areas of the economy is reflected in estimated weights approximating Mon., . . . , Fri. = 1.4; Sat., Sun. = 0.0 for some types of series. Another instance is the retail sale of nondurable goods where the daily-sales pattern tends to be reflected in the estimated weights, Mon. = 0.9, Tues. = 0.9, Wed. = 0.9, Thurs. = 1.0, Fri. = 1.4, Sat. = 1.4, Sun. = 0.5.

When none of the monthly activity varies with the calendar composition (there is no trading-day variation), the estimated daily weights all equal "1", regardless of the pattern of activity within the week. When only a portion of the series is independent of calendar composition, the estimated daily weights can be considered as composed of two parts, the first part having differential weights for each day and the other consisting of equal weights for each day. Consider the following hypothetical set of daily weights:

Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.	Total of weights	Percent
1.20	1.20	1.20	1.20	1.20	0.50	0.50	7	100

These weights can be separated into two parts:

A..	.70	.70	.70	.70	.70	0.0	0.0	3.5	50
B..	.50	.50	.50	.50	.50	.50	.50	3.5	50

where "A" reflects that part of the series which varies with the number of different-type days and accounts for 50 percent of the series, and "B" is independent of calendar composition and accounts for 50 percent of the total series. Part B of the weights may be divided into two types of activity: (B1) activity which varies only with the length of the month, and (B2) activity which is constant from month to month.[8] In addition, bookkeeping practices can modify the actual variations represented by the A and B parts of the weights.

While a particular series may not contain all the types of variations mentioned above, such types of variations are widespread. Consider, for example, activity at the manufacturing level. First, in many industries a 5-day week is prevalent, which exerts a tendency for the total monthly activity to vary with the number of weekdays. This activity is represented by "A". Second, there is some overtime work and continuous processes that operate 7 days a week. There may also be monthly schedules or continuing demand which results in activity that is independent of calendar composition and varies with the length of the month. This activity is represented in "B1". Third, there may be some activities such as those operating under contracts calling for one-twelfth of the year's scheduled activity to be completed each month, that are independent of calendar composition and the length of the month, represented in "B2". Finally, bookkeeping practices can be represented in either "A" or "B".

Bookkeeping reports that do not cover precisely the calendar month affect trading-day variation and thereby the estimated daily weights. For example, suppose a firm that operates on a 6-day week, (Mon., . . . , Sat. = 1.17, Sun. = 0.0) follows the practice of closing its books for the month on Friday whenever the month ends on Saturday (which occurs only about twice a year) and including the Saturday activity in the following month. This practice would apply to 31-day months beginning on Thursday and 30-day months beginning on Friday. For such months, reported sales would be decreased by almost 4 percent. Conversely, reported sales are increased by almost 4 percent for the months following, which always begin on Sunday. Reporting sales in this manner results in a monthly series which yields daily weights of Mon., . . . , Fri. = 1.17, Sat. = 0.0 and Sun. = 1.17 rather than the actual daily rates, which had a zero Sunday.[9]

Another bookkeeping practice that is used by some firms and affects trading-day variation is the plan known as the 4-4-5 plan in which the first and second months of each quarter always contain exactly 4 weeks and the third month 5 weeks. This practice eliminates trading-day variation, since each period has exactly the same number of each type of day of the week as the corresponding periods in earlier years. When some activities are reported under such plans and others are reported on a strict calendar-month basis, the combined variation can be represented by the above example where the "B" weights represent the portion of the series that is independent of calendar composition.

From the information contained in the monthly data, we cannot actually determine how much activity occurs on each day of the week. All we can determine is how the monthly series varies with the composition of the calendar. For example, the activity represented by the Sunday "B" weight does not necessarily indicate that the activity occurs on Sunday. It may be distributed among the other days in some undeterminable fashion. Without recourse to other information, it is impossible to separate the various factors affecting the daily weights. What the weights do represent, after they are combined into monthly factors, is the monthly activity.

From the above discussion, it is apparent that the quality of a trading-day adjustment should be determined by its effect upon the monthly series, not by comparing the estimated weights with what is either supposed or known to be the actual daily rate of activity.

D. Constructing Monthly Trading-Day Adjustment Factors

The monthly trading-day adjustment factors are simply the estimated values, $\hat{I}_i$, shown in equation 3. These factors are divided into the data to remove the estimated trading-day variaton.

[8] The difference between B1 and B2 is that activity represented by B1 varies with the length of the month. Since the seasonal removes length-of-month differences (except for differences between Februaries), variations explained by B1 and B2 are indistinguishable in the irregular component. (See footnote 10 for a further discussion.)

[9] The effect upon the Sunday weight may appear extreme since the Saturday is only 1/30 or 1/31 of the month. However, 4 full weeks, 28 days, are found in all months, and only 2 or 3 days are unique to each month. The shift of one day's activity is 1/2 or 1/3 of the monthly variation attributable to calendar composition.

Even though length-of-month variation can be removed by the seasonal, a common practice has been to include an allowance for it in the trading-day factors. When such factors are divided into the unadjusted data, followed by a recomputation of the seasonal factors, based upon the trading-day adjusted data, a compensating revision occurs in the seasonal factors. The same seasonally adjusted data, therefore, is obtained as would have been obtained if no allowance for the length of the month had been included in the trading-day factor. When it is desired to follow such a procedure, the following form is often used.[10]

(Equation 4)

$$\hat{I}_i' = \frac{X_{1i}b_1 + X_{2i}b_2 + \cdots + X_{7i}b_7}{30.4375},$$

where $30.4375 = \frac{365.25}{12}$ (the average length of month);

b_j's are those estimated in equation 2.

[10]Further consideration suggests that if part of the series is independent of the length of the month, equation 4 is incorrect conceptually for all months and equation 3 for Februaries.

In the example in section III, C, the "B" part of the series that does not vary with the calendar composition was explained in part by "B1" which varies only with the length of the month and "B2" which is constant from month to month. Statistically "B1" and "B2" can only be distinguished by comparing leap-year and non-leap-year Februaries, since the seasonal factor compensates for other differences in the length of month. This distinction was ignored in estimating the daily weights because only 1 out of every 48 values is involved.

To construct the adjustment factors when part of the series is independent of the length of the month, the following forms are correct:

Where a_1 is the portion of series dependent upon calendar composition and length of month (this corresponds to the "A" weights in section III, C);

a_2 is the portion of series dependent only upon the length of month, i.e., (B1);

a_3 is the portion of series which does not vary from month to month, i.e., (B2);

$a_1 + a_2 + a_3 = 1.0$;

N_i is either 31, 30, or 28.25;

c_j = the (A) part of the b_j weights in IIIc; and

$$a_1 = \frac{\sum_1^7 c_j}{7};$$

then equation 4 becomes—

(Equation 4A)

$$\hat{I}_i'' = \frac{X_{1i}c_1 + X_{2i}c_2 + \cdots + X_{7i}c_7}{30.4375} + \frac{N_i a_2}{30.4375} + a_3.$$

Likewise, for leap-year Februaries, revise equation 3—

(Equation 3A)

$$\hat{I}_i''' = \frac{X_{1i}c_1 + X_{2i}c_2 + \cdots + X_{7i}c_7}{28.25} + \frac{29\,a_2}{28.25} + a_3;$$

and for non-leap-year Februaries,

(Equation 3A)

$$\hat{I}_i''' = \frac{X_{1i}c_1 + X_{2i}c_2 + \cdots + X_{7i}c_7}{28.25} + \frac{28\,a_2}{28.25} + a_3.$$

Since for a single series there are not enough observations to make a reliable determination of a_2 and a_3, it is necessary to make an arbitrary assumption. What evidence that has been obtained by comparing leap-year and non-leap-year Februaries suggests the variation is usually closer to $a_2 = 1.0 - a_1$ and $a_3 = 0.0$ than to $a_2 = 0.0$ and $a_3 = 1.0 - a_1$. The former is usually assumed and the equations reduce to those in the text. Such distinctions are academic in the seasonally adjusted series, since assuming either $a_2 = 0.0$ or $a_3 = 0.0$ affects only slightly the seasonally adjusted data for Februaries.

If one is interested in the trading-day adjusted data or in a study of the seasonal fluctuations, then the distinctions become important. For example, one could include the length-of-month variation in the seasonal pattern as did Kuznets (4). Or, if one wished to study seasonal patterns excluding length-of-month variation, it is necessary to attempt to estimate a_2 and a_3.

IV. RESULTS

A. Artificial Series

Tests with artificial series help us examine two questions: (1) What is the effect of the Census Bureau seasonal adjustment process upon known trading-day variation? and (2) How accurately can trading-day variation be estimated in the presence of various amounts of random variation?

To construct the artificial series, a set of monthly trading-day factors covering a period of 10 years were derived from daily weights in the manner described by equation 3. The following daily weights, which we shall refer to as "input weights," were used:

Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.	Total of weights
0.80	0.90	1.00	1.20	1.45	1.65	0.00	7.00

The set of monthly trading-day factors was taken as an artificial series. In addition, 18 artificial series were constructed by combining the monthly trading-day factors multiplicatively with 18 different rando[illegible] series designed to be normally distributed about a me[illegible] 1.00. The average month-to-month change (without [illegible] to sign) in the 18 random series ranged from 0.1 percent to 20.0 percent.

These artificial series resemble real economic series in two respects. First, they include a trading-day component that resembles the trading-day variation we might expect to find in some economic series. Second, the random components cover the range of variation most often found in the irregular components of economic series. In order to observe the effect of the seasonal adjustment process upon trading-day variation we did not introduce any seasonal or cyclical pattern, preferring to assume that they were constant. In this respect, the series do not resemble economic series and we may assume the test results to be somewhat better than if various seasonal and cyclical patterns had been introduced.

Three sets of daily weights were estimated from each artificial series. The first set was estimated directly from the artificial series and affords a basis for determining the effect of the seasonal adjustment process. Estimates corresponding with the first set could not be made from a real series which contained a seasonal component. The second set was estimated from the irregular components after seasonal adjustments of the artificial series were made with the X-9 version of Census Method II. If the seasonal adjustment process had no effect, the weights estimated before and after seasonal adjustment would be identical. The third set of weights was estimated from irregular components after iteration. Sometimes it has been considered necessary to iterate the estimation of the seasonal and trading-day variations in order to remove all the trading-day variation from the seasonal (see reference 5). If iteration improves the estimates, we would expect the third set of weights to correspond more closely to the input weights than do the second set. To obtain the third set,

Table 6.--DAILY WEIGHTS ESTIMATED FROM 10-YEAR ARTIFICIAL SERIES THAT CONTAIN KNOWN INPUT WEIGHTS AND A RANDOM COMPONENT

Series[1]	Average month-to-month percent change (without regard to sign) for--			Estimated daily weights for known input weights of--							Average deviation (without regard to sign) from input weights	Ratio of column 11			F-ratio[2]	Average standard error[2]
	Artificial series	Trading-day component	Random component	0.80 for Mon.	0.90 for Tues.	1.00 for Wed.	1.20 for Thurs.	1.45 for Fri.	1.65 for Sat.	0.00 for Sun.		b/a	c/b	c/a		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1[3].......a				0.80	0.90	1.00	1.20	1.46	1.65	0.00	.002				326,000	.00
b	3.54	3.54	0.00	0.88	0.88	1.03	1.17	1.32	1.64	0.07	.054	27.00	.37	10.00	209	.05
c				0.84	0.90	1.01	1.17	1.41	1.65	0.02	.020				10,400	.01
2........a				0.81	0.89	1.01	1.20	1.46	1.64	-0.01	.01				12,600	.01
b	3.56	3.54	0.11	0.89	0.87	1.04	1.17	1.33	1.65	0.07	.05	5.00	.40	2.00	206	.05
c				0.85	0.89	1.01	1.18	1.41	1.65	0.00	.02				7,000	.01
3........a				0.90	0.82	1.08	1.11	1.55	1.58	-0.04	.08				160	.07
b	3.71	3.54	0.81	0.95	0.80	1.10	1.09	1.44	1.56	0.07	.09	1.12	.78	.88	140	.06
c				0.93	0.80	1.09	1.10	1.48	1.60	0.00	.07				420	.04
4........a				0.88	0.97	0.80	1.42	1.33	1.80	-0.22	.15				25.2	.18
b	4.61	3.54	2.74	0.84	0.96	0.81	1.39	1.28	1.76	-0.05	.12	.80	1.08	.87	28.4	.15
c				0.82	0.96	0.75	1.45	1.40	1.78	-0.14	.13				46.6	.14
5........a				1.10	0.64	1.34	1.03	1.49	1.74	-0.34	.22				20.3	.20
b	4.97	3.54	3.04	1.01	0.68	1.39	0.94	1.45	1.72	-0.19	.19	.86	1.26	1.09	27.9	.16
c				0.97	0.64	1.44	0.92	1.56	1.75	-0.29	.24				39.9	.15
6........a				0.75	0.60	1.36	0.59	1.95	1.11	0.64	.43				4.62	.33
b	6.09	3.54	5.30	0.95	0.45	1.44	0.55	1.79	1.19	0.62	.44	1.02	1.05	1.07	6.46	.25
c				0.89	0.46	1.46	0.51	1.91	1.18	0.59	.46				8.22	.24
7........a				0.70	1.62	0.19	2.28	0.73	1.39	0.09	.54				1.52	.70
b	9.91	3.54	9.66	1.21	1.10	0.56	1.99	0.68	1.51	-0.06	.40	.74	1.12	.83	2.36	.45
c				1.25	1.10	0.50	2.14	0.67	1.52	-0.18	.45				3.09	.45
8........a				3.41	0.43	0.30	0.89	1.68	1.17	-0.88	.81				.77	1.27
b	21.93	3.54	22.07	2.77	1.07	0.84	0.18	2.12	0.55	-0.52	.80	.99	1.11	1.10	.96	1.04
c				2.97	1.10	0.69	0.17	2.20	0.52	-0.65	.89				1.17	1.04

[1]Weights for "a" estimated from artificial series. Weights for "b" estimated from irregular component of X-9 seasonal adjustment of artificial series. Weights for "c" estimated from irregular component after iteration in which artificial series was adjusted by "b" weights prior to second X-9 seasonal adjustment.
[2]See section IV, C, for description of tests of significance and for the levels above which the F- and t-ratios indicate significant variation.
[3]Series contains only trading-day variation.

the artificial series were first adjusted for trading-day variation with the second set. Then the series were adjusted with X-9 to obtain improved estimates of the seasonals and trend-cycle upon which to base the irregular, from which the weights were reestimated.

The estimated weights and the average absolute deviations from the input weights are shown in table 6 for 8 of the 19 series. (Also shown are F-ratios and standard errors which are discussed in section IV, C.) From table 6 the following conclusions are made:

1. In general, the seasonal adjustment process does not have much effect. The weights estimated before and after seasonal adjustment are similar.

2. What effect the seasonal adjustment process does have is generally concentrated on the series with small random variations. For these series, the weights estimated after seasonal adjustment deviate more from the input weights than do those estimated before seasonal adjustment. For these series, it is possible to improve the weights estimated after seasonal adjustment by iteration.

3. There are indications that, for series with larger random variation, better estimates are made after seasonal adjustment than before. This may indicate that for highly irregular series the seasonal adjustment process dampens the irregular variations more than the trading-day variations and thereby improves the prospect for estimating the trading-day variation. For these series iteration appears to make the estimates slightly worse. To determine whether it actually makes them worse might require a test with more series.

These conclusions are similar to Census Bureau experience with economic series. Experience has also suggested that trading-day variation can be usefully estimated from the irregular component and that iteration often does not yield substantial improvements.[11]

Now let us consider our second question: How accurately can trading-day variation be estimated in the presence of random variation? It is apparent from examining the estimated daily weights and their average deviations from the input weights that when there are moderate or large random fluctuations, the estimates are quite poor. When $\overline{I}$ (the average month-to-month change, without regard to sign, in the irregular component) is above 5 percent, the deviation of the estimates from the input weights is greater than the variation among the input weights that we are trying to estimate.

[11]The improvement achieved by iteration also depends upon the sensitivity to trading-day variation of the moving averages used to estimate the seasonal component and whether the identification of extreme values is improved by iteration. These two factors are not considered in detail in this paper.

It is useful to extend this approach and determine theoretically the levels of $\overline{I}$ for various lengths of series, time periods, and trading-day patterns above which the deviation of the estimates from the input or true weights is greater than the variation among the input weights. Table 7 shows such levels of $\overline{I}$ above which—

(Equation 5)

$$E\left[\sum_{1}^{7} (b_j - B_j)^2\right] > \sum_{1}^{7} (B_j - 1.0)^2 ,$$

where b_j's are the estimated weights;

B_j's are the true or input weights (where $\sum_{1}^{7} B_j = 7$).

The critical level of $\overline{I}$ for our input weights with a 10-year series for the period 1953-63 is 6.9 percent and for a 6-year series for the period 1958-63, 5.5 percent. For the 5-day-week weight pattern and the same periods, it is 8.8 percent and 7.1 percent.

The levels of $\overline{I}$ shown in table 7 can be taken as theoretical upper limits above which trading-day variation cannot be usefully estimated. These limits actually are not of much practical use, since the true trading-day pattern is not known when applying the method to a real series. They do suggest, however, that estimates made from highly irregular series cannot be expected to be useful.

In making a decision about whether a particular series should be trading-day adjusted, the F-test discussed in section IV C is more useful than the theoretical relationship shown in table 7. The F-test does not require that the true trading-day pattern be known. Basing decisions on the F-test at the 5 percent or 1 percent confidence levels usually leads to the same decisions as those suggested by the levels of $\overline{I}$ in table 7. Sometimes the F-test tends to suggest significant variation at levels of $\overline{I}$ somewhat above those shown in table 7.

Regardless of the approach taken to the question of how accurately trading-day variation can be estimated, one conclusion emerges: The method cannot estimate trading-day variation in highly irregular series.[12]

[12]This conclusion also holds for other methods of making trading-day adjustments. Consider the possible gain in adjusting a highly irregular series for trading days. Assume that we are able to estimate the trading-day variation exactly and that we then adjust for it. For example, series 3 in table 6, a series with little irregular, has an average month-to-month change before seasonal or trading-day adjustment of 3.71 percent (column 1). An exact, or perfect trading-day adjustment would reduce the average change to 0.81 percent, a 78 percent reduction (column 3). For series 7, a highly irregular series, the adjustment would yield a much smaller reduction. The average change is reduced from 9.91 percent to 9.66 percent, only a 3 percent reduction, illustrating that even if the trading-day pattern could be estimated exactly, very little gain is possible. The error associated with any alternative method of trading-day adjustment is probably almost as large as the possible gain.

Table 7.--LEVELS OF IRREGULAR VARIATIONS ABOVE WHICH DEVIATION OF ESTIMATES FROM TRUE WEIGHTS IS GREATER THAN DEVIATION OF TRUE WEIGHTS FROM 1.0

Known daily weights								Average month-to-month percent change without regard to sign					
								Irregular component			Combined irregular and trading-day component		
Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Variance	6-year series (1958-63)	10-year series (1953-62)	28-year series	6-year series (1958-63)	10-year series (1953-62)	28-year series
1.40	1.40	1.40	1.40	1.40	0.00	0.00	0.400	7.05	8.81	15.70	8.85	10.30	16.58
0.80	0.90	1.00	1.20	1.45	1.65	0.00	0.245	5.52	6.90	12.30	6.56	7.76	12.80
1.17	1.17	1.17	1.17	1.17	1.17	0.00	0.167	4.56	5.70	10.17	5.33	6.33	10.54
1.17	1.17	1.17	1.17	1.17	0.58	0.58	0.071	2.97	3.71	6.62	3.72	4.33	6.99

NOTE: See appendix B for the derivation of these measures.

B. Economic Series

Tests with economic series help us examine three related questions that cannot be readily answered with tests on artificial series: (1) How does the method compare with alternative techniques? (2) Do the estimates break down when applied to data outside the time period from which they were made? and (3) Is there evidence that the characteristics of trading-day variation change substantially over time or from one season to another and that our model which assumes no change, is inappropriate?

The criterion used to evaluate alternative trading-day adjustments of economic series is that the best method is the one resulting in the smallest month-to-month change, without regard to sign, in the irregular or unexplained variation (referred to as $\bar{I}$). It is not sufficient, however, to apply this criterion only for the "historical" period from which the estimates were made. A historical comparison is biased if the estimates of one or the other method explain, not only the trading-day variation, but part of the irregular variation. A more effective test is to apply the estimates made for the historical period to the current period, where methods which are too sensitive to the historical irregular fluctuations and those which inadequately allow for the characteristics of trading-day variation will both yield large fluctuations that are included in the computed $\bar{I}$.

Evaluation, therefore, consists of the following steps: (1) Estimate trading-day variation with each method from a historical period (For the regression method, estimates were made from irregular components of Method II, X-9 seasonal adjustments covering the historical period.) (2) make the trading-day adjustment to the historical and current data with the historical estimate; (3) obtain an irregular component by seasonally adjusting the combined historical and current data and compute $\bar{I}$ for each period (In each case the irregular component was from a Method II, X-9 seasonal adjustment of the combined historical and current data); and (4) compare the $\bar{I}$'s from the various methods giving particular attention to the current period.

Alternative trading-day adjustment methods are compared in table 8 for selected retail trade segments, bank debits, building permits and manufacturers' shipments and new orders for selected goods. A brief description of the results is as follows:

1. Retail sales.—For the period 1953-63, sales of eight retail kinds of business were adjusted for trading-day variation by regression estimates computed from the period 1953-61 and also 1957-61. These regression adjustments are compared with trading-day adjustments based upon average rates of sales on each day of the week computed from unpublished daily retail sales for 1962 that are available at the Bureau of the Census. They are also compared with series not adjusted for trading-days.

The regression estimates for 1953-61 yield the smallest $\bar{I}$'s, even for the current period of 1962-63 where we might expect the results to be biased in favor of the 1962 daily sales rates. The 1957-61 estimates are next best for both periods, followed by the daily sales rates and no adjustment.

Even though the unpublished estimates of the 1962 daily sales are not up to the Census Bureau publication standards,[13] the seven average daily rates are based on more evidence than is often available for an external adjustment and they appear to be reasonably close to what our experience would suggest as the daily sales pattern (see table 8). This comparison, therefore, supports the hypothesis that the customary external observation of the daily pattern of activity does not provide an adequate basis for a trading-day adjustment.

2. Bank debits.—For bank debits outside New York City, regression estimates computed from the period 1951-60 are compared with a 5-day week similar to the Federal Reserve adjustment used for part of this period. Also included are adjustments made with the Eisenpress method (3). By estimating a separate regression for each month, the Eisenpress method allows for seasonal characteristics in trading-day variation. It includes estimates for at least 12 coefficients (one for each month) and usually 24 (two for each month) or more, rather than seven. Two modifications of the original Eisenpress method are also compared. The first modification allows for a moving seasonal. The second, taking advantage of the evidence supplied by the Census Bureau regression, allows for the number of Mondays and Fridays as well as Saturdays and Sundays. The regression estimates yield smaller $\bar{I}$'s than the Eisenpress or 5-day week alternatives for both the current and historical period.

3. Building permits.—For "U.S. building permits," regression estimates computed from the period 1954-61 are compared with a 5-day week which might be selected a priori and with the series not adjusted for trading-days. The regression estimates and the 5-day-week adjustment yield approximately the same results for the historical period and for the current period of 1962-63. Both substantially reduce the irregular fluctuations found in the series not adjusted for trading days.

4. Manufacturers' shipments and new orders.—For manufacturers' shipments in four and new orders in five selected industries, regression estimates computed from the period 1953-61 are compared with two sets of weights that might be selected a priori: (1) A 5-day week and (2) weights where Saturday and Sunday receive partial weights. They are also compared with the series not adjusted for trading days. The a priori weights where Saturday and Sunday receive partial weights have been found to be appropriate for the aggregate series and thus might be expected to be appropriate for each component.

The regression estimates yield the smallest $\bar{I}$ in the historical period for seven of the nine series. In the current period, 1962-63, they are best for three series while the a priori weights where Saturday and Sunday receive partial weights are best for three series and the a priori 5-day week weights are best for three series.

Some of these manufacturing series contain larger irregular variations than do most of the other test series. Where the irregular variations are large, it is possible to discern the tendency for the differences between alternatives to be small relative to the magnitude of $\bar{I}$. The results are less conclusive, illustrating the fact that as irregular variations increase, the possible gains from trading-day adjustments decrease.

On the basis of these tests our conclusion is that the regression method performs quite well in comparison with other methods. There is little evidence that the estimates break down substantially in the current period or that changes in the characteristics of the trading-day variation over time or from one season to another invalidate our stable model.

C. Tests of Significance

The regression routine being added to Census Method II includes two tests which are useful guides in determining

---

[13] The 1962 daily sales were voluntarily reported by about a fourth of the 1,600 respondents in the weekly survey of stores with 10 or less outlets. The number of reports by kind of business is quite small, ranging from 2 for liquor stores and 8 for variety stores to 36 for lumber and building material dealers and 38 for eating and drinking places. The daily data, therefore, contains some inaccuracies. From the daily sales figures seven daily weights were computed by determining, for each week, the percentage of the week's sales occurring on each day and then averaging the daily

Table 8.--COMPARISON OF ALTERNATIVE TRADING-DAY ADJUSTMENTS FOR SELECTED BUSINESS ACTIVITIES

Series	Daily weights								Ī, average monthly change, without regard to sign, in irregular component				Rank (lowest Ī equals one)		F-ratio[1]	Average standard error
	Total	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	1953-61	1961	1962	1963	Historical 1953-61	Current 1962-63		
RETAIL SALES																
Eating and drinking places:																
Regression, 1953-61..........	7.00	0.95	1.04	0.97	0.91	1.12	1.10	0.92	.76	.52	.72	.78	2	2	2.25	.04
Regression, 1957-61..........	7.00	0.86	1.02	0.98	0.92	1.15	1.07	1.00	.74	.55	.53	.78	1	1	1.60	.04
1962 daily sales............	7.00	0.92	0.92	0.97	1.02	1.36	1.32	0.49	1.05	.94	.84	.99	4	4	14.42	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	.88	.79	.75	.95	3	3	9.23	x
Variety stores:																
Regression, 1953-61..........	7.00	0.96	0.96	0.97	1.01	1.15	1.59	0.35	1.54	.76	.83	.72	1	1	1.24	.10
Regression, 1957-61..........	7.00	0.90	1.11	0.94	0.85	1.30	1.56	0.34	1.55	.99	.83	.82	2	2	1.07	.12
1962 daily sales............	7.00	1.04	0.94	0.87	0.94	1.32	1.88	0.01	1.61	1.38	1.18	.88	3	3	2.14	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	2.61	1.96	1.75	1.58	4	4	18.07	x
Men's and boys' wear stores:																
Regression, 1953-61..........	7.00	0.79	1.02	1.13	0.70	1.59	1.43	0.34	2.27	2.65	2.82	2.70	1	2	.41	.14
Regression, 1957-61..........	7.00	0.93	0.75	1.18	0.97	1.45	1.09	0.62	2.33	1.87	3.04	3.36	2	4	2.11	.18
1962 daily sales............	7.00	1.04	0.97	0.94	1.15	1.36	1.53	0.01	2.46	2.87	2.21	2.39	3	1	.99	x
No adustment[2]...............	-	-	-	-	-	-	-	-	2.96	3.38	3.02	2.80	4	3	15.39	x
Shoe stores:																
Regression, 1953-61..........	7.00	0.78	1.07	0.91	0.74	1.57	1.82	0.11	2.23	2.33	2.86	3.05	1	3	.63	.14
Regression, 1957-61..........	7.00	0.42	1.30	1.15	0.48	1.86	1.27	0.52	2.44	2.08	2.77	3.08	2	2	2.82	.21
1962 daily sales............	7.00	0.95	0.87	0.76	0.91	1.28	2.19	0.04	2.45	2.26	2.99	3.35	3	4	4.28	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	3.32	1.52	1.39	3.20	4	1	26.91	x
Furniture, home furnishings stores:																
Regression, 1953-61..........	7.00	1.24	0.92	0.84	1.23	1.10	1.11	0.57	1.25	1.14	1.08	2.15	1	1	.48	.07
Regression, 1957-61..........	7.00	1.23	0.89	0.97	1.14	1.15	1.07	0.55	1.27	1.08	1.26	2.06	2	3	1.30	.09
1962 daily sales............	7.00	1.30	1.04	0.95	1.13	1.18	1.39	0.01	1.45	1.28	1.88	1.36	3	2	6.74	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	1.75	1.36	1.46	2.26	4	4	12.95	x
Lumber yards, and building materials dealers:																
Regression, 1953-61..........	7.00	1.13	1.11	1.16	1.22	1.15	0.67	0.56	1.69	1.32	.86	.88	1	1	.29	.10
Regression, 1957-61..........	7.00	1.00	0.99	1.40	1.11	1.24	0.64	0.63	1.77	1.28	1.01	1.09	2	2	1.15	.14
1962 daily sales............	7.00	1.29	1.26	1.16	1.22	1.19	0.87	0.01	1.81	1.74	1.76	1.10	3	3	3.56	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	2.86	2.94	2.15	2.46	4	4	19.79	x
Hardware stores:																
Regression, 1953-61..........	7.00	1.08	0.96	0.78	1.29	1.22	1.15	0.51	1.58	1.81	1.80	2.03	1	1	1.02	.10
Regression, 1957-61..........	7.00	0.96	0.84	0.96	1.19	1.44	0.90	0.71	1.62	1.73	2.24	2.20	2	3	2.14	.15
1962 daily sales............	7.00	1.11	1.04	1.10	1.05	1.22	1.48	0.00	2.22	2.47	2.24	2.12	3	2	5.02	x
No adjustment[2]..............	-	-	-	-	-	-	-	-	2.79	2.83	2.25	3.09	4	4	15.02	x
Liquor stores:																
Regression, 1953-61..........	7.00	0.87	0.96	0.75	1.04	1.32	1.53	0.52	1.15	.70	1.05	1.06	2	2	1.36	.07
Regression, 1957-61..........	7.00	0.78	1.05	0.79	0.90	1.46	1.53	0.50	1.11	.61	.88	.71	1	1	1.06	.08
1962 daily sales............	7.00	1.24	0.76	0.83	0.86	1.15	2.17	0.00	1.55	1.56	1.78	1.39	3	3	13.97	x
No adustment[2]...............	-	-	-	-	-	-	-	-	2.46	1.80	2.08	2.52	4	4	57.34	x
Average rank of retail sales; 8 components:																
Regression, 1953-61..........	x	x	x	x	x	x	x	x	x	x	x	x	1.2	1.6	x	x
Regression, 1957-61..........	x	x	x	x	x	x	x	x	x	x	x	x	1.8	2.2	x	x
1962 daily sales............	x	x	x	x	x	x	x	x	x	x	x	x	3.1	2.8	x	x
No adjustment	x	x	x	x	x	x	x	x	x	x	x	x	3.9	3.4	x	x

Table 8.--COMPARISON OF ALTERNATIVE TRADING DAY ADJUSTMENTS FOR SELECTED BUSINESS ACTIVITIES--Con.

Series	Daily weights: Total	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Ī, average monthly change, without regard to sign, in irregular component: base period	1961	1962	1963	Rank (lowest Ī equals one): Historical	Current	F-ratio[1]	Average standard error
									1951-60	1961	1962	1963	1951-60	1961-63		
BANK DEBITS																
Regression, 1951-60	7.00	1.52	1.14	1.24	1.24	1.38	0.19	0.29	.90	.56	.48	[3].62	1	1	1.56	.07
Regression, 1951-60[4]	7.00	1.53	1.06	1.31	1.19	1.46	0.14	0.32	.88	.60	.53	[3].72			2.16	.20
Eisenpress, 1951-60[5]	-	-	-	-	-	-	-	-	1.30	1.74	1.86	[3]2.54	4	5	3.77	x
Eisenpress[5] [6]	-	-	-	-	-	-	-	-	.95	1.02	.61	[3]1.62	3	4	5.95	x
Eisenpress[5] [7]	-	-	-	-	-	-	-	-	.90	1.12	.40	[3]1.12	2	3	2.77	x
A priori weights	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	1.39	.88	.72	[3].78	5	2	21.25	x
No adjustment[2]	-	-	-	-	-	-	-	-	3.71	3.39	3.02	[3]4.88	6	6	143.74	x
									1954-61		1962	1963	1954-61	1962-63		
BUILDING PERMITS																
Regression, 1954-61	7.00	1.20	1.11	1.48	1.25	1.60	0.32	0.69	3.17		2.55	3.28	2	1	.91	.17
A priori	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	3.13		3.81	3.62	1	2	2.42	x
No adjustment[2]	-	-	-	-	-	-	-	-	5.19		3.38	7.08	3	3	33.90	x
									1953-61		1962	1963	1953-61	1962-63		
MANUFACTURERS' SHIPMENTS AND NEW ORDERS																
Other petroleum products shipments:																
Regression, 1953-61	7.00	1.13	1.19	1.01	1.24	1.11	0.84	0.46	1.39		1.61	1.87	1.5	2	.36	.08
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	3.29		2.58	4.46	4	4	59.59	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	1.39		1.89	1.32	1.5	1	2.60	x
No adjustment[2]	-	-	-	-	-	-	-	-	2.42		3.18	2.81	3	3	29.89	x
Construction, mining, and material handling machinery shipments:																
Regression, 1953-61	7.00	0.95	1.53	0.93	0.95	1.46	0.64	0.54	2.63		2.00	1.17	1	2	.15	.16
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	3.95		2.66	4.32	4	4	13.80	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	2.86		1.84	1.21	2	1	1.72	x
No adjustment[2]	-	-	-	-	-	-	-	-	3.65		2.92	2.10	3	3	9.11	x
Tobacco shipments:																
Regression, 1953-61	7.00	1.89	0.96	1.42	1.18	1.24	0.00	0.31	1.54		1.28	1.89	1	1	2.63	.11
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	2.68		2.40	2.22	2	2	21.80	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	2.92		2.43	3.83	3	3	32.90	x
No adjustment[2]	-	-	-	-	-	-	-	-	4.35		3.68	6.23	4	4	76.05	x
Complete aircraft shipments:																
Regression, 1953-61	7.00	0.81	1.08	0.94	1.66	1.34	0.56	0.60	3.66		6.18	5.82	1	3	.80	.18
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	4.71		5.16	6.15	4	1	10.00	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	3.98		5.10	6.49	2	2	3.46	x
No adjustment[2]	-	-	-	-	-	-	-	-	4.69		6.07	6.48	3	4	14.77	x
Average rank of manufacturers shipments; 4 components, Ī >5.00:																
Regression, 1953-61	x	x	x	x	x	x	x	x	x		x	x	1.1	2.0	x	x
A priori (1)	x	x	x	x	x	x	x	x	x		x	x	3.5	2.8	x	x
A priori (2)	x	x	x	x	x	x	x	x	x		x	x	2.1	1.8	x	x
No adjustment	x	x	x	x	x	x	x	x	x		x	x	3.3	3.5	x	x

Metal working machinery, new orders:															
Regression, 1953-61	7.00	1.33	0.39	1.93	0.91	1.62	1.04	-0.21	9.29	10.99	13.58	1	2	.12	.49
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	9.32	11.65	12.75	2	1	3.59	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	9.42	10.24	16.08	3	3	2.03	x
No adjustment[2]	-	-	-	-	-	-	-	-	9.93	10.04	17.60	4	4	3.54	x
Motor vehicle parts, new orders:															
Regression, 1953-61	7.00	1.04	1.12	1.83	0.45	2.44	0.37	-0.26	10.84	5.84	4.88	2	1	.21	.48
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	11.07	6.67	5.85	4	3	3.46	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	10.87	6.38	5.76	3	2	2.57	x
No adjustment[2]	-	-	-	-	-	-	-	-	10.80	6.97	6.68	1	4	5.90	x
Blast furnaces, steel mills, new orders:															
Regression, 1953-61	7.00	0.52	1.62	1.21	1.44	1.13	0.72	0.35	10.27	10.43	12.69	1	2	.32	.43
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	10.79	12.66	9.37	3	1	2.31	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	10.72	10.90	12.31	2	3	.86	x
No adjustment[2]	-	-	-	-	-	-	-	-	11.54	11.63	12.19	4	4	4.06	x
Communication equipment, new orders:															
Regression, 1953-61	7.00	1.17	1.72	-0.46	2.10	1.43	1.40	-0.37	11.05	12.92	5.50	1	1	.24	.56
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	12.86	12.37	6.19	3	2	4.78	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	12.69	13.55	6.34	2	3	4.14	x
No adjustment[2]	-	-	-	-	-	-	-	-	13.00	15.23	7.41	4	4	3.53	x
Electrical industrial apparatus new orders:															
Regression, 1953-61	7.00	1.02	1.71	-0.87	3.27	0.84	0.40	0.63	9.54	5.97	5.02	1	2	1.28	.52
A priori (1)	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00	10.95	6.08	5.16	2	3	4.11	x
A priori (2)	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58	11.04	6.22	4.38	3	1	4.49	x
No adjustment[2]	-	-	-	-	-	-	-	-	11.42	7.54	4.65	4	4	5.17	x
Average rank of manufacturers' new orders; 5 components, $\bar{I} > 10.00$:															
Regression, 1953-61	x	x	x	x	x	x	x	x	x	x	x	1.2	1.6	x	x
A priori (1)	x	x	x	x	x	x	x	x	x	x	x	2.8	2.0	x	x
A priori (2)	x	x	x	x	x	x	x	x	x	x	x	2.6	2.4	x	x
No adjustment	x	x	x	x	x	x	x	x	x	x	x	3.4	4.0	x	x

x--Not applicable.

[1]The F-ratio shown for "No adjustment" was computed from the combined irregular-trading day component. It is a test for the presence of trading-day variation in the unadjusted data. F-ratios shown for the other adjustments were computed from irregular components after the series was adjusted with the daily weights. Thus, they are a test for residual trading-day variation, that is, for the adequacy of the daily weights. Section IV, C describes these tests of significance and shows levels above which these ratios indicate significant variation.

[2]"No adjustment" is shown without weights. An alternative would be to show a weight of 1.0 for each day of the week.

[3]Excludes the extreme value of November 1963. Although using the October to November and November to December changes results in higher $\bar{I}$'s for 1963, it does not affect the current rankings.

[4]No extremes deleted from regression in order to be comparable with Eisenpress regressions.

[5]The Eisenpress method does not yield 7 daily weights. A separate regression is made for each month.

[6]Modified to include a moving seasonal.

[7]Modified to include a moving seasonal and dummy variables of Saturday-Sunday combined and Monday-Friday combined where X_1 = 0, 1, or 2 depending on whether there are 8, 9, or 10 Saturdays and Sundays in a month and X_2 = 0 or 1 depending on whether there are 8 or 9 Mondays and Fridays in a month. This is in contrast to the variables used by Eisenpress where X_1 = 0 or 1 depending on whether there are 4 or 5 Saturdays a month and X_2 = 0 or 1 depending on whether there are 4 or 5 Sundays in a month.

whether trading-day variation is present in a series. One of these is a standard F-test to determine the significance of the regression and the other a t-test to determine if a daily weight differs significantly from a specified value. These tests are described more fully in appendix A.

The F-ratio and the average of the standard errors for the daily weights[14] are shown in tables 6 and 8 to provide a convenient reference when assessing results of other series. Under assumptions of normality, the F-ratio and t-ratio indicate significant variation when they are above the levels shown in table 9.

The assumptions of such tests are not always exactly fulfilled. In particular, the seasonal-adjustment process results in an irregular component which tends to be autocorrelated, even if the original series was an artificial series with a random irregular. Also, and perhaps more importantly, the adjustment process, by identifying values outside a specified sigma limit as extreme and excluding them from the computation of the daily weights, sometimes excludes some good values or does not succeed in excluding all extremes. Including or excluding a few values in the tail of the distribution can substantially affect the residual variance and the F- and t-ratios. For example, in table 8, values outside 2 sigma were designated as extreme for the retail sales series. If the limit had been 2.8 sigma instead, the F-ratios for "Eating and drinking places" would have been 1.63, .76, 9.45 and 8.34 rather than 2.25, 1.60, 14.42 and 9.23. In table 6, no extremes were deleted from the artificial series since the random values all belong to the same distribution. If a 2.8 or 2.0 sigma limit had been used, the F-ratio for series 6a would have been 6.02 or 8.05 rather than 4.62; for series 7a the F-ratio would have been 1.66 or 2.38 rather than 1.52.

In general, a choice of a sigma limit of about 2.8 appears reasonable in the new Method-II routine. In some instances, however, a limit of 2.0 or even less is required to identify the extremes. The analyst, therefore, should give close attention to the choice of the limit and to whether the assumption of a normal distribution provides valid test results.

In spite of such violations of the assumptions upon which the tests are based, the tests are of assistance in developing a good trading-day adjustment.

V. LIMITATIONS

A. Reliability

An important limitation of the regression and other methods is the amount of error associated with the estimates. This is discussed in section IV and will not be considered further.

[14] Since the standard errors of the 7 daily weights are approximately equal in practice, only their average is shown.

B. Variation Due to Holidays

Retail sales increase in the fall because of the occurrence of Christmas and, at the manufacturing level, dips occur in many series during July, because of the Fourth of July and plant-wide vacations. Most such variations due to holidays are accounted for by the seasonal factors but, in some instances, the effect of a holiday is not the same each year, varying with the calendar composition. In these cases, a residual variation, referred to simply as holiday variation, is present in the irregular component which, if it is not recognized and allowed for, can distort the trading-day estimates. Its estimation and removal can also reduce the irregular variation. Our method is limited in that it does not include an allowance for holiday variation. A sequential adjustment, however, can be made for holidays.

Now that a trading-day adjustment method has been computerized, it may be possible to study more thoroughly the relation of trading-day and holiday variation and develop techniques that allow for both variations simultaneously. Following is a summary of our experience to date:

1. Holiday variation is unimportant in many series that contain trading-day variation and also in most or all series that do not contain trading-day variation.[15] Where there is apparent holiday variation, it does not seem to seriously distort the estimates of trading-day variation, with the possible exception of Easter in series such as the retail sales of apparel.

2. The practice of assigning a zero weight to holidays when constructing trading-day adjustment factors is incorrect. The essential element is not whether the activity is shut down for the holiday, but whether the variation in the monthly series is related to the holiday.[16]

3. With the exception of Easter, major U.S. holidays are correlated with calendar composition in such a way that it is difficult to separate trading-day and holiday variation. For example, every time Christmas falls on Monday, December begins on Friday and contains five Fridays, Saturdays, and Sundays.

4. The effect of a holiday often occurs in 2 adjacent months. For example, when Labor Day is early, some "back to school" shopping occurs in August, but when Labor Day is late, such shopping occurs in September.

[15] One type of series where holiday variation can be a problem is that based upon a survey covering one week of the month. For example, the series on the average workweek can be affected by holidays that fall in the survey week.

[16] The Federal Reserve Board discontinued this practice for the Index of Industrial Production in 1953. "It is not always clear that holidays have an impact on output proportional to their number in the month, as was assumed under the old procedure. In some cases output 'lost' on account of holidays may be made up on contiguous days, particularly where the rate of purchase or consumption of the product is not influenced by the holiday. In other cases, as in connection with Christmas Day or July 4, output losses may be more than proportional to the 1 day of holiday time" (2, p. 1261). The current Federal Reserve practice is to make no allowance for holidays.

Table 9.--SIGNIFICANT F- AND t-RATIOS FOR SERIES OF VARIOUS LENGTHS

Length of series	Total degrees of freedom	Regression degrees of freedom	Residual degrees of freedom	F		t (2-tailed test)	
				5 Percent	1 Percent	5 Percent	1 Percent
5 years....................	60	6	54	2.27	3.15	2.00	2.66
6 years....................	72	6	66	2.24	3.09	2.00	2.66
7 years....................	84	6	78	2.21	3.04	1.98	2.62
8 years....................	96	6	90	2.19	2.99	1.98	2.62
10 years....................	120	6	114	2.17	2.95	1.98	2.62
12 years....................	144	6	138	2.16	2.92	1.96	2.58
16 years or more (n years)..	12n	6	12n-6	2.14	2.90	1.96	2.58

5. The only series seasonally adjusted by Census where significant holiday variation has been found are sales of certain types of retail business where adjustments are currently being made for Easter, Labor Day, and Thanksgiving-Christmas. For Easter a technique similar to the standard Easter adjustment is used, see references 5 and 8. A similar technique is used for Labor Day and Thanksgiving-Christmas. Essentially, it consists of arranging the irregular component, after adjustment for trading days, in an order that appears to bear a relationship to the date of the holidays and then fitting a smooth curve to estimate the relationship. In the case of Easter, the relationship between the date of the holiday and the variation in the data is obvious, the later Easter occurs, the higher are April sales and the lower are March sales and vice versa. In other instances, the relationship, if it exists, is not so obvious either a priori or in the data.

C. Changes Over Time

Our method of estimating trading-day variation makes no provision for changes in the characteristics of trading-day variation over a period of years. It is based upon the assumption that trading-day variation is a fairly constant, deep-seated phenomenon in the economy. However, there are several factors in the economy which can be assumed to cause changes in trading-day variation: (a) A half day of work or no work on Saturday is more common now than a few years ago; (b) fewer banks are open on Saturdays than before; (c) more retail stores are open evenings than in the past; (d) the amount of overtime worked on Saturdays, Sundays, and other days varies over the business cycle; (e) the introduction of electronic computers has changed bookkeeping practices in various industries and collection and processing of data in organizations such as the Bureau of the Census.

A relatively simple way of handling possible changes over time in trading-day variation is to restrict the analysis to a reasonably short period. At the Census Bureau, a period of about 8 or 10 years is usually used. (For the retail sales adjustments shown in table 8, better results were obtained for the years 1962-63 when estimates were made from 1953-61 rather than restricting the period to 1957-61. In this case, the gain from including 4 years, 1953-56 is larger than any loss arising from changes in the trading-day variation between 1953-56 and 1957-61.) It appears doubtful that including an explicit allowance for changes over time is necessary.[17]

In new Census Bureau seasonal-adjustment programs various options will be available to control the time period upon which the computations are based. For example, if the series covers 1948-64, the trading-day regression could be computed from say, 1957-64 and if the results are significantly different than previous results computed from say, 1953-60, the new estimates would be applied. Such options should be an adequate tool for allowing for changes over time when necessary.

[17] Canadian retail sales, however, shows some evidence of changes over time in trading-day variation, reported in reference 5.

D. Changes Within the Year

Our method of estimating trading-day variation assumes that the characteristics of trading-day variation are the same in each season of the year. In some instances, this assumption may not be entirely warranted. For example, many department stores have different hours of business in summer months than in the other months of the year.

Where there is a seasonal pattern in the trading-day variation it is possible to adapt the above method by separating some months from others and developing two or more sets of daily weights. Such procedures, though, are pretty much restricted by the lack of sufficient data to provide reliable estimates.

REFERENCES

1. Bongard, J., "Notes on the Characteristics of Seasonal Variation In Europe," Seasonal Adjustment on Electronic Computers, Organization for Economic Cooperation and Development, 1960, pp. 179-192.
2. Broida, Arthur L., et al, "Revised Federal Reserve Monthly Index of Industrial Production," Federal Reserve Bulletin, vol. 39, no. 12, December 1953, pp. 1260-1264.
3. Eisenpress, Harry, "Regression Techniques Applied to Seasonal Corrections and Adjustments for Calendar Shifts," Journal of the American Statistical Association, vol. 51, no. 276, December 1956, pp. 615-620.
4. Kuznets, Simon, Seasonal Variations in Industry and Trade, National Bureau of Economic Research, Inc., 1933, pp. 10, 25, New York.
5. Marris, Stephen N., "The Measurement of Calendar Variation," Seasonal Adjustment on Electronic Computers, Organization for Economic Cooperation and Development, 1960, pp. 345-360.
6. Method of Adjusting Time Series for Inequalities in the Number and Nature of Working Days, Research Paper No. 6, Organization for Economic Cooperation and Development, ESD/S/228, February 1961.
7. Shiskin, Julius, Electronic Computers and Business Indicators, Occasional Paper 57, National Bureau of Economic Research, Inc., 1957.
8. Thompson, Orville K., and Smith, Howard G., "Revised Indexes of Department Store Sales and Stocks," Federal Reserve Bulletin, vol. 37, no. 12, December 1951, pp. 1463-1515.
9. U.S. Bureau of the Census, Specifications for the X-9 version of Census Method II seasonal-adjustment program. Office of the Chief Economic Statistician, Bureau of the Census, Washington, D.C., 20233. (Specifications for the X-11 version that will include the trading-day routine will be made available.)
10. Young, Allan, "Census Trading-Day Adjustment Method," Business Cycle Developments, U.S. Bureau of the Census, May 1964, pp. 59-64. (The present paper covers the same ground and is a more detailed discussion.)

Appendix A. Mathematical Statement of Regression Method

The regression method of estimating daily weights rests on two fundamental assumptions:

(Assumption 1) All residual trading-day variation appears in the irregular; and

(Assumption 2) This variation may be expressed in terms of 7 daily weights.

The steps in the method are as follows:

Step 1.—Make a preliminary seasonal adjustment to obtain an irregular component, $\bar{I}$, which is assumed to contain all trading-day variation.

Step 2.—Delete extreme values in $\bar{I}$ which would tend to distort the estimated daily weights.[1] Call the resulting modified irregular I^*. In Census Method II, the mean of I^* is approximately 100.0.

Step 3.—Divide each I^* by 100, multiply by the number of days in that month, and subtract the number of days in the month so that the estimated daily weights will sum to zero.[2] Call this transformed series Y.

We assume that—

(Assumption 3) $Y = XB + E$,

where $Y = [Y_1\ Y_2 \ldots Y_n]'$ is the vector of the transformed irregular component and trading-day variation and n is the number of months included in the regression,

$E = [E_1\ E_2 \ldots E_n]'$ is the vector of the "true" irregular series,

$B = [B_1\ B_2 \ldots B_7]'$ is the vector of the daily weights to be estimated[3] and $\sum_1^7 B_j = 7$,

X is the matrix of independent variables with $X_{1i}, X_{2i}, \ldots, X_{7i}$ corresponding to the number of Mondays, Tuesdays, . . ., Sundays in a given month.

Step 4.—Modify the weights to sum to 0 for purposes of testing for the existence of trading-day variation. Since $\sum_1^7 B_j = 0$ by definition, $B_7 = 0 - \sum_1^6 B_j$. Define $B_1, \ldots, B_6$ as the Monday, . . ., Saturday weights and B_7 as the Sunday weight. Hence, $Y = X\,B + E$ becomes $Y = \hat{X}\,\hat{B} + E$,

where $\hat{B} = [B_1\ B_2 \ldots, B_6]'$

and $\hat{X}_{ji} = X_{ji} - X_{7i}$ $(j = 1, \ldots, 6;\ i = 1, \ldots, n)$.

Then

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \begin{bmatrix} \hat{X}_{11} & \hat{X}_{21} & \cdots & \hat{X}_{61} \\ \hat{X}_{12} & \hat{X}_{22} & \cdots & \hat{X}_{62} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \hat{X}_{1n} & \hat{X}_{2n} & \cdots & \hat{X}_{6n} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ \cdot \\ B_6 \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \\ \cdot \\ \cdot \\ \cdot \\ E_n \end{bmatrix}$$

or

$$Y_i = \sum_1^6 \hat{X}_{ji} B_j + E_i \qquad (i = 1, 2, \ldots n).$$

Step 5.—Assume:

(Assumption 4) $E(E) = O$, where O is an n-term vector of zeroes;

(Assumption 5) $V(E) = \sigma^2 I$, where I is an n x n identity matrix and σ^2 is the variance of the E_i;

(Assumption 6) $\hat{X}$ is fixed;

(Assumption 7) $\hat{X}$ has rank = 6.

Compute $\hat{b} = [b_1\ b_2 \ldots b_6]' = (\hat{X}'\hat{X})^{-1}\hat{X}'Y$, which is the least-squares estimate of $\hat{B}$. As an estimate of B_7, form $b_7 = 0 - \sum_1^6 b_j$.

Step 6.—Estimate the standard error of each weight as follows:

$$S_{b_j} = \left[\frac{(\hat{X}'\hat{X})^{-1}_{jj} \cdot e'e}{n-6}\right]^{\frac{1}{2}} \qquad (j = 1, 2, \ldots, 6),$$

where $e = [e_1\ e_2 \ldots, e_n]' = Y - \hat{X}\hat{b}$ is the vector of residuals from the fitted regression.

Now

$$\sigma^2_{b_7} = \sum_1^6 \sigma^2_{b_j} + 2 \sum_{\substack{1 \\ j \neq j'}}^6 \sum_1^6 \sigma_{b_j b_{j'}}.$$ As an estimate of $\sigma^2_{b_7}$, form

$$S^2_{b_7} = \frac{e'e}{n-6} \sum_1^6 \sum_1^6 (\hat{X}'\hat{X})^{-1}_{ji}.$$

Step 7.—To make inferences about the estimated weights b_j, assume that—

(Assumption 8) The E_i have a joint normal distribution.

To test whether b_j differs from 0, form $t_{b_j} = b_j/S_{b_j}$, which has a t-distribution with n-6 degrees of freedom. To test whether b_j differs from any specified k, form $t'_{b_j} = (b_j - k)/S_{b_j}$, which has the same distribution as t_{b_j}.

A test for the existence of trading-day variation in I^* may be made as follows. Form

$$F = \frac{\hat{b}'\hat{X}'\hat{X}\hat{b}/6}{e'e/(n-6)}.$$

which has an F-distribution with 6 and n-6 degrees of freedom. If this ratio is sufficiently low to conclude that the regression is not significant, we may conclude that there is no trading-day variation present.

To derive weights to be used in the seasonal program, add 1 to each b_j. If a trading-day adjustment is made prior to the seasonal adjustment, the b_j explain the residual trading-day variation. To combine the b_j with prior weights, use the formula $D_j = P_j + b_j$ where D_j are the final weights and P_j are the prior weights. Use the standard errors of the b_j as the standard errors of the D_j for inferences about the D_j.

[1]The specific procedure for identifying extreme values will be included in specifications for new Census Bureau seasonal programs. Essentially it consists of removing values where the residual computed in the regression falls outside a given sigma limit and then recomputing the regression. In section IV, C of the text it is suggested that sigma limits of 2.0 or 2.8 are often satisfactory.

[2]Februaries are handled in a special manner. The seasonal adjustment procedure eliminates variation due to the length of the month by, in effect, dividing each monthly value other than Februaries by 30 or 31. To reintroduce this length-of-month weighting, simply multiply each I^*_j by 30 or 31. Februaries are divided in the seasonal program by the average length of February for the period covered, which will be approximately 28.25. Hence, Februaries are multiplied here by 28.25. However, the actual number of days in the month (28 or 29) is subtracted from the result.

[3]The constant term in the regression is forced to equal zero.

Tests on artificial series suggest that assumptions 1 and 4-7 are not seriously violated. Assumption 2 may be violated when a single set of daily weights is estimated from a series where the trading-day pattern changes seasonally, cyclically or secularly.

Assumption 3 states that the residual trading-day and "true irregular" parts of the transformed irregular Y are related in an additive fashion ($Y = XB + E$). However, the daily weights estimated from the regression are combined into monthly calendar adjustment factors, which are used to remove trading-day variation in the unadjusted series in a multiplicative fashion. While it may seem inconsistent to estimate trading-day variation additively and apply it multiplicatively, this seems to be the best available method.

Purely additive and purely multiplicative alternatives were rejected because they do not allow for both the concept of daily weights and a multiplicative relationship of the "trading-day and irregular" component to the trend-cycle. The various alternatives are close approximations to each other, as illustrated by the fact that when two numbers X and Y are in the range of 0.95 to 1.05 (as most monthly calendar adjustment factors and irregulars are), the difference between $Z = X \cdot Y$ and $Z' = X + Y$ is negligible.

Assumption 8 is sometimes violated. The distribution of some irregular components may be skewed. Also, many economic series contain extreme values caused by strikes or unusual events which cannot be considered a part of the "true" irregular distribution. The process of identifying and removing these values may cause a truncation of one or both tails of the distribution, thus leaving a residual with a nonnormal distribution. Since the F-and t-tests are robust against nonnormality, these violations are not considered to be serious.

Standard Errors for Monthly Calendar Adjustment Factors

The regession method provides estimates of 7 daily weights $B_1, \ldots, B_7$ and estimates of their standard errors $\sigma_{B_1}, \ldots, \sigma_{B_7}$.

Monthly calendar adjustment factors are derived from the 7 daily weights as follows (assuming length-of-month variation is left in the seasonal):

$$\text{31-day months: } M_{31} = \frac{4\sum_{1}^{7} B_k + B_j + B_{j+1} + B_{j+2}}{31} = \frac{28 + B_j + B_{j+1} + B_{j+2}}{31},$$

where the month begins on day j and $B_{j+7} = B_j$;

$$\text{30-day months: } M_{30} = \frac{28 + B_j + B_{j+1}}{30};$$

$$\text{Leap-year Februaries: } M_{29} = \frac{28 + B_j}{28.25};$$

$$\text{Non-leap-year Februaries: } M_{28} = \frac{28}{28.25} = .991.$$

The standard errors of these factors are then:

$$\sigma^2_{M_{31}} = \left(\frac{1}{31}\right)^2 \left[\sigma^2_{B_j} + \sigma^2_{B_{j+1}} + \sigma^2_{B_{j+2}} + 2\left(\sigma_{B_j B_{j+1}} + \sigma_{B_j B_{j+2}} + \sigma_{B_{j+1} B_{j+2}}\right)\right];$$

$$\sigma_{M_{31}} = \frac{1}{31}\left[\sigma^2_{B_j} + \sigma^2_{B_{j+1}} + \sigma^2_{B_{j+2}} + 2\left(\sigma_{B_j B_{j+1}} + \sigma_{B_j B_{j+2}} + \sigma_{B_{j+1} B_{j+2}}\right)\right]^{\frac{1}{2}};$$

$$\sigma_{M_{30}} = \frac{1}{30}\left[\sigma^2_{B_j} + \sigma^2_{B_{j+1}} + 2\sigma_{B_j B_{j+1}}\right]^{\frac{1}{2}};$$

$$\sigma_{M_{29}} = \frac{1}{28.25}\sigma_{B_j};$$

$$\sigma_{M_{28}} = 0.$$

If length-of-month variation is included in the M's, the denominator of $\sigma_{M_{31}}$, $\sigma_{M_{30}}$ and $\sigma_{M_{29}}$ will be 30.4375.

Appendix B. Derivation of Relationship Between Trading-Day and Irregular Variations

Text table 6 gives levels of $\bar{I}$ above which a set of theoretical "true" trading-day weights with known variance cannot be reliably estimated from observations over a specified time period. These are the levels at which the expected value of the squared deviations between the estimated and true weights are greater than the squared deviations between the true weights and zero. In other words, the critical level is where

$$E\left[\sum_1^7 (b_j - B_j)^2\right] = \sum_1^7 B_j^2,$$

where $E(B_j) = 0$ and $E(b_j) = B_j \quad (j = 1, \ldots, 7)$.

Recalling assumptions 3 to 7 of appendix A,

(Assumption 3') $Y = XB + E$,

(Assumption 4') $E(E) = O$,

(Assumption 5') $V(E) = \sigma^2 I$,

(Assumption 6') X is fixed,[1]

(Assumption 7') X has rank = 7,[1]

proceed as follows:

$$b - B = (X'X)^{-1} X'Y - B$$
$$= (X'X)^{-1} X'(XB + E) - B$$
$$= (X'X)^{-1} X'E.$$

$$\text{Then } (b - B)^2 = (b - B)(b - B)'$$
$$= (X'X)^{-1} X'EE'X (X'X)^{-1}$$

$$\text{and } E\left[(b - B)^2\right] = \sigma^2 (X'X)^{-1}.$$

$$\text{Hence, } E\left[(b_j - B_j)^2\right] = \sigma^2 (X'X)_{jj}^{-1} \quad (j = 1, \ldots, 7)$$

$$\text{and } E\left[\sum_1^7 (b_j - B_j)^2\right] = \sigma^2 \sum_1^7 (X'X)_{jj}^{-1}.$$

Let n denote the number of months included in the regression and n_i the number of days in month i $(i = 1, \ldots, n)$.

Make the simplifying assumption:

(Assumption 8')

All $n_i = 30.4375$ $\left(\text{average length of month} = \frac{365\ 1/4}{12}\right)$.

Now $Y_i = n_i \left(\frac{I_i}{100} - 1\right)$,

where I_i is the "true" irregular and $E(I_i) = 100$ $(i = 1, \ldots, n)$.

$$\text{Then } \sigma^2 = \sigma_E^2 = \sigma_Y^2 = \frac{n_i^2}{(100)^2} \sigma_I^2$$
$$\doteq \frac{n_i^2}{(100)^2} \left(\frac{(100)^2 \cdot \pi}{4}\right) \bar{I}^2$$
$$= 727.625227\ \bar{I}^2.$$

(For proof that $\sigma_I^2 \doteq \frac{(100)^2 \cdot \pi}{4} \bar{I}^2$, see lemma at end of this appendix.)

Let $K_1 = \sum_1^7 B_j^2$.

$$\text{Now } E\left[\sum_1^7 (b_j - B_j)^2\right] = 727.625227 \sum_1^7 (X'X)_{jj}^{-1}\ \bar{I}^2$$
$$= K_2 \bar{I}^2.$$

[1] Assumptions 6' and 7' concerning X are analogous to assumptions 6 and 7 in appendix A concerning $\hat{X}$.

20

Then solve for $\bar{I} = \sqrt{\frac{K_1}{K_2}}$

Lemma. Proof of Approximation $\sigma_I^2 \doteq \frac{(100)^2 \cdot \pi}{4} \bar{I}^2$ (due to Harry Rosenblatt of Census):

$$\bar{I} = \frac{1}{n-1} \sum_{t=1}^{n-1} |\delta(I_t)|, \text{ where } \delta(I_t) = \frac{I_{t+1} - I_t}{I_t}.$$

Assume: (Assumption 9') To a satisfactory approximation, the $\delta(I_t)$ are independent, normal random variables, with zero mean and common variance σ_δ^2.

From assumption 9' it follows that $\bar{I}$ is a mean deviation with

$$\bar{I} = E|\delta(I)| = \sqrt{\frac{2}{\pi}}\, \sigma_\delta.$$

From assumption 5' (the I_t are independent with mean 100 and common variance σ_I^2) and a Taylor expansion approximation to the variance of the ratio represented by $\delta(I_t)$, it follows that

$$\sigma_\delta^2 \doteq \frac{\sigma_{I_{t+1}}^2}{\left[E(I_{t+1})\right]^2} + \frac{\sigma_{I_t}^2}{\left[E(I_t)\right]^2} - \frac{2\,\mathrm{Cov}(I_{t+1}\ I_t)}{E(I_{t+1})E(I_t)}$$
$$\doteq \frac{2\sigma_I^2}{(100)^2}.$$

Hence

$$\sigma_I^2 \doteq \frac{(100)^2}{2} \sigma_\delta^2 = \frac{(100)^2 \cdot \pi}{4} \left[E|\delta(I)|\right]^2,$$

and replacing $E|\delta(I)|$ by its estimate the mean deviation $\bar{I}$ results in the approximation $\sigma_I^2 \doteq \frac{(100)^2 \cdot \pi}{4} \bar{I}^2$.

Appendix C. Importance of Trading-Day Variation in Census Series

The relative contributions of the various components to the month-to-month variation in the unadjusted series, shown in text table 1, are derived from the relation $\bar{O}^2 = \overline{TD}^2 + \bar{E}^2 + \bar{S}^2 + \bar{C}^2 + \bar{I}^2$, where O, TD, . . ., I are the Method II summary measures of average absolute percent changes shown in table C1, (i.e., $\bar{X} = \frac{1}{n-1} \sum_{1}^{n-1} \left| \frac{X_{t+1} - X_t}{X_t} \right|$, where X = O, TD, E, S, C, I) given the assumptions below:

(Assumption 1) O = TD x E x S x C x I; where O, TD, . . ., I designate the original series and its components;

(Assumption 2) $E\left[\delta(X)\right] = $ (zero), where $\delta(X) = \left| \frac{X_{t+1} - X_t}{X_t} \right|$ and X = O, TD, E, S, C, I;

(Assumption 3) the $\delta(X)$ have the same distribution, where X = O, TD, E, S, C, I;

(Assumption 4) TD, E, S, C, and I are independent.

Assumption 2 is violated for O and C when the series contains a trend. However, the errors tend to be offsetting. Since the assumptions do not hold exactly, the following relation is used in order to force the results to add to exactly 100 percent. $\bar{O}^2 \doteq \bar{O}'^2 = \overline{TD}^2 + \bar{E}^2 + \bar{S}^2 + \bar{C}^2 + \bar{I}^2$. These formulations were developed by Bongard at the OECD (1) and Rosenblatt at Census.

Table C2 shows the daily weights from which the monthly trading-day components are derived.

Table C1.--SUMMARY MEASURES OF CENSUS SERIES

(Average month-to-month percent change, without regard to sign)

Census series	Unadjusted series $\bar{O}$	Trading-day component $\overline{TD}$	Holiday component $\bar{E}$	Seasonal component[2] $\bar{S}$	Trend-cycle component $\bar{C}$	Irregular component $\bar{I}$
Sales of retail business, 1953-63	7.50	1.96	0.32	6.75	0.43	0.65
Sales of wholesale business, 1960-63	5.56	3.52	....	3.55	0.38	0.87
Manufacturers' shipments, 1953-62[1]	5.27	2.24	....	4.24	0.79	1.18
Manufacturers' new orders, 1953-62[1]	5.10	2.24	....	4.04	1.00	1.93
U.S. exports, 1953-63	6.18	2.61	....	4.85	1.02	2.43
U.S. imports, 1953-63	7.00	5.20	....	4.53	0.91	2.32
New building permits, private housing units, 1954-62[1]	11.74	5.27	....	9.52	1.56	3.20

[1]Summary measures obtained from seasonal adjustment of aggregate series rather than from sum of seasonally adjusted components.

[2]Length-of-month variation is included in the seasonal component.

Table C2.--DAILY WEIGHTS FOR CENSUS TRADING-DAY ADJUSTMENTS

Census series	Total	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.
Sales of retail business, 1953-64:								
Nondurable goods[1]	7.00	0.87	0.92	0.93	0.98	1.43	1.38	0.49
Durable goods[1]	7.00	1.27	0.97	1.20	1.01	1.55	0.67	0.33
Sales of wholesale business, 1960-64[1]	7.00	0.96	1.26	1.48	1.16	1.21	0.53	0.40
Manufacturers' shipments, 1953-64	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58
Manufacturers' new orders, 1953-64	7.00	1.17	1.17	1.17	1.17	1.17	0.58	0.58
U.S. exports:								
1953-60	7.00	1.00	1.00	1.05	1.10	1.85	0.75	0.25
1961-64	7.00	0.75	1.10	1.00	1.05	1.70	0.90	0.50
U.S. imports, 1953-64	7.00	1.61	1.47	1.33	1.33	1.26	0.00	0.00
New building permits, private housing units, 1954-64	7.00	1.40	1.40	1.40	1.40	1.40	0.00	0.00

[1]Daily weights are weighted averages of weights used for individual kinds of business.

WORKING PAPER SERIES TECHNICAL PAPER SERIES

WORKING PAPERS

No. 1. Raw Materials in the United States Economy: 1900-52 (Preliminary)
No. 2. The Role of the 1954 Census of Manufactures in Overcoming Problems of Industry Data
No. 3. Tract Data Compared for a 25-Percent Sample and a Complete Census
No. 4. Sampling in the 1950 Census of Population and Housing
No. 5. Occupational Trends in the United States: 1900 to 1950
No. 6. Raw Materials in the United States Economy
No. 7. Papers Presented at the Census Tract Conference, December 29, 1958
No. 8. Materials on the Preparation and Conduct of the U.S.S.R. All-Union Population Census of 1959
No. 9. Historical Comparability of Census of Manufactures Industries: 1929-1958
No. 10. Papers Presented at the Census Tract Conference, December 29, 1959
No. 11. Papers Presented at the Census Tract Conference, August 25, 1960
No. 12. Papers Presented at the Census Tract Conference, August 28, 1961
No. 13. Papers Presented at the Census Tract Conference, September 8, 1962
No. 14. The Spectral Analysis of Economic Time Series
No. 15. Methodology and Scores of Socioeconomic Status
No. 16. Procedural Report on the 1960 Censuses of Population and Housing
No. 17. Papers Presented at the Census Tract Conference, September 5, 1963
No. 18. The Measurement of Performance Potential in Manufacturing Establishments
No. 19. Tests of Use of Post Office Resources to Improve Coverage of Censuses

TECHNICAL PAPERS

No. 1. The Sample Survey of Retail Stores: A Report on Methodology
... Description of the Sample for the Monthly Retail Trade Report (unnumbered revision)
No. 2. Accuracy of Census Statistics With and Without Sampling
No. 3. Farm Population, 1880-1950
No. 4. The Post-Enumeration Survey: 1950
No. 5. Tests and Revisions of Bureau of the Census Methods of Seasonal Adjustments
No. 6. The Current Population Survey Reinterview Program, Some Notes and Discussion
No. 7. The Current Population Survey, A Report on Methodology
No. 8. Trends in the Income of Families and Persons in the United States: 1947 to 1960
No. 9. Reconciliation of the 1958 Census of Retail Trade with the Monthly Retail Trade Report
No. 10. Population Trends in the United States: 1900 to 1960
No. 11. Response Errors in Collection of Expenditures Data by Household Interviews: An Experimental Study
No. 12. Estimating Trading-Day Variation in Monthly Economic Time Series

To receive further information or to order copies of these reports, contact the nearest field office of the Department of Commerce or write to—

BUREAU OF THE CENSUS
WASHINGTON, D.C. 20233

Modeling Time Series With Calendar Variation

W. R. BELL and S. C. HILLMER*

The modeling of time series data that include calendar variation is considered. Autocorrelation, trends, and seasonality are modeled by ARIMA models. Trading day variation and Easter holiday variation are modeled by regression-type models. The overall model is a sum of ARIMA and regression models. Methods of identification, estimation, inference, and diagnostic checking are discussed. The ideas are illustrated through actual examples.

KEY WORDS: Calendar variation; Trading day variation; Easter holiday variation; ARIMA models; Monthly time series.

1. INTRODUCTION

Suppose we observe a time series Z_t that follows the model (perhaps after transformation)

$$Z_t = f(\mathbf{X}_t;\xi) + N_t. \tag{1.1}$$

Here f is a function of ξ, a vector of parameters, and of $\mathbf{X}_t$, a vector of fixed independent variables observed at time t, and N_t is a noise series. If N_t is white noise, then (1.1) is the familiar linear or nonlinear regression model. However, when one deals with time series, N_t will generally be autocorrelated and frequently nonstationary. Numerous authors have warned against the consequences of using standard regression theory when N_t is autocorrelated, the problem being well established as long ago as Anderson (1954).

In this article we are concerned with the converse problem—that of the effects of ignoring $f(\mathbf{X}_t; \xi)$ when analyzing a time series. In the particular case we consider, $f(\mathbf{X}_t; \xi)$ represents trading day and holiday effects. For this case we illustrate the important points that (a) pure ARIMA models should not be applied blindly to all time series, (b) to ignore known, relevant independent variables is to invite difficulties, and (c) substantial improvements in models can be obtained when relevant independent variables are incorporated in the model.

* W. R. Bell is Mathematical Statistician, U.S. Census Bureau, Washington, DC 20233. S. C. Hillmer is Assistant Professor, School of Business, University of Kansas, Summerfield Hall, Lawrence, KS 66045. This research was partially supported while the authors were participants in the ASA-Census Research Fellowship Program. This program was funded by the U.S. Census Bureau and the National Science Foundation, and their support was instrumental in completing this research. The authors also acknowledge the many stimulating discussions with Census Bureau employees that helped shape the ideas of the research, and they wish to thank the referees and an associate editor for a number of helpful suggestions. The second author was partially supported by the Department of Commerce (Census Bureau) through JSA 81-2.

2. MODEL-BUILDING PROCEDURES

In developing models of the form (1.1) for a specific set of data we follow the three-stage model-building procedure of identification, estimation, and diagnostic checking presented in Box and Jenkins (1976). In (1.1) we assume that N_t follows the ARIMA model

$$\phi(B)\delta(B)N_t = \theta(B)a_t, \tag{2.1}$$

where B is the backshift operator ($BN_t = N_{t-1}$), $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ have all their zeros outside the unit circle, $\phi(B)$ and $\theta(B)$ have no common zeroes, $\delta(B)$ is a differencing operator (all zeroes on the unit circle) such as $(1 - B)$ or $(1 - B)(1 - B^{12})$, and $\{a_t\}$ is a sequence of independent, identically distributed (iid) random variables with mean 0 and variance σ^2. Some of the ϕ's and θ's may be 0 or otherwise constrained, so that (2.1) could be a multiplicative seasonal model.

2.1 Model Identification

The regression portion of the model, $f(\mathbf{X}_t; \xi)$, can be identified by consideration of the nature of the independent variables, which in our case are describing the trading day or holiday variation. To identify the noise model (2.1) we first examine the sample autocorrelation function (SACF) of the time series Z_t. In our experience with series containing trading day or holiday variation, examination of the SACF of Z_t is useful for determining the degree of differencing, $\delta(B)$, in N_t. We believe this is so because the effect of the nonstationary N_t on the computed sample autocorrelations dominates the effect of the trading day or holiday variation. In contrast, after Z_t (and thus N_t) has been appropriately differenced, the effect of the differenced N_t on the computed sample autocorrelations no longer dominates the effect of the differenced $f(\mathbf{X}_t; \xi)$. The SACF and sample partial autocorrelation function (SPACF) of the differenced Z_t series are usually confused. At this stage we must at least approximately remove the effects of $f(\mathbf{X}_t; \xi)$ from Z_t. To do this we fit the model

$$\delta(B)Z_t = \delta(B)f(\mathbf{X}_t; \xi) + e_t \tag{2.2}$$

by least squares regression (linear or nonlinear) and examine the SACF and SPACF of the residuals from this regression in order to tentatively identify the noise model. A justification for this procedure is that the sample au-

© Journal of the American Statistical Association
September 1983, Volume 78, Number 383
Applications Section

tocorrelations and hence the sample partial autocorrelations of the residuals from the least squares fit of (2.2) differ from those of $\delta(B)N_t$ by an amount that converges in probability to zero (see Fuller 1976, p. 399). This procedure is illustrated by two examples later in this article.

2.2 Model Estimation

Combining (1.1) and (2.1), we can write our model as

$$\delta(B)Z_t = \delta(B)f(\mathbf{X}_t; \xi) + \frac{\theta(B)}{\phi(B)} a_t. \quad (2.3)$$

We can then estimate ξ, ϕ, and θ, in (2.3) by maximum likelihood methods assuming normality of the a_t's We estimate σ^2 by $\hat{\sigma}^2 = (n - r)^{-1} \sum \hat{a}_t^2$ where n is the number of observations less the degree of $\delta(B)\phi(B)$, r is the number of parameters in (2.3), and

$$\hat{a}_t = \hat{\theta}(B)^{-1}\hat{\phi}(B)\delta(B)[Z_t - f(\mathbf{X}_t, \hat{\xi})].$$

Since the model for N_t is invertible this is asymptotically equivalent to nonlinear least squares.

Pierce (1971) discusses inference for the model (1.1) for the case in which $f(\mathbf{X}_t; \xi)$ is linear in ξ. He shows that under some conditions on the a_t's and the $\mathbf{X}_t$'s that the least squares estimates $\hat{v} = (\hat{\xi}, \hat{\phi}, \hat{\theta})$ are consistent and asymptotically normal, $\hat{\xi}$ is asymptotically independent of $(\hat{\phi}, \hat{\theta})$, and $\hat{\sigma}^2$ is a consistent estimator of σ^2. Also, the (i, j)th element of the inverse of the asymptotic covariance matrix of $\hat{v}$ can be approximated numerically by $-(\partial^2 L/\partial v_i \partial v_j) \mid \hat{v}$, where L is the log-likelihood. Hannan (1971) and Gallant and Goebel (1976) obtain results analogous to those of Pierce for the case in which $f(\mathbf{X}_t; \xi)$ is nonlinear in ξ, although they do not explicitly consider the asymptotic properties of $\hat{\phi}$ and $\hat{\theta}$. They require the additional assumptions of continuity of $f(\mathbf{X}_t; \xi)$ for the consistency of $\hat{\xi}$ (Hannan 1971) and twice differentiability for the asymptotic normality.

2.3 Diagnostic Checking

In general, the adequacy of both the assumed formulation of $f(\mathbf{X}_t;\xi)$ and the assumed noise model $\phi(B)\delta(B)N_t = \theta(B)a_t$ should be checked. To check the form of $f(\mathbf{X}_t; \xi)$ the residuals, $\hat{a}_t$, can be plotted against the X_{it} and any other possible independent variables. The $\hat{a}_t$ should be plotted against time to check for outliers, constancy of variance, and trends. The SACF of the residuals should be examined for any large autocorrelations. Ljung and Box (1978) show that under the hypothesis that the model is correct, for large n the statistic

$$Q = n(n + 2) \sum_{k=1}^{L} r_k(\hat{a})^2/(n - k)$$

has approximately a $\chi^2(L - s)$ distribution, where $r_k(\hat{a})$ is the lag k sample autocorrelation of $\hat{a}_t$, and s equals the number of parameters in the noise model. The noise model is judged inadequate if Q exceeds $\chi^2{}_\gamma(L - s)$ for some suitable γ.

3. TRADING DAY AND HOLIDAY VARIATION

The variation in a monthly time series that is due to the changing number of times each day of the week occurs in a month is called *trading day variation*. Trading day variation occurs when the activity of a business or industry varies with the days of the week so that the activity for a particular month partially depends on which days of the week occur five times. In addition, Young (1965) notes that accounting and reporting practices can create trading day effects in a time series. For example, stores that perform their bookkeeping activities on Fridays tend to report higher sales in months with five Fridays than in months with four Fridays. *Holiday variation* refers to fluctuations in economic activity due to changes from year to year in the composition of the calendar with respect to holidays. The primary example of this for U.S. economic series is the increased buying that takes place in some retail sales series just before Easter. This is a holiday effect since Easter falls on various dates in March and April. Holiday effects must be distinguished from seasonal effects, which are attributable to the same month every year. For instance, the increase in retail sales in December prior to Christmas each year is a seasonal effect and not a holiday effect.

Almost all of the previous research on trading day and holiday effects has dealt with their relation to seasonal adjustment. Young (1965) describes the procedures that are used in the Census X-11 seasonal adjustment program to adjust time series for trading day variation, and briefly discusses the adjustments made for holiday variation. Cleveland and Devlin (1980,1982) have reported on methods to identify times in which trading day effects are present in a time series and on methods to remove these effects. Pfefferman and Fisher (1980) discuss adjustments for both trading day and holiday variation. All of these authors use a two-stage approach in which a regression model is fitted to data that have been preprocessed to remove the trend and seasonality. We prefer to postulate a model of the form (1.1) and ARIMA noise structure and simultaneously estimate the regression and ARIMA parameters. Once a model of the form (1.1) has been developed, it can be used for a variety of purposes including forecasting and seasonal adjustment.

4. MODELING TRADING DAY VARIATION IN TIME SERIES

Trading day variation arises in part because the activity for a monthly time series varies with the days of the week. We assume that trading day effects can be approximated by a deterministic model. We deal only with flow series for which the data are the accumulation of the daily values (flows) over the calendar months. (Cleveland and Grupe (1982) discuss modeling of trading day effects for other types of series such as stock series, e.g., inventories.) If ξ_i, $i = 1, \ldots, 7$, represent the average rates of activity on Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday for the series being modeled (i.e., the daily effects), then the effect attributable to the number of times each day of the week occurs in month t is

$$TD_t = \sum_{i=1}^{7} \xi_i X_{it}, \tag{4.1}$$

where X_{it}, $i = 1, \ldots, 7$, are, respectively, the number of Mondays, Tuesdays, and so on in month t. A similar model was used by Young (1965), Cleveland and Devlin (1982), and Pfefferman and Fisher (1980). The model (4.1) accounts for variations in level due to differing month lengths, and allows for variations in level due to differing day of the week compositions for months of the same length. A model for the time series that incorporates trading day effects is

$$Z_t = TD_t + N_t, \tag{4.2}$$

where TD_t is as defined in (4.1) and N_t as in (2.1).

We obtain a useful reparameterization of (4.1) as follows. Let $\bar{\xi} = 1/7 \sum_{i=1}^{7} \xi_i$, $T_{it} = X_{it} - X_{7t}$, $i = 1, \ldots, 6$, and let $T_{7t} = \sum_{i=1}^{7} X_{it}$ denote the length of month t. Then we can write (4.1) as

$$TD_t = \sum_{i=1}^{7} (\xi_i - \bar{\xi})(X_{it} - X_{7t})$$
$$+ \sum_{i=1}^{7} (\xi_i - \bar{\xi})X_{7t} + \bar{\xi} \sum_{i=1}^{7} X_{it}$$
$$= \sum_{i=1}^{7} \beta_i T_{it}, \tag{4.3}$$

where $\beta_i = \xi_i - \bar{\xi}$, $i = 1, \ldots, 6$, and $\beta_7 = \bar{\xi}$. Our model then becomes

$$Z_t = \sum_{i=1}^{7} \beta_i T_{it} + \frac{\theta(B)}{\phi(B)\delta(B)} a_t. \tag{4.4}$$

We get the same estimate for TD, whether we use the parameterization (4.1) or (4.3); however, we have observed that estimates of the ξ_i's tend to be highly correlated while estimates of $\beta_1, \ldots, \beta_6$ are less so and are not highly correlated with the estimate of β_7. The parameters $\beta_i = \xi_i - \bar{\xi}$, $i = 1, \ldots, 6$, measure the differences between the Monday, Tuesday, $\ldots$, Saturday effects and the average of the daily effects, $\beta_7 = \bar{\xi}$. The difference between the Sunday effect and the average of the daily effects is then

$$\xi_7 - \bar{\xi} = \sum_{1}^{7} \xi_i - \bar{\xi} - \sum_{1}^{6} \xi_i$$
$$= 6\bar{\xi} - \sum_{1}^{6} (\beta_i + \bar{\xi}) = -\sum_{1}^{6} \beta_i,$$

and one may solve for the Sunday effect, ξ_7, using $\beta_7 - \sum_1^6 \beta_i$.

4.1 An Example

As an example, consider the series retail sales of lumber and building materials from January 1967 to September 1979, (the data for which may be obtained from the U.S. Census Bureau). Examination of a plot of the series reveals that the amplitude of the seasonality increases with the level. Therefore, we have determined that it is appropriate to model the natural logarithms, which we denote by Z_t. Examination of the SACF of the logged data and the SACF of the first differenced logged data indicated that first and twelfth differences are needed to achieve stationarity. The SACF of $(1 - B)(1 - B^{12})Z_t$ together with plus and minus two standard error limits are reported in Figure 1. Figure 1 does not exhibit a recognizable pattern. In order to identify the noise model we note that (2.2) for this example can be written

$$(1 - B)(1 - B^{12})Z_t = \sum_{1}^{7} \beta_i(1 - B)(1 - B^{12})T_{it} + e_t,$$

so we examine the SACF of the residuals from the regression of $(1 - B)(1 - B^{12})Z_t$ on $(1 - B)(1 - B^{12})T_{it}$ for $i = 1, \ldots, 7$. From this SACF, Figure 2, the presence of the large negative value at lag 12 suggests the noise

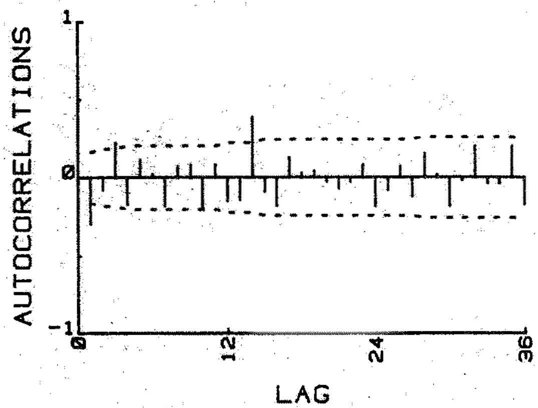


Figure 1. SACF of $(1 - B)(1 - B^{12})Z_t$.

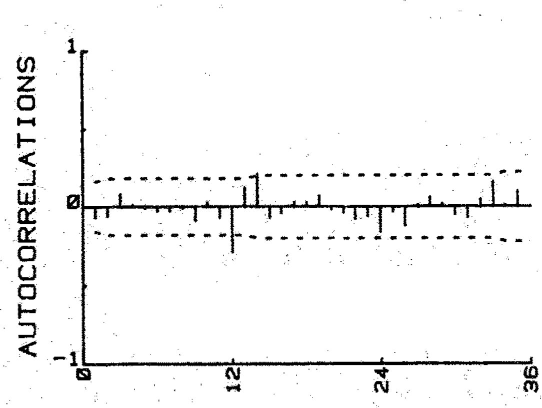


Figure 2. SACF of Regression Residuals.

model $(1 - B)(1 - B^{12})N_t = (1 - \theta_{12}B^{12})a_t$. Therefore, a tentatively entertained model for this series is

$$Z_t = \sum_{i=1}^{7} \beta_i T_{it} + \frac{(1 - \theta_{12}B^{12})}{(1 - B)(1 - B^{12})} a_t. \quad (4.5)$$

The BMDQ2T program (Liu 1979) was used to estimate the parameters in the model (4.5). The parameter estimates and corresponding standard errors are as follows:

$$\hat\beta_1 = \begin{matrix}.0055\\(.0042)\end{matrix} \quad \hat\beta_2 = \begin{matrix}.0068\\(.0041)\end{matrix} \quad \hat\beta_3 = \begin{matrix}.0017\\(.0042)\end{matrix}$$

$$\hat\beta_4 = \begin{matrix}.0103\\(.0041)\end{matrix} \quad \hat\beta_5 = \begin{matrix}.0061\\(.0041)\end{matrix} \quad \hat\beta_6 = \begin{matrix}-.0098\\(.0041)\end{matrix}$$

$$\hat\beta_7 = \begin{matrix}.037\\(.014)\end{matrix} \quad \hat\theta_{12} = \begin{matrix}.87\\(.028)\end{matrix} \quad \hat\sigma^2 = .00101.$$

The sample autocorrelations of the residuals from this model are all within plus or minus two standard errors of zero, and other diagnostic checks reveal no inadequacies with this model. The correlation matrix for the parameter estimates $\hat\beta_i$ $i = 1, \ldots, 7$ are reported in Table 1. The parameter estimates $\hat\beta_1, \ldots, \hat\beta_6$ are correlated so that individual inferences about these parameters must be made with caution. In contrast, $\hat\beta_7$ appears to be nearly uncorrelated with $\hat\beta_1, \ldots, \hat\beta_6$. This correlation pattern is typical of others that we have observed.

Inferences about the parameters in (4.5) can be made based on the asympotic theory referenced in Section 2. We first examine whether the daily effects (ξ_i) are different for the different days of the week by testing

$$H_0:\ \xi_1 = \cdots = \xi_7 \text{ vs.}$$

$$H_1:\ \text{not all } \xi_i \text{ are equal.}$$

This is equivalent to testing

$$H_0:\ \beta_1 = \cdots = \beta_6 = 0 \text{ vs.}$$

$$H_1:\ \text{not all } \beta_i = 0\ i = 1, \ldots, 6. \quad (4.6)$$

If **A** is the estimated covariance matrix of $\hat{\boldsymbol\beta} = (\hat\beta_1, \ldots, \hat\beta_6)'$, then under H_0 in (4.6) the asymptotic distribution of $\hat{\boldsymbol\beta}' \mathbf{A}^{-1}\hat{\boldsymbol\beta}$ is chi-squared with 6 degrees of freedom. Because $\hat{\boldsymbol\beta}' \mathbf{A}^{-1}\hat{\boldsymbol\beta} = 147.63$ is larger than 12.6, which is the .05 critical value of a chi-squared distribution with 6 degrees of freedom, we reject H_0 in (4.6) and conclude that the different days of the week have significantly different effects.

Table 1. Correlation Matrix of $\hat\beta$

	$\hat\beta_1$	$\hat\beta_2$	$\hat\beta_3$	$\hat\beta_4$	$\hat\beta_5$	$\hat\beta_6$	$\hat\beta_7$
$\hat\beta_1$	1.						
$\hat\beta_2$	−.54	1.					
$\hat\beta_3$	−.12	−.50	1.				
$\hat\beta_4$	.14	−.09	−.53	1.			
$\hat\beta_5$	.07	.14	−.12	−.51	1.		
$\hat\beta_6$	−.04	.05	.15	−.07	−.55	1.	
$\hat\beta_7$	.12	−.11	.13	−.14	.05	.09	1.

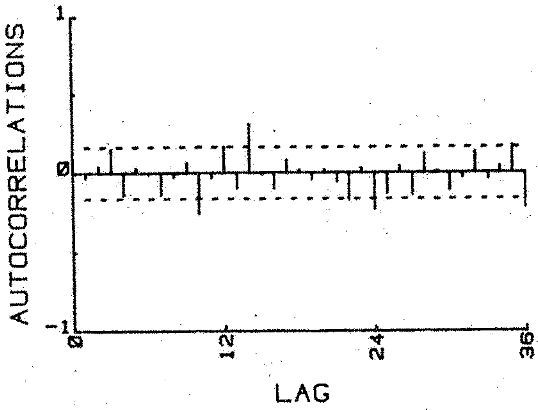


Figure 3. SACF of Residuals From (4.9).

It is also of interest to test

$$H_0:\ \beta_7 = 0 \text{ vs.}$$

$$H_1:\ \beta_7 \neq 0. \quad (4.7)$$

Since $\hat\beta_7$ divided by its standard error equals 2.6, we reject the null hypothesis in (4.7). When $\beta_7 \neq 0$, the term $\beta_7 T_{7t}$ in (4.5) accounts for an effect due to leap-year Februaries. To see this, notice from (4.5) that when we apply $1 - B^{12}$ to the data Z_t, we obtain

$$(1 - B^{12})Z_t = \sum_{i=1}^{7} \beta_i(1 - B^{12})T_{it} + \frac{(1 - \theta_{12}B^{12})}{1 - B} a_t.$$

Since T_{7t} equals the length of month t, $(1 - B^{12})T_{7t} = 0$ except in a leap-year February and the February of the following year.

4.2 Ignoring Trading Day Effects

From the preceding analysis it is clear that the model (4.5) is an adequate description of this time series. It is of interest to get an idea of the effect of ignoring the trading day variables in this particular example. With this idea in mind we estimated the model

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta_{12}B^{12})a_t. \quad (4.8)$$

Examination of the residual autocorrelations from (4.8) revealed a number of significant values, including a significant autocorrelation at lag one. We therefore tried the model

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B)(1 - \theta_{12}B^{12})a_t. \quad (4.9)$$

The parameter estimates for (4.9) are $\hat\theta_1 = .40$, $\hat\theta_{12} = .88$, and $\hat\sigma^2 = .0017$. The residual autocorrelations are plotted in Figure 3. By comparing the results of the fit for (4.5) with those of the fit for (4.9), we can judge the impact of the trading day effects upon this data set. While the residual autocorrelations from model (4.9) did not reveal any specific pattern, there are a number of moderately large sample autocorrelations. Futhermore, the value of the Ljung-Box Q based on 36 lags is 96.9, which greatly exceeds $\chi^2_{.01}(34) = 56.1$. We conclude that the residuals from (4.9) are not random. In contrast the model (4.5) passes the diagnostic checks and there is about a 40 percent reduction in the residual sum of squares from model (4.9) to (4.5). For this particular example the trading day effects are substantial and ignoring these effects is inappropriate.

5. MODELING HOLIDAY (EASTER) EFFECTS IN TIME SERIES

The Census Bureau adjusts certain retail sales series for holiday effects due to Easter, Labor Day, and Thanksgiving-Christmas (Young 1965). However, the Labor Day and Thanksgiving-Christmas adjustments are rather negligible, so we deal here only with modeling the effects of changing Easter dates. Techniques similar to those discussed here could be used to model other holiday effects, if necessary. For example, Liu (1980) discussed the problems involved with modeling a time series affected by the varying placement of the Chinese New Year.

The earliest and latest dates on which Easter can fall are March 22 and April 25. Thus, for series in which increased buying takes place before Easter we expect the March and April values in any particular year to depend on the date of Easter.

Specifying a functional form for the effect of Easter is not as simple as doing so for trading day effects. To be rather general, let $\tilde{\alpha}_i$ denote the effect on the series being modeled on the ith day before Easter; let $h(i, t)$ be 1 when the ith day before Easter falls in the month corresponding to time point t, and 0 otherwise. Then the Easter effect at t, E_t, is

$$E_t = \sum_{i=1}^{K} \tilde{\alpha}_i h(i, t), \qquad (5.1)$$

where K denotes some suitable upper bound on the length of the effect in days. Since many time series that contain Easter variation also contain trading day variation, we consider the model

$$Z_t = TD_t + E_t + N_t, \qquad (5.2)$$

where TD_t is given by (4.3), N_t by (2.1), and E_t by (5.1), although we will need to simplify E_t.

The relationship (5.1) was derived by consideration of the daily impact of Easter on the level of the series. Unfortunately, in most situations the only data available are monthly values of the series; as a consequence, in practice we cannot estimate effects as general as (5.1). To illustrate, consider the placement of Easter for the years 1967 to 1979. We chose these particular years because they correspond to the time frame of an actual set of data that is considered later; however, conclusions similar to those that we draw for these years are relevant for other time periods. Figure 4 shows the Easter dates for these years and constitutes the experimental design for determining the effect of Easter. From the diagram it is evident

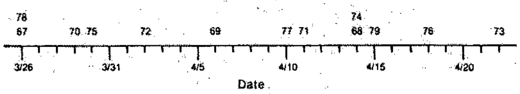


Figure 4. The Placement of Easter for 1967 to 1979.

that not all of the $\tilde{\alpha}_i$ in (5.1) can be estimated. For example, in these years whenever the fourth day before Easter fell in March so did the fifth day before Easter; otherwise, they both fell in April. Thus we cannot distinguish the effect of $\tilde{\alpha}_4 h(4, t)$ from that of $\tilde{\alpha}_5 h(5, t)$, using the data from 1967–1979. Since we cannot estimate all of the $\tilde{\alpha}_i$ in (5.1), we must look for special patterns in the $\tilde{\alpha}_i$.

We initially use the simple pattern $\tilde{\alpha}_1 = \cdots = \tilde{\alpha}_\tau = \tilde{\alpha}$, $\tilde{\alpha}_{\tau+1} = \cdots = \tilde{\alpha}_K = 0$ for some value τ. This implies

$$E_t = \alpha \cdot H(\tau, t), \qquad (5.3)$$

where $\alpha = \tilde{\alpha}\tau$ and $H(\tau, t) = 1/\tau \sum_{i=1}^{\tau} h(i, t)$. Given τ, $H(\tau, t)$ can be defined as the proportion of the time period τ days before Easter that falls in the month corresponding to time point t. With this definition $H(\tau, t)$ can be defined for any $\tau > 0$. For fixed t, $H(\tau, t)$ is in general a continuous but nondifferentiable function of τ. Figure 5 shows $H(\tau, t)$ for t corresponding to March 1969 and April 1969, Easter having been on April 6 that year.

Patterns other than that leading to (5.3) are possible. However, because Easter seldom occurred in early April from 1967 to 1979 (see Figure 4), it is unlikely that complex patterns can be detected from the data. This situation may change as additional data covering different Easter dates become available. We illustrate an approach to checking the adequacy of our assumed pattern in Section 5.3.

5.1 Noise Model Identification

It is of interest to consider the effect of E_t on the ACF of the original series and its differences. Figure 6 shows the SACF of $(1 - B)(1 - B^{12})H(14, t)$ (using January 1967 through September 1975 data), its most unusual features being the spikes at and near lags 36 and 48. Patterns in the SACF's for $H(\tau, t)$ for other τ and other time periods are similar. The degree to which these characteristics are transmitted to the original series depends on the magnitude of the Easter effect relative to TD_t and N_t. However, spikes at these lags can be taken as a possible indication of Easter effects in a series, especially when they show up in the SACF of a residual series from a model that has no terms to account for Easter effects.

To illustrate noise model identification, we consider the example of monthly retail sales of shoe stores (U.S.) from January 1967 through September 1979, which is available from the Census Bureau. (The observation for January 1970 ($t = 37$) was found to be an outlier and was modified from 243 to 270.3 (millions of dollars). The effect of the outlier was estimated by fitting the model with an indicator variable at $t = 37$.) We found it appropriate to analyze natural logarithms (denoted by Z_t) and to take $(1 - B)(1 - B^{12})Z_t$. Figure 7 gives the SACF of the differenced series, which exhibits behavior very similar to that in Figure 6, reflecting the Easter effect. To approximately remove E_t we choose a preliminary value of τ in (5.3), such as $\tau = 14$, and regress $(1 - B)(1 - B^{12})Z_t$ on $(1 - B)(1 - B^{12})H(14, t)$ and $(1 - B)(1 - B^{12})T_{it}$, $i = 1, \ldots, 7$. The SACF of the residuals from this regression, shown in Figure 8, does not show any influence of Easter or trading day effects. From this we identify a tentative noise model:

$$(1 - B)(1 - B^{12})N_t = (1 - \theta_1 B)(1 - \theta_{12}B^{12})a_t. \quad (5.4)$$

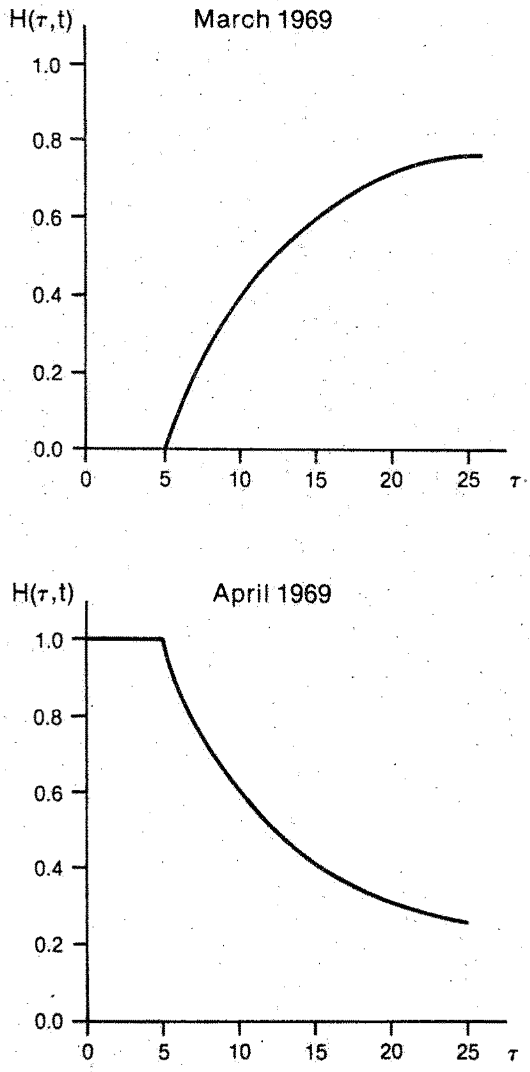


Figure 5. $H(\tau, t)$.

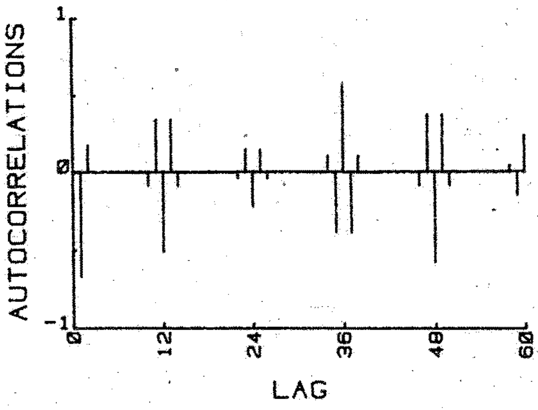


Figure 6. SACF of $(1 - B)(1 - B^{12})H(14, t)$.

5.2 Estimation of the Holiday Model

We demonstrate estimation of the model

$$Z_t = \sum_{i=1}^{7} \beta_i T_{it} + \alpha H(\tau, t) + \frac{\theta(B)}{\phi(B)\delta(B)} a_t \quad (5.5)$$

with the shoe stores example begun in Section (5.1). Notice that (5.5) is linear in $\beta_1, \ldots, \beta_7$, and α for fixed τ, so for fixed τ estimation may proceed in a manner analogous to that for the trading day model (4.4). We can obtain maximum likelihood estimators for the parameters of (5.5), including τ, by defining the asymptotic log-likelihood

$$L_{\max}(\tau) = \max_{\beta,\alpha,\phi,\theta,\sigma^2} L(\beta,\alpha,\tau,\phi,\theta,\sigma^2)$$

$$= -\frac{n}{2}\ln\hat{\sigma}^2(\tau) + \text{constant}$$

(where $\hat{\sigma}^2(\tau)$ is the estimate of σ^2 for fixed τ) and maximizing this over τ. Table 2 gives $\hat{\sigma}^2(\tau)$ for the shoe store

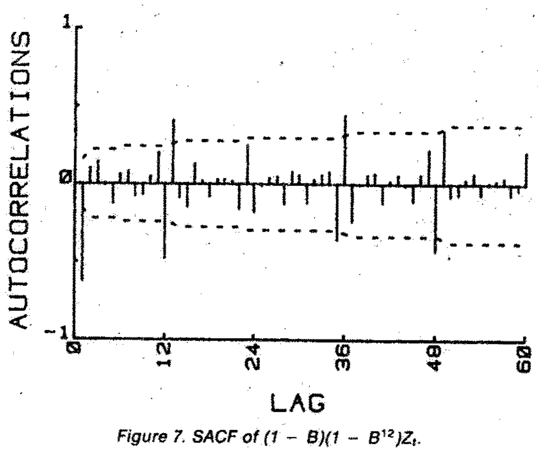


Figure 7. SACF of $(1 - B)(1 - B^{12})Z_t$.

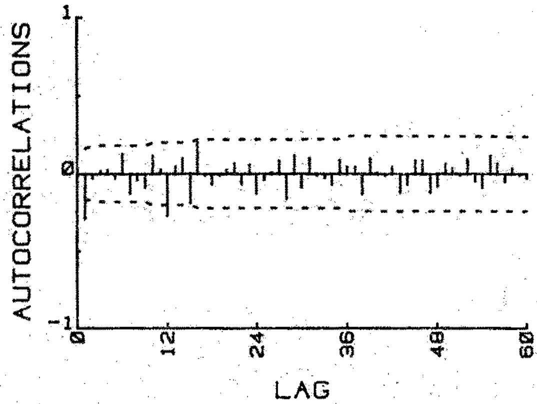


Figure 8. SACF of Regression Residuals.

series for $\tau = 1, \ldots, 25$. The MLE of τ is $\tau = 10$ (approximately) and the estimates of the other parameters together with their standard errors are (from the fit with $H(10, t)$ in the model) as follows:

$$\hat\beta_1 = \underset{(.0066)}{.0036} \quad \hat\beta_2 = \underset{(.0062)}{-.0024} \quad \hat\beta_3 = \underset{(.0063)}{-.0035}$$

$$\hat\beta_4 = \underset{(.0062)}{-.0087} \quad \hat\beta_5 = \underset{(.0063)}{.0213} \quad \hat\beta_6 = \underset{(.0065)}{.0131}$$

$$\hat\beta_7 = \underset{(.021)}{.045} \quad \hat\theta_1 = \underset{(.079)}{.32} \quad \hat\theta_{12} = \underset{(.029)}{.86}$$

$$\hat\alpha = \underset{(.012)}{.166} \quad \hat\sigma^2 = .00160.$$

It is of interest to note that for this series there is not much difference in the values of $\hat\sigma^2(\tau)$ for values of τ near 10.

Since $H(\tau, t)$ is continuous, $\hat\tau$ should be a consistent estimator of τ; however, since $H(\tau, t)$ is not differentiable for all t, the asymptotic normality need not hold. For fixed τ the results cited in Section 2.2 apply to the estimators of the other parameters, so that we can make inferences conditional on τ. If $\hat\tau$ were approximately independent of the estimators of the other parameters, then we could fix τ at 10 to make inferences. This can be checked by examining the parameter estimates and their standard errors for various τ. (In computing the standard errors, the parameter σ was estimated using the residual standard error for $\tau = 10$ because we considered this value to be a better estimator of σ than the residual standard error for other values of τ.) For this example the standard errors of $\hat\beta_1(\tau), \ldots, \hat\beta_7(\tau)$ are quite nearly constant for $\tau = 1, \ldots, 25$. Also, $\hat\beta_3(\tau), \ldots, \hat\beta_7(\tau)$ vary little for $\tau = 2, \ldots, 25$ and $\hat\beta_1(\tau)$ and $\hat\beta_2(\tau)$ vary little for $\tau = 7, \ldots, 25$. The estimates at the lower values of τ differ more from the others, although the differences are not large relative to the standard errors. The standard errors for $\hat\theta_1(\tau)$ and $\hat\theta_{12}(\tau)$ show little variation (no more than 10 percent) for $\tau = 2, \ldots, 25$ and $\tau = 2, \ldots, 16$ respectively, and are slightly lower outside these ranges. The estimates $\hat\theta_1(\tau)$ and $\hat\theta_{12}(\tau)$ vary little with τ, except possibly for $\tau = 1$. It seems for this series that $\hat\beta_1, \ldots, \hat\beta_7, \hat\theta_1, \hat\theta_{12}$, and their standard errors are relatively independent of $\hat\tau$, at least for a large part of the range of τ considered; thus we can make inferences on $\beta_1, \ldots, \beta_7, \theta_1$, and θ_{12} conditional on $\tau = 10$.

Table 2 shows that $\hat\alpha(\tau)$ and its estimated standard error depends more on τ. Still, we note that for this example there appears to be a fairly wide range of values of τ for which the estimates of α and their standard errors are fairly constant. These results are partially due to the experimental design given in Figure 4. Thus for data covering approximately the same years as this particular example, it may be reasonable to choose an approximate value for τ (for example, $\tau = 10$) and proceed with the inference conditional upon the value of τ chosen.

5.3 Checking the Easter Model

One way that the model (5.5) can be inadequate is if the Easter effect is more complex than that described by the simple pattern $\alpha_1 = \cdots = \alpha_\tau, \alpha_{\tau+1} = \cdots = \alpha_k = 0$. We cannot estimate all the α_i in (5.1), but can estimate a somewhat general pattern by grouping some of the terms in (5.1) together. We used the grouping

$$\begin{aligned} E_t = {} & \alpha_1[h(1, t) + h(2, t)] \\ & + \alpha_2[h(3, t) + \cdots + h(6, t)] \\ & + \alpha_3[h(7, t) + \cdots + h(10, t)] \\ & + \alpha_4[h(11, t) + \cdots + h(14, t)] \\ & + \alpha_5[h(15, t) + \cdots + h(18, t)] \\ & + \alpha_6[h(19, t) + \cdots + h(22, t)]. \end{aligned} \quad (5.6)$$

Table 2. Estimation of τ

τ	1	2	3	4	5	6	7	8	9	10	11	12	13
$100\hat\sigma^2(\tau)$	.212	.182	.177	.176	.175	.167	.163	.162	.161	.160	.161	.164	.167
$\hat\alpha(\tau)$	.13	.15	.15	.15	.15	.16	.16	.16	.16	.17	.17	.17	.17
$\hat\sigma(\hat\alpha(\tau))$	.0129	.0127	.0125	.0124	.0123	.0123	.0123	.0122	.0122	.0123	.0126	.0128	.0130

τ	14	15	16	17	18	19	20	21	22	23	24	25
$100\hat\sigma^2(\tau)$	.171	.171	.173	.175	.176	.178	.180	.183	.183	.183	.183	.183
$\hat\alpha(\tau)$	.17	.18	.18	.19	.19	.20	.21	.21	.22	.23	.24	.25
$\hat\sigma(\hat\alpha(\tau))$	.0135	.0141	.0146	.0151	.0157	.0164	.0169	.0176	.0183	.0192	.0200	.0209

Any grouping of the $h(i, t)$ can be used as long as it produces explanatory variables that are linearly independent over the span of the data. For our example we have 26 March and April observations for estimating the Easter effect. So as not to spread the observations too thin, we decided to use six groups, and chose the grouping in (5.6) to yield groups of equal length, except for a first group of length two to allow for a possibly important effect immediately before Easter.

Our general model at this point is (5.2) with TD_t given by (4.3), E_t by (5.6), and N_t by (2.1). We investigate how complex an Easter effect is needed by sequentially testing

$$H_0: \quad \alpha_j = \alpha_{j+1} = \cdots = \alpha_6 = 0 \text{ vs.}$$

$$H_1: \quad \alpha_j \neq 0, \alpha_{j+1} = \cdots = \alpha_6 = 0$$

for $j = 1, \ldots, 6$. When H_0 is rejected we can investigate whether a simple pattern of the form $\bar{\alpha}_1 = \cdots = \bar{\alpha}_\tau = \bar{\alpha}, \bar{\alpha}_{\tau+1} = \cdots = \bar{\alpha}_k = 0$ is adequate by testing (for $j > 1$)

$$H_0': \quad \alpha_1 = \cdots = \alpha_j, \alpha_{j+1} = \cdots = \alpha_6 = 0$$

against H_1. Table 3 presents (asymptotic) likelihood ratio test statistics for the shoe stores example computed as

$$\frac{[\text{RSS}(H_0) - \text{RSS}(H_1)]/\nu_1}{\text{RSS}(H_1)/\nu_2}$$

and similarly for H_0', where RSS denotes the residual sum of squares. The numerator degrees of freedom, ν_1, is 1 for testing H_0 and $j - 1$ for testing H_0'. The denominator degrees of freedom, ν_2, is 153 − 13 (for differencing) −1 (outlier) − 7 (TD parameters) −2 (θ_1 and θ_{12}) $-j = 130 - j$. The 5% and 1% critical values for the $F(\nu_1, \nu_2)$ distribution are also reported. The test statistics do not indicate that $\alpha_j \neq 0$ for $j > 3$. Also, there is no reason to reject the assumption that $\alpha_1 = \alpha_2 = \alpha_3$. We conclude that for this example the data give no evidence that the simplified Easter effect given by (5.3) is inadequate.

In addition to checking the Easter effect, we also should check the adequacy of the noise model (5.4). The sample autocorrelations of the residuals for the shoe stores series (using the model (5.5) with $\tau = \hat{\tau} = 10$) are all within plus or minus two standard errors of zero with the exception of $r_8(\hat{a})$, which is 2.7 standard errors below zero. The Ljung-Box Q statistic for 36 lags is 46.3. Since this is less than 48.6, the $\chi^2_{.05}(34)$ critical value, we conclude that the residuals appear to be white noise.

Table 3. Investigating $\alpha_1 = \cdots = \alpha_j, \alpha_{j+1} = \cdots \alpha_6 = 0$

j	F-statistic for H_0	F-statistic for H_0'	$F_{.05}(j-1, 130)$	$F_{.01}(j-1, 130)$
1	90.3	—	—	—
2	12.1	.9	3.9	6.8
3	5.7	.6	3.0	4.8
4	.0	—	—	—
5	2.7	—	—	—
6	.3	—	—	—

$F_{.05}(1,130) = 3.9 \quad F_{.01}(1,130) = 6.8$

NOTE: The $F(\nu_1, 130-j)$ critical values are very close to the $F(\nu_1, 130)$ critical values.

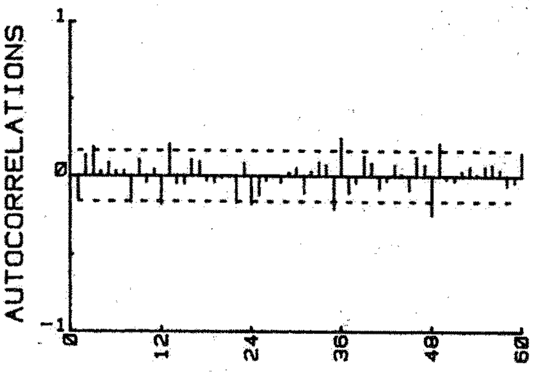


Figure 9. SACF of Residuals From (5.7).

5.4 Ignoring Trading Day and Easter Effects

As in the example of Section 4 it is of interest to investigate the influence of the trading day and Easter holiday terms in model (5.5) by fitting the model without these terms, which is

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B)(1 - \theta_{12}B^{12})a_t. \quad (5.7)$$

The parameter estimates for model (5.7) are $\hat{\theta}_1 = .68$, $\hat{\theta}_{12} = .93$, and $\hat{\sigma}^2 = .00333$. The residual autocorrelations are plotted in Figure 9. From Figure 9 there appear to be a number of moderately large $r_k(\hat{a})$'s at low lags, but there is not a recognizable pattern that would suggest a modification if a pure ARIMA model is to be used. Also, the behavior of the $r_k(\hat{a})$'s near lags 36 and 48 resembles that in Figure 6, indicating the presence of the Easter effect. The Ljung-Box Q statistic based upon 36 lags is 79.5, which is larger than $\chi^2_{.01}(34) = 56.1$. Thus, we would reject the hypothesis that the residuals from model (5.7) were white noise. The model (5.7) has obvious inadequacies, and there is about a 50 percent reduction in the residual sum of squares when the trading day and Easter influences are appropriately modeled.

6. CONCLUSIONS

In the time series literature the model (1.1) has been considered from a theoretical viewpoint; however, in many applications there has been an apparent tendency either to consider pure regression models or to consider pure ARIMA time series models. We have argued that there are situations in which a combination of these two models is superior. As particular examples we considered in detail the cases of time series that include trading day variation and Easter holiday variation. These two particular examples are important because there are many time series that contain one or both of these effects. The actual time series we considered indicate that substantial improvements over pure ARIMA models can be achieved if trading day and Easter effects are appropriately modeled. We hope that from this research more model builders will become aware of trading day and Easter variations and, as a result, will be in a better position to handle them.

[*Received March 1981. Revised February 1983.*]

REFERENCES

ANDERSON, R.L. (1954), "The Problem of Autocorrelation in Regression Analysis," *Journal of the American Statistical Association*, 49, 113–129.

BOX, G.E.P., and JENKINS, G.M. (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.

CLEVELAND, W.P., and GRUPE, M.R. (1982), "Modeling Times Series When Calendar Effects are Present," to appear in *Proceedings of the Conference on Applied Time Series Analysis of Economic Data*, ed. Arnold Zellner.

CLEVELAND, W.S., and DEVLIN, S.J. (1980), "Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods," *Journal of the American Statistical Association*, 75, 487–496.

——— (1982), "Calendar Effects in Monthly Time Series: Modeling and Adjustment," *Journal of the American Statistical Association*, 77, 520–528.

FULLER, W. (1976), *Introduction to Statistical Time Series*, New York: John Wiley and Sons.

GALLANT, A.R., and GOEBEL, J.J. (1976), "Non-linear Regression with Autocorrelation Errors," *Journal of the American Statistical Association*, 71, 961–967.

HANNAN, E.J. (1971), "Non-linear Time Series Regression," *Journal of Applied Probability*, 8, 767–780.

LIU, L.M. (1979), "User's Manual for BMDQ2T Time Series Analysis," BMDP Technical Report.

——— (1980), "Analysis of Time Series with Calendar Effects," *Management Science*, 26, 106–112.

LJUNG, G.M., and BOX, G.E.P. (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

PFEFFERMAN, D., and FISHER, J. (1980), "Festival and Working Days Prior Adjustments in Economic Time Series," in *Time Series*, ed. O.D. Anderson, New York: North-Holland.

PIERCE, D.A. (1971), "Least Squares Estimation in the Regression Model With Autoregressive-moving Average Errors," *Biometrika*, 58, 299–312.

YOUNG, A.H. (1965), "Estimating Trading-Day Variation in Monthly Economic Time Series," Technical Paper 12, Bureau of the Census.

O.D. Anderson & M.R. Perryman, eds., TIME SERIES ANALYSIS

A BAYESIAN APPROACH TO THE TRADING-DAY ADJUSTMENT OF MONTHLY DATA

Makio Ishiguro
Hirotugu Akaike

The Institute of Statistical Mathematics
Tokyo, Japan

The construction of a simple trading-day adjustment procedure developed for the Bayesian seasonal adjustment program BAYSEA is discussed. The practical utility of the procedure is demonstrated by numerical examples.

1. Introduction

Monthly economic data usually represents the monthly aggregate of an economic activity. Accordingly its spectral characterictic is controlled by that of the daily series. A series of monthly data may be approximated by a series which is obtained by first taking simple moving averages over (365x4+1)/48 days and then sampling at every (365x4+1)/48 day, where (365x4+1)/48 is the average number of days within a month. Thus the spectrum of monthly data is approximated by the spectrum of daily data deformed by the moving average filter and the sampling effect.

Actually, it is well known that the spectrum of monthly data shows a peak at the frequency 0.348 cycles/month (which is the alias of weekly cycles), provided that the original daily series contains a weekly component (Tukey, 1978). Conversely, if the spectrum of a monthly data shows a clear peak at frequency 0.348 cycles per month, this can be regarded as a strong evidence for the existence of the trading-day effect.

From (1978, p.28) suggested the desirability of a procedure which simultaneously, rather than sequentially, adjust for trading-day and seasonal variation. The Bayesian approach to the seasonal adjustment proposed by Akaike (1979) leads to a regression type procedure which can easily realize this suggestion. In this paper we will discuss the construction of a trading-day adjustment procedure developed for the Bayesian seasonal adjustment program BAYSEA (Akaike and Ishiguro, 1980) and demonstrate its practical use by examples.

2. Numerical examples of trading-day effect

We expect that department store sales data will exhibit a strong trading-day effect. Fig. 1a,b,c and d are log-transformed department store sales data of Japan, its trend, seasonal and irregular component, respectively. The decomposition into the three components is realized by the program BAYSEA. The power spectrum of the irregular component illustrated in Fig. 1e, obtained by fitting an ARMA model. shows a sharp peak at the frequency 0.348 cycle per month, the fundamental frequency expected for the trading-day effect.

The second example is concerned with the retail sales data of U.S.A. Fig. 2a, b, c, and e give the results corresponding to those given in Fig. 1.

The results obtained by applying the additive x-11 variant of the Census Method II without trading-day adjustment to the same data are illustrated in Fig. 3. The trading-day effect is also apparent in the power spectrum of the irregular component.

3. Modification of BAYSEA for trading-day adjustment

A. Original basic model for BAYSEA.
The basic model of the program BAYSEA is a slightly modified version of the model discussed by Akaike (1979). We assume that ($y(i)$, $i=1,2,\ldots,N$) admits the decomposition

$$y(i) = T(i) + S(i) + I(i), \qquad (1)$$

where $T(i)$ and $S(i)$ are the trend and seasonal component, respectively, and $I(i)$'s are independently identically distributed as Gaussian with mean zero and variance σ^2. The likelihood of the model (1) is then given by

$$f(y|a) = \left(\frac{1}{\sqrt{2\pi}\ \sigma}\right)^N \exp[-\frac{1}{2\sigma^2}\ \| y - xa \|^2], \qquad (2)$$

where

$$y = (y(1),y(2),\ldots,y(N))',$$

$$a = (T(1),T(2),\ldots,T(N),S(1),S(2),\ldots,S(N))'$$

and

$$x = \begin{pmatrix} 1 & & 0 & 1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & 1 & 0 & & 1 \end{pmatrix} \qquad (3)$$

(each block $N \times N$)

We also assume that the parameter vector a has a prior distribution with probability density function defined by

$$\pi(a|d) = (\frac{1}{2\pi})^N\ (\frac{1}{\sigma})^{2N}\ |D'(d)D(d)|^{\frac{1}{2}}\ \exp[-\frac{1}{2\sigma^2}\| D(d)(a - a_0)\|^2] \qquad (4)$$

where, $|\cdot|$, $'$ and $\|\cdot\|$ denotes determinant, transpose and Euclidian norm, respectively. a_0 is a suitably chosen initial guess of a, and $D(d) = dD$ where D is a 3Nx2N matrix defined by

$$D = \begin{pmatrix} D_k & 0 \\ 0 & wE_\ell \\ 0 & uF_M \end{pmatrix} \qquad (5)$$

(each block $N \times N$)

and d is a hyperparameter. Here D_k is introduced to control the smoothness of the k-th order differences of the trend component and is given by

$$D_1 = \begin{bmatrix} 1 & & & 0 \\ -1 & 1 & & \\ & \ddots & \ddots & \\ 0 & & -1 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 & & & & 0 \\ -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ 0 & & 1 & -2 & 1 \end{bmatrix}, \quad \text{etc.}$$

Similarly, to control the smoothness of the seasonal components, E_ℓ is defined as one of the matrices

$$E_1 = \begin{bmatrix} I & & & 0 \\ -I & I & & \\ & \ddots & \ddots & \\ 0 & & -I & I \end{bmatrix}, \quad E_2 = \begin{bmatrix} I & & & & 0 \\ -2I & I & & & \\ I & -2I & I & & \\ & \ddots & \ddots & \ddots & \\ 0 & & I & -2I & I \end{bmatrix}, \quad \text{etc.},$$

where I is an MxM identity matrix and M is the number of seasons in a period or a year. To keep the sum of the seasonal components within a year close to zero F_M is defined by

$$F_M = \begin{bmatrix} 1 & & & & & & 0 \\ 1 & 1 & & & & & \\ \vdots & & \ddots & & & & \\ 1 & \cdots & 1 & 1 & & & \\ & \ddots & & \ddots & \ddots & & \\ 0 & & 1 & & \cdots & 1 & 1 \end{bmatrix} \quad (\text{first } M \text{ rows marked } \updownarrow M)$$

w and u are suitably chosen constants. When the hyperparameter d is known, the posterior distribution of the parameter a is given by

$$\pi(a|y,d) \propto \exp[-\frac{1}{2\sigma^2}(a - a_*)'\tilde{X}'\tilde{X}(a - a_*)],$$

where

$$\tilde{X} = \begin{bmatrix} X \\ D(d) \end{bmatrix}$$

and a_* is the posterior mean of a. a_* is the value of a which minimizes $\| z(a|d) \|^2$, where $z(a|d)$ is defined by

$$z(a|d) = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \\ dc(1) \\ \vdots \\ dc(L) \end{bmatrix} - \begin{bmatrix} X \\ \\ D(d) \end{bmatrix} \begin{bmatrix} a(1) \\ \vdots \\ a(2N) \end{bmatrix} \qquad (6)$$

where L=3N and a(i) and c(i) denote the i-th element of the vector a and $D(d)a_0$ respectively. Our estimates of the trend and seasoned components are defined by the posterior mean

$$a_* = (T_*(1), T_*(2), \ldots, T_*(N), S_*(1), S_*(2), \ldots, S_*(N))'$$

The choice of the hyperparameter d is realized by maximizing the likelihood of the model, or by minimizing the Bayesian information criterion ABIC(d) defined by

$$\begin{aligned} \mathrm{ABIC}(d) &= -2 \log \text{(marginal likelihood)} \\ &= -2 \log \int f(y|a)\, \pi(a|d)\, da \\ &= N \log \left[\frac{1}{N} \| z(a_*|d) \|^2 \right] \\ &\quad + \log |D'(d)D(d) + X'X| - \log |D'(d)D(d)|. \end{aligned}$$

The choice of the matrices D_k, E_ℓ and the constants w, u can also be realized by minimizing ABIC.

B. Modification for trading-day adjustment.

Instead of the decomposition (1), we assume the decomposition of the form

$$y(i) = T(i) + S(i) + \sum_{j=1}^{7} d_j(i) w_j + I(i), \qquad (1)'$$

where $d_j(i)$ denotes the number of j-th day-of-week within the i-th month and w_j ($j=1,\ldots,7$) are respectively the weights for each day-of-week to be estimated under the restriction

$$\sum_{j=1}^{7} w_j = 0. \qquad (7)$$

This condition makes the term $\Sigma\, d_j(i) w_j$ sensitive only to the variation between $d_j(i)$'s ($j=1,2,\ldots,7$). The modified model (1)' contains six more parameters than the original model (1). The prior distribution of the expanded parameter vector $a = (T(1),\ldots,T(N),S(1),\ldots,S(N),w_1,\ldots,w_6)$ is defined by (4) with the modifications

$$D(d) = \begin{bmatrix} dD_k & 0 & 0 \\ 0 & dwE & 0 \\ 0 & duF & 0 \\ 0 & 0 & vI_6 \end{bmatrix} \qquad (5)'$$

where I_6 is the 6x6 identity matrix. The estimate a^* of a is obtained by minimizing $\| z(a|d) \|^2$, where $z(a|d)$ is given by

$$z(a|d) = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \\ dc(1) \\ \vdots \\ dc(3N) \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} X \\ \\ D(d) \end{bmatrix} \begin{bmatrix} T(1) \\ \vdots \\ T(N) \\ S(1) \\ \vdots \\ S(N) \\ w_1 \\ \vdots \\ w_6 \end{bmatrix}$$

where X is an $N \times (2N+6)$ matrix defined by

$$X = \begin{bmatrix} 1 & & 0 & 1 & & 0 & d_1(1) - d_7(1), \ldots, d_6(1) - d_7(1) \\ & \ddots & & & \ddots & & \\ 0 & & 1 & 0 & & 1 & d_1(N) - d_7(N), \ldots, d_6(N) - d_7(N) \end{bmatrix} \qquad (3)'$$

Thus the trading-day adjustment is realized simply by modifying the matrices D(d) and X.

4. Performance of the modified BAYSEA

The modified BAYSEA was applied to two sets of data treated in Section 2. The values of the hyperparameter d were selected automatically by BAYSEA. Several sets of values of the hyperparameters w, u and v were tried for each case and the one which gave the minimum of ABIC was chosen. The intital selection of the set of possible values of the parameters can easily be realized by watching the spectrum of the irregular component for each choice of the parameter values. The spectra of the irregular components of both data are shown in Fig.'s 4a and 5. The peaks at the frequency 0.348 cycle per month are now completely eliminated for both cases. These results clearly demonstrate the feasibility of the trading-day adjustment by the Bayesian approach realized by BAYSEA.

5. Further modification for leap year effect

Here we consider further modification of BAYSEA for leap year type effect which is due to the change of the numbers of effective trading days within each month. The necessary modification for this effect is realized as follows: First, the condition (7) is ignored and vI_6 of (5)' is replaced by

$$v\begin{bmatrix} I_7 - (1/7)ee' \\ \delta e' \end{bmatrix}$$

where I_7 denotes a 7x7 identity matrix and e denotes the seven dimensional vector defined by $e = (1,1,\ldots,1)'$, and where δ is a suitably chosen constant.
The parameter vector a is further augumented by w_7 and $d_j(i) - d_7(i)$ of (3)' is replaced by $d_j(i) - \bar{d}$ where $\bar{d}$ is equal to 4.348, the average number of a day of week in a month. Obviously (3)' is augumented by the column composed of $d_7(i) - \bar{d}$.

With this modification the Bayesian model can handle both the trading-day and leap year effect. This fact was confirmed by numerical experiments.

This subsection was added after the time of The International Time Series Meeting.

Acknowledgement

The research reported in this paper was partly supported by a grant from the Asahi Shinbun Publishing Company.

References

[1] Akaike, H. (1979) Likelihood and the Bayes procedure. A paper presented at the International Meeting on Bayesian Statistics, May 28 - June 2, 1979, Valencia, Spain. To be published in Trabajos de Estadistica.

[2] Akaike, H. and Ishiguro, M. (1980) BAYSEA, A Bayesian seasonal adjustment program, Computer Science Monographs, No.13, The Institute of Statistical Mathematics, Tokyo.

[3] Fromm, G. A. (1978) Comments on "An overview of the objective and framework of seasonal adjustment" by Shirley Kallek. In Arnold Zellner (ed.), Seasonal Analysis of Economic Time Series, Economic Research Report, ER-1, U.S. Department of Commerce, Bureau of the Census, 26-29.

[4] Tukey, J. W. (1978) Comments on "Seasonality: Causation, Interpretation and Implications" by Clive W. J. Granger. In Arnold Zellner (ed.), Seasonal Analysis of Economic Time Series, Economic Research Report, ER-1, U.S. Department of Commerce, Bureau of the Census, 50-53.

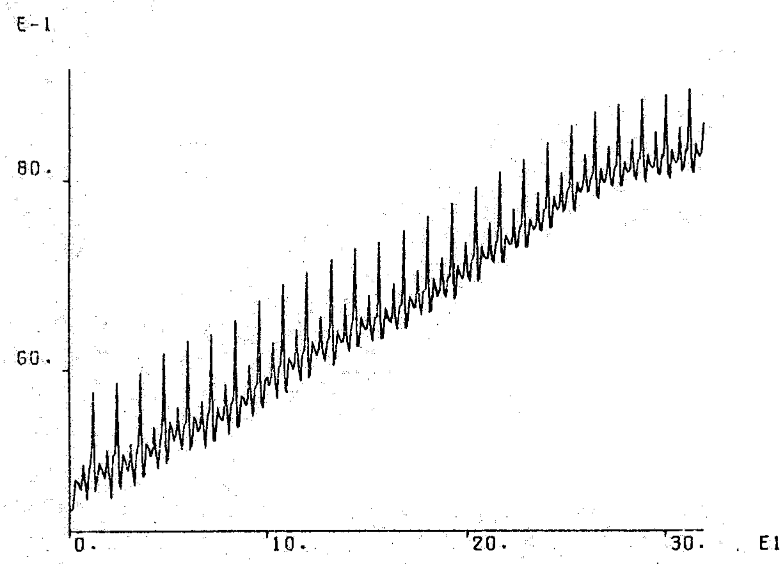


Fig. 1a Total sales at department stores in Japan (Data by Department Stores Association)

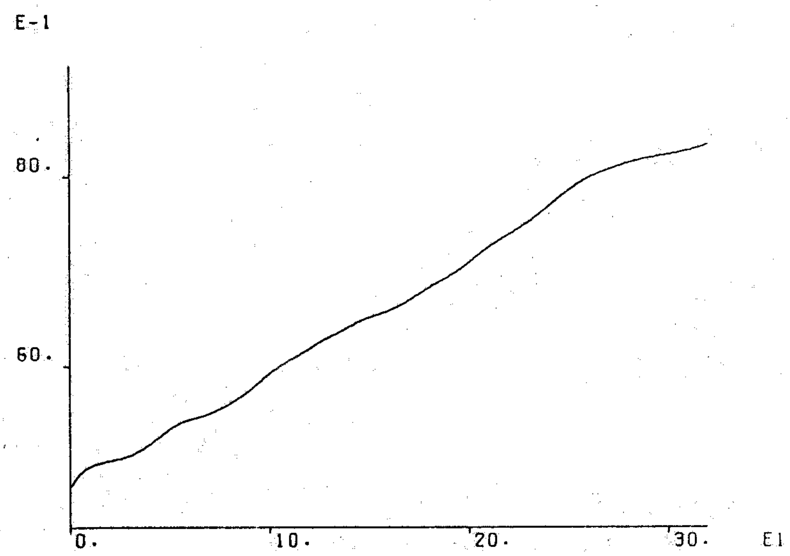


Fig. 1b Estimated trend

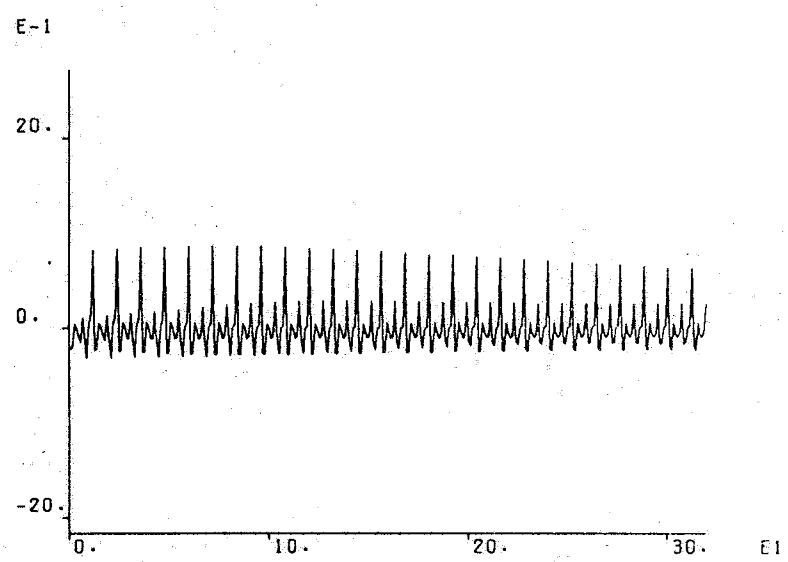


Fig. 1c Estimated seasonal component

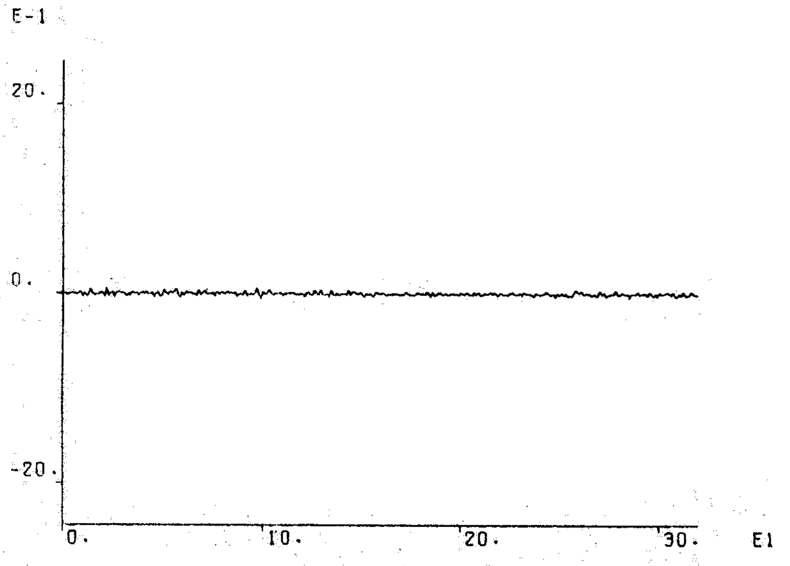


Fig. 1d Estimated irregular component

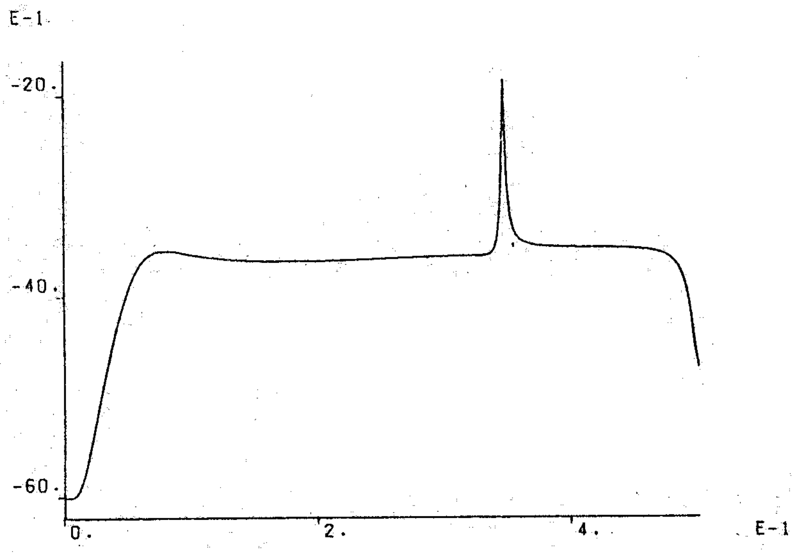


Fig. 1e Spectrum of the estimated irregular component

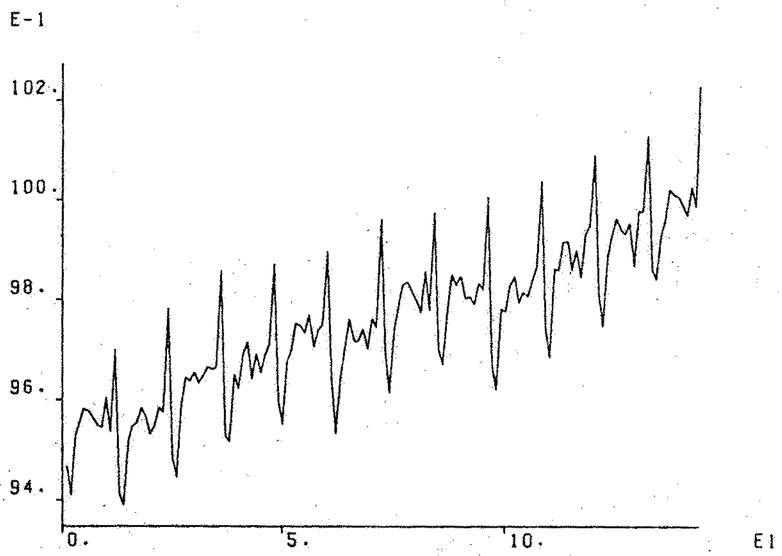


Fig. 2a Retail sales in U.S.A.
(Data by Census Bureau)

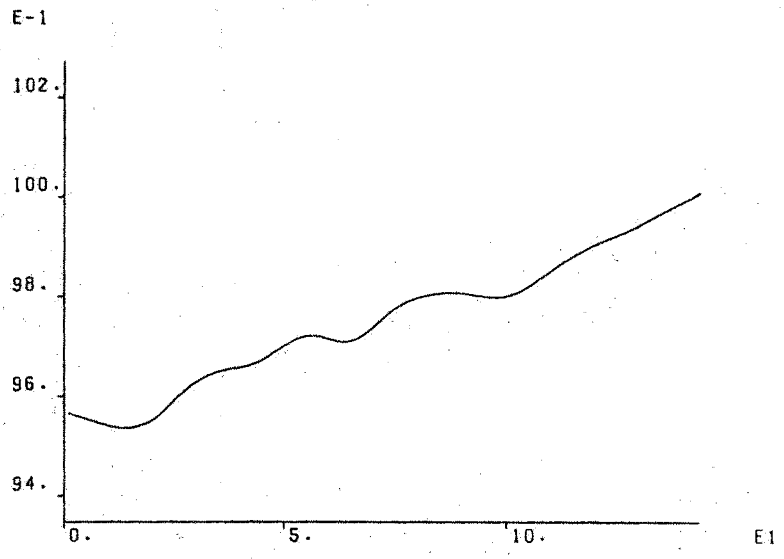


Fig. 2b Estimated trend (by BAYSEA)

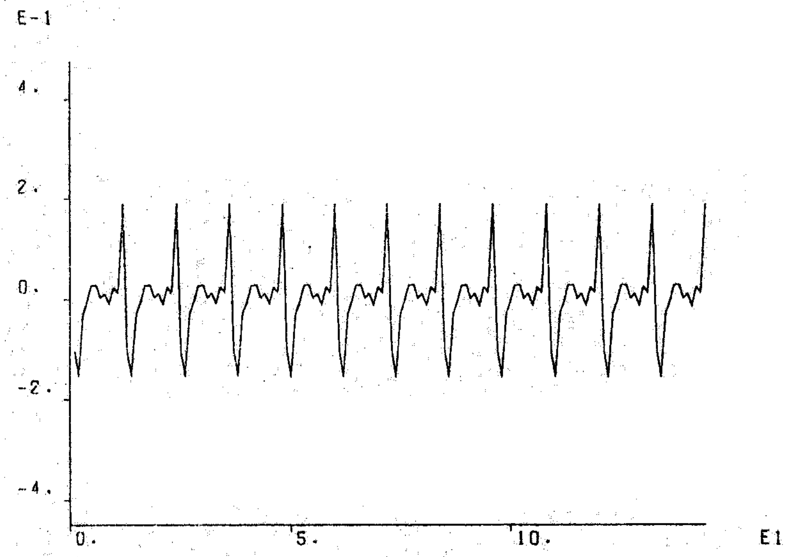


Fig. 2c Estimated seasonal component (by BAYSEA)

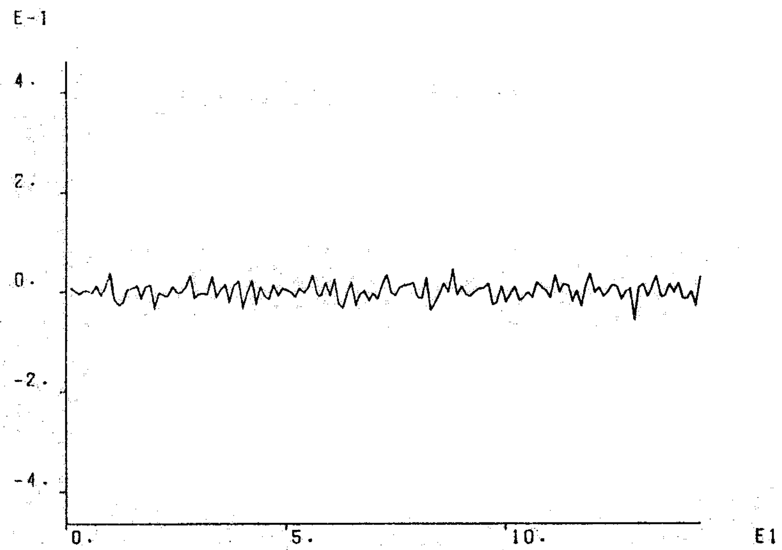


Fig. 2d Estimated irregular component (by BAYSEA)

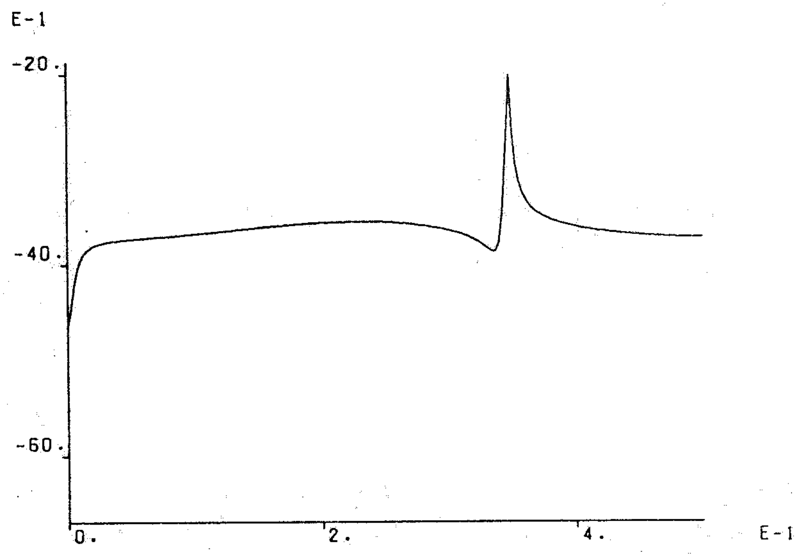


Fig. 2e Spectrum of the estimated irregular component

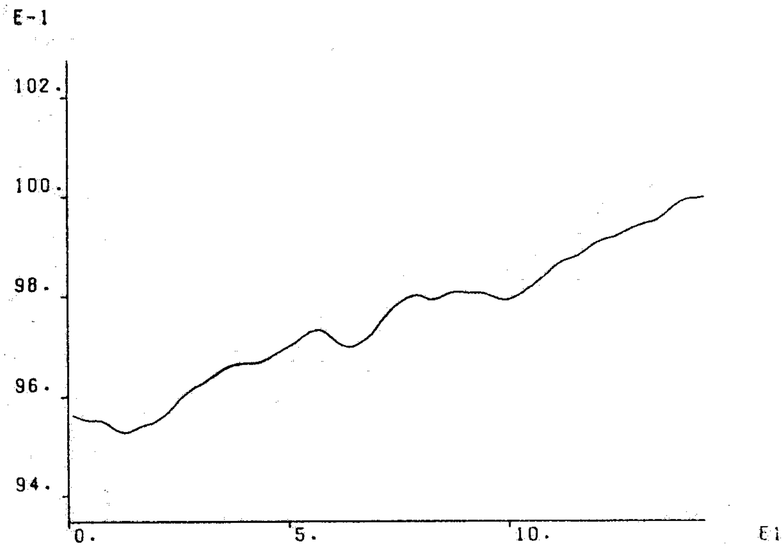


Fig. 3b Estimated trend (by X-11)

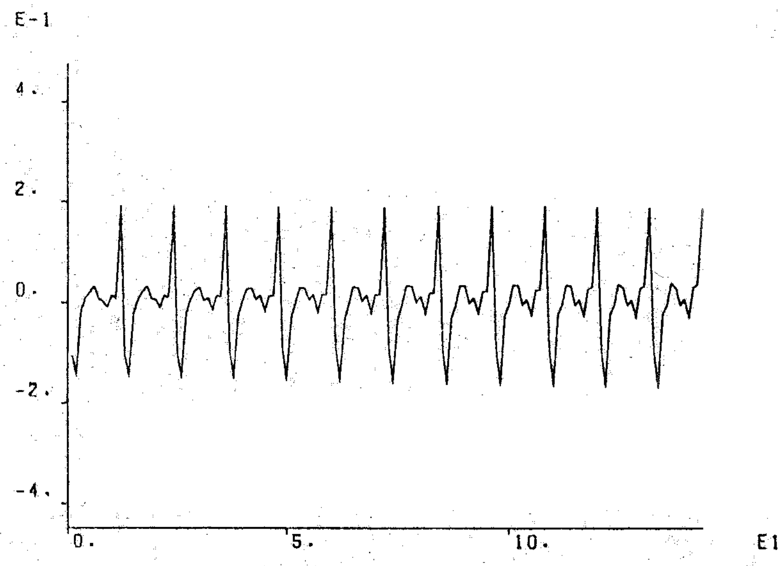


Fig. 3c Estimated seasonal component (by X-11)

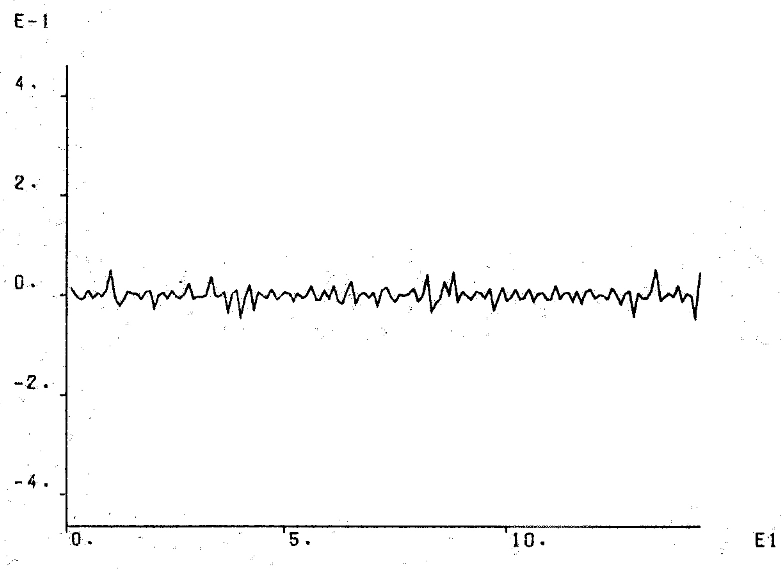


Fig. 3d Estimated irregular component (by X-11)

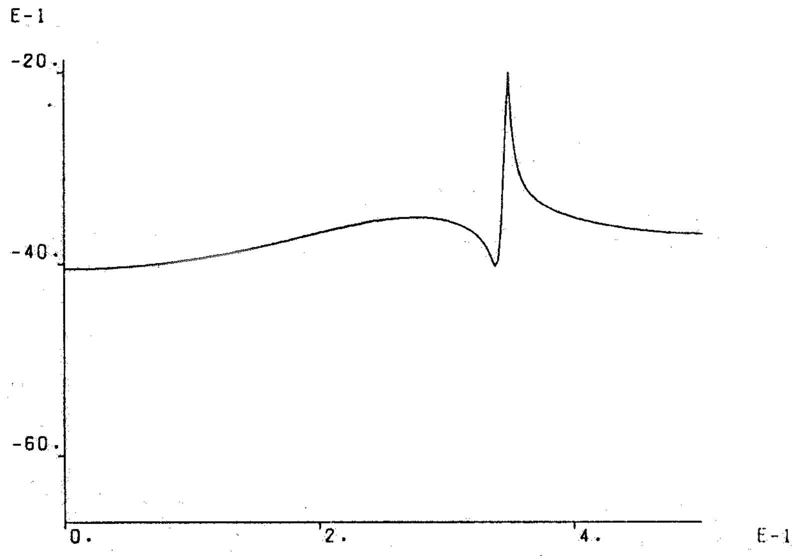


Fig. 3e Spectrum of the estimated irregular component

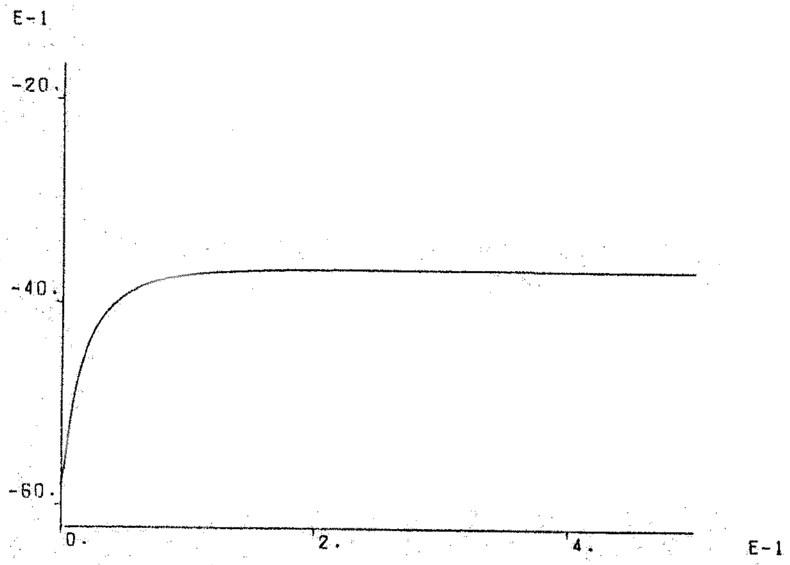


Fig. 4a Spectrum of the irregular component of department store sales data, which is estimated by BAYSEA with trading-day adjustment

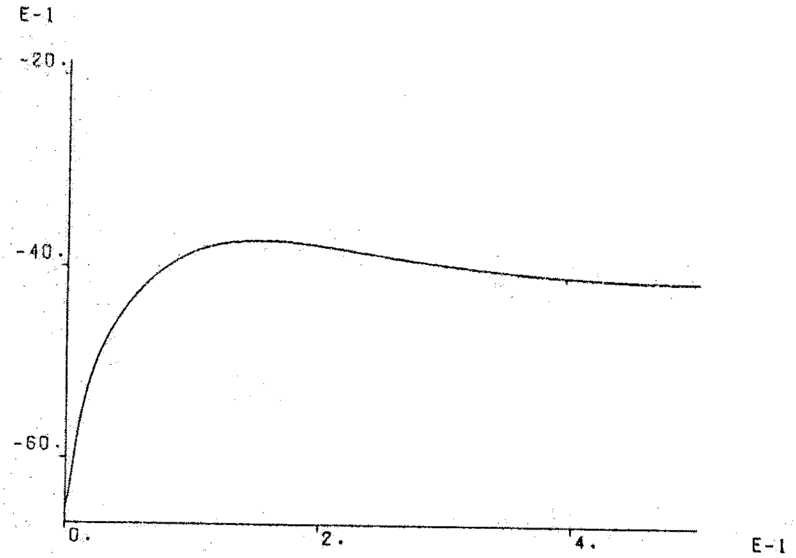


Fig. 4b Spectrum of the irregular component of retail sales data, which is estimated by BAYSEA with trading-day adjustment

Calendar Effects in Monthly Time Series: Modeling and Adjustment

WILLIAM S. CLEVELAND and SUSAN J. DEVLIN*

Most monthly time series that represent a total of some variable for each month contain calendar effects due to changing month length, day-of-the-week effects, and holidays. It is important to remove the calendar variation to allow an effective assessment of the variation due to important factors. New procedures for calendar adjustment are presented in this article. A plausible model for the daily data is used to derive a model for the monthly data in which the power transformed, month-length corrected data are equal to trend plus seasonal plus calendar plus irregular. The procedure for fitting the calendar component in the monthly model is (a) divide the aggregated monthly data by month length and multiply by 30.4375, the average month length; (b) choose a power transformation; (c) remove trend and seasonal components; (d) estimate the calendar parameters by robust regression. Since the model is only a hypothesis it is important to check its validity. This can be done by various residual plots in the regression analysis and by using spectrum analysis and time-domain graphical methods to detect residual calendar effects in the adjusted series. This approach is compared to the X-11 calendar estimation procedures.

KEY WORDS: Trading day adjustment; Seasonal adjustment; Spectrum analysis; Graphics.

1. INTRODUCTION

1.1. Calendar Effects

Most monthly time series that represent a total of some variable for each month contain calendar effects due to holidays, changing month length, and day-of-the-week effects. Such series also usually contain three other components: a trend component, which describes the long-term change in the level of the series; a yearly seasonal component, which describes the periodic or nearly periodic variation with a period of 12 months; and an irregular component, which describes the noisy variation.

One example is the number of interstate toll messages (i.e., phone calls) in the Bell System. The monthly series has a calendar component since the number of messages on a particular day depends on the day of the week and whether the day is a holiday; on Saturdays, Sundays, and many holidays, for example, there are fewer calls since business calling is substantially reduced. Thus the monthly messages series will have variation due to the changing fraction of each day of the week in the months and variation due to the changing fraction of holidays; the first type of variation together with part of the second is the calendar component. The remainder of the second, as we will see later, is in the seasonal, provided certain assumptions about the holidays are true. For the messages series, for example, months with a greater fraction of workdays than average will tend to have a greater number of messages.

In seasonal and calendar adjustment the seasonal and calendar components are removed since they represent variation that is typically not of primary interest and can obscure movement that is of primary interest. The series with the components removed is the adjusted series. For example, enormous numbers of business and national economic series are calendar adjusted and seasonally adjusted (c.f. Bureau of Economic Analysis 1976), since the calendar and seasonal variation is often a large fraction of the total variation in the series.

1.2 Current Methodology for Modeling and Adjustment

The most-used calendar adjustment procedure is that contained in the Census X-11 package (Shiskin, Young, and Musgrave 1967). The methodology is based on a discussion of calendar effects by Young (1965). The X-11 procedure carries out adjustment of the monthly data using only the monthly data itself. However, if the daily data are available one can rather straightforwardly determine the effect of each day of the week and holidays, and then adjust the monthly data. Frequently, however, daily data are not available or, if they are, the processing can be expensive since data handling increases by a factor of approximately 30 over the monthly series. Thus, procedures that use only the monthly data are typically the relevant ones for practice.

1.3 Overview

The purpose of this article is to present new methodology for dealing with calendar effects. The description begins in Section 2 with a model for daily data from which a model for the aggregated monthly data is derived. The fitting of the derived model in Section 3 begins by dividing

* William S. Cleveland and Susan J. Devlin are statisticians at Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974.

September 1982, Volume 77, Number 379
Applications Section

the aggregated monthly data by month length, multiplying by the average month length, 30.4375, and then applying a power transformation to the result to make the model additive. Next, the trend and seasonal components are removed from the transformed data and the result is regressed on certain calendar variables.

In Section 4 methods are suggested for checking the adequacy of the fitted model and the adjusted series. Residual plots are used to assess the assumptions made in the regression. Calendar detection procedures, which are discussed in detail by Cleveland and Devlin (1980), are used to determine if there is any residual calendar variation in the adjusted series.

Finally, the X-11 procedures and the procedures presented here are compared in Section 5 by Monte Carlo experimentation.

1.4 Examples

Several examples will be used to illustrate the methodology. The first is the interstate toll messages series from January 1964 to December 1978. The second is the number of housing starts for privately owned dwellings from January 1960 to December 1974 (Bureau of Economic Analysis 1976). The third is the number of international airline passengers from January 1949 to December 1960 (Box and Jenkins 1976).

2. MODELS FOR THE DAILY AND MONTHLY SERIES

Let $X(D)$ be the aggregated daily series for the Dth day. Then the aggregated monthly series for the mth month is

$$\sum_m X(D),$$

where $\sum_m$ means the sum over all days in month m. In this section we hypothesize a model for $X(D)$ and then derive the properties of the aggregated monthly series.

2.1 The Daily Model

A class of power transformations (Box and Cox 1964; Tukey 1957) for positive u with parameter p is defined by

$$u^{(p)} = \begin{cases} u^p & \text{if } p > 0 \\ \log u & \text{if } p = 0 \\ -u^p & \text{if } p < 0. \end{cases}$$

For all p $u^{(p)}$ is an increasing function of u. The inverse transformation of $u^{(p)}$, which takes us back to the original scale, is

$$v^{[p]} = \begin{cases} v^{1/p} & \text{if } p > 0 \\ 10^v & \text{if } p = 0 \\ (-v)^{1/p} & \text{if } p < 0. \end{cases}$$

We shall suppose that a power transformation of the daily data has four additive components,

$$X^{(p)}(D) = T(D) + S(D) + C(D) + I(D). \quad (2.1)$$

$T(D)$ is the trend component in the daily series; $S(D)$ is the seasonal component with a period of one year; $I(D)$ is the irregular component; and $C(D)$ accounts for a day-of-the-week effect (i.e., a weekly cycle) in the series so that $C(D) = \alpha_k$, if D is the kth day of the week, where $\sum_{k=1}^{7} \alpha_k = 0$. Modifications of this model to account for the effect of holidays are given in Section 2.3.

2.2 The Monthly Model

Unfortunately, to go from the daily to the monthly model we need to introduce more notation. A lowercase letter with a bar will be used to denote an average daily value of a series for each month. Thus

$$\bar{x}(m) = \frac{\sum_m X(D)}{\text{number of days in month } m};$$

and $\bar{t}(m)$, $\bar{s}(m)$, $\bar{c}(m)$, and $\bar{i}(m)$ are similarly defined. We let $x(m) = 30.4375\bar{x}(m)$ be the month-length corrected monthly series, where $30.4375 = 365.25/12$ is the average month length. We also let $t(m) = (30.4375)^{(p)}\bar{t}(m)$, $s(m) = (30.4375)^{(p)}\bar{s}(m)$, $c(m) = (30.4375)^{(p)}\bar{c}(m)$, and $i(m) = (30.4375)^{(p)}\bar{i}(m)$. Finally we let $\bar{d}_k(m)$ be the fraction of times the kth day of the week occurs in month m and $d_k(m) = (30.4375)^{(p)}\bar{d}_k(m)$.

Let us first consider the case $p = 1$. Summing both sides of (2.1), dividing each side by the number of days in month m, and multiplying by 30.4375 gives

$$x(m) = t(m) + s(m) + c(m) + i(m),$$

where

$$c(m) = \sum_{k=1}^{7} \alpha_k d_k(m). \quad (2.2)$$

In the case where $p \neq 1$, the situation is somewhat more complicated but it does turn out that an equation similar to the last one holds approximately. This will be shown for the case $p > 0$; a nearly identical argument holds for $p < 0$, and an argument similar in approach but somewhat different in detail holds for $p = 0$. The critical assumptions are the following:

The values of $S(D)$, $C(D)$, and $N(D)$ are small compared with $T(D)$. (2.3)

$T(D)$ is nearly constant over a period of one month. (2.4)

From (2.1)

$$X(D) = T^{[p]}(D)\left(1 + \frac{S(D) + C(D) + I(D)}{T(D)}\right)^{[p]}.$$

Using (2.3) and the power-series expansion around one of the inverse power transformation,

$$u^{[p]} \approx 1 + \frac{1}{p}(u - 1),$$

we have

$$X(D) \approx T^{[p]}(D)\left(1 + \frac{S(D) + C(D) + I(D)}{pT(D)}\right). \quad (2.5)$$

In view of assumption (2.4), the values of $T(D)$ for all D in the mth month are nearly the same. Thus for D in the mth month, $T(D)$ can be replaced by $\bar{t}(m)$. Now summing both sides of (2.5) and dividing by month length we have

$$\bar{x}(m) \approx \bar{t}^{(p)}(m)\left(1 + \frac{\bar{s}(m) + \bar{c}(m) + \bar{i}(m)}{p\bar{t}(m)}\right).$$

Now, using assumption (2.3) and a power-series expansion around one of the power transformation,

$$x^{(p)} \approx 1 + p(x - 1),$$

we have

$$\bar{x}^{(p)}(m) \approx \bar{t}(m) + \bar{s}(m) + \bar{c}(m) + \bar{i}(m),$$

and multiplying both sides of this approximate equation by $(30.4375)^{(p)}$ we have

$$x^{(p)}(m) \approx t(m) + s(m) + c(m) + i(m), \qquad (2.6)$$

where $c(m)$ is defined by (2.2).

2.3 Holidays

To take account of holidays we need to modify the model in (2.1) in the following way:

$$X^{(p)}(D) = T(D) + S(D) + C(D) + I(D) + H(D), \qquad (2.7)$$

where $C(D)$ is now modified to be zero if D is a holiday, and where $H(D)$ is β_j when D is the day of the jth holiday of the year and $H(D)$ is zero otherwise.

Model (2.7) assumes that the holiday effects are additive and that the effect of each holiday is the same from year to year and does not change, for example, if the day of the week on which the holiday occurs changes.

Suppose the jth holiday always occurs within the same month; then it is clear that the holiday's effect is a purely seasonal effect and can therefore be included as part of the seasonal component. In the United States the only holiday of any consequence that changes months is Easter. Thus we will redefine the $S(D)$ and $H(D)$ components so that $S(D)$ includes all holiday effects except Easter and $H(D)$ describes Easter.

Suppose the jth holiday always occurs on the same day of the week and within the same month. Let D be a day on which the jth holiday occurs and suppose it is the kth day of the week. Then we can alter $C(D)$ by changing it back to α_k, and we can alter $S(D)$ by changing it from $S(D)$ to $S(D) - \alpha_k$; Equation (2.7) still holds and our new $S(D)$ is still legitimately a seasonal component since the value of α_k subtracted is the same each year. Of course, this same change cannot be done for a holiday that occurs on different days of the week from one year to the next (such as Christmas and January 1), since then the α_k subtracted would be different from one year to the next.

When we aggregate (2.7) over months and divide by month length we now get (in place of (2.6))

$$x^{(p)}(m) = t(m) + s(m) + c(m) + i(m) + h(m),$$

where $c(m)$ is defined as in (2.2), but $\bar{d}_k(m)$ now equals the fraction of occurrences of the kth day of the week in the mth month minus the fraction of days on which the kth day of the week is a holiday that does not occur within the same month and on the same day of the week. Let γ be the Easter effect, then $h(m)$ is γ if Easter occurs in the mth month and is zero otherwise; thus if Easter is likely to have an effect very different from that of other Sundays, its effect can be included in the calendar model as a dummy variable and its occurrence should not be counted in $\bar{d}_k(m)$ for Sunday.

In the remainder of the article we shall suppose that the occurrence of Easter has a negligible effect on the series and will not consider it in our modeling. The "moving" holidays used in computing $\bar{d}_k(m)$ are New Years Day, Memorial Day, July 4, and Christmas.

3. FITTING THE MONTHLY MODEL

3.1 Month-Length Correction

The first step is the month-length correction of the aggregated monthly series: dividing by month length and multiplying by 30.4375. The division corrects the aggregated series for month length and the multiplication simply brings the units back to amount per month. Month-length correction, which is a rather simple operation, results in a much less complex model for the series. From (2.6) we can see that not correcting would result in the equation

$$\text{aggregated series for month } m = v^{(p)}(m)t(m) + v^{(p)}(m)s(m) + v^{(p)}(m)c(m) + v^{(p)}(m)i(m),$$

where $v(m)$ is the length of the mth divided by 30.4375. Now the trend and seasonal are confounded with a month-length effect, which complicates the structure since the separate effects are no longer additive.

3.2 Choosing a Power Transformation

The next step in the fitting procedure is to choose a power transformation of the month-length corrected series, $x(m)$, so that the decomposition will be the additive one, in terms of the four components, in (2.6). If an incorrect value of p is chosen, the transformed series involves interactions between the components. For example, suppose $p = .5$ and p is incorrectly chosen to be 1; then

$$x(m) = t^2(m) + s^2(m) + c^2(m) + i^2(m) + 2t(m)s(m) + 2s(m)i(m) + 2t(m)i(m) + 2t(m)c(m) + 2s(m)c(m) + 2c(m)i(m),$$

so that interaction terms, such as $2t(m)c(m)$, are present. There is not, of course, any guarantee that a power transformation will make the decomposition additive; in some cases it might be that a transformation not in this class is needed or it might be that no simple transformation will produce additivity. But our experience has been that the class of power transformations is a rich enough one to satisfy most situations in practice.

One procedure for estimating the power p is to minimize a measure of the interaction between the trend and the seasonal (Cleveland, Devlin, and Terpenning 1981). This is done by computing a t statistic for the interaction term and choosing a power that makes the t statistic close to zero. For the three examples in Section 1.4 the resulting values of p (chosen from the values 0, ± .25, ± .5, ± 1) are .25 for messages, −.25 for the airline series, and .5 for housing starts.

3.3 Removing Trend and Seasonal Components; Matched Processing

Suppose L is a processor that removes the trend and seasonal components and suppose that L is linear. Then from (2.6)

$$Lx^{(p)}(m) = Lt(m) + Ls(m) + \sum_{k=1}^{7} \alpha_k Ld_k(m) + Li(m) \doteq \sum_{k=1}^{7} \alpha_k Ld_k(m) + Li(m). \tag{3.1}$$

Thus α_k can be estimated by regressing $Lx^{(p)}(m)$ on the seven explanatory variables, $Ld_k(m)$, $k = 1, \ldots, 7$, subject to the constraint that $\sum_{k=1}^{7} \alpha_k = 0$. Applying L to $d_k(m)$ as well as to $x^{(p)}(m)$ will be referred to as *matched processing*. It is important to process $d_k(m)$ so that (3.1) is true, but we shall see in Section 5 that not all procedures use matched processing.

One method for eliminating the trend and seasonal components is to decompose $x^{(p)}(m)$ into trend plus seasonal plus irregular components using a decomposition procedure such as that in X-11 (Shiskin, Young, and Musgrave 1967), X-11 ARIMA (Dagum 1978), or SABL (Cleveland, Devlin, and Terpenning 1981), and take $Lx^{(p)}(m)$ to be the irregular component. For the three examples in Section 1.4 the SABL procedure was used with the length of the trend smoothing window equal to 11 and the length of the seasonal smoothing window equal to 15. To guard against end effects having a distorting effect on the results, the first six and last six values of $Lx^{(p)}(m)$ were discarded before carrying out the estimation described in the next section.

When one of the three trend-seasonal-irregular decompositions just mentioned is used to remove the trend and seasonal components, L is only approximately linear. To do the matched processing we would not apply the decomposition directly to $d_k(m)$ since there are weighting schemes in which outliers are identified and downweighted in the smoothing procedures. But the effect of L on $d_k(m)$ can be approximated by using an $\tilde{L}$ in which the weighting for $x^{(p)}(m)$ is used in the processing of $d_k(m)$. $\tilde{L}$ for the SABL decomposition is described in Cleveland, Devlin, and Terpenning (1981). $\tilde{L}$ for other decompositions can be defined in an analogous fashion.

3.4 Estimation of α_k

After the trend and seasonal components are removed from $x^{(p)}(m)$, α_k can be estimated by regressing $Lx^{(p)}(m)$ on the seven explanatory variables $\tilde{L}d_k(m)$, $k = 1, \ldots, 7$, with the constraint that $\sum_{k=1}^{7} \alpha_k = 0$. For the major domain of application of calendar-adjustment methods, which is business and economic series, outliers frequently occur and can have a substantially distorting effect on parameter estimates unless robust methods are used. Least squares is nonrobust (Hampel 1973; Huber 1964; Kuznets 1933). Thus we have used iterated weighted least squares with the bisquare weight function and $k = 6$ (Andrews 1971; Mosteller and Tukey 1977) to estimate the parameters. Probability plots demonstrate that error distributions with tails appreciably longer than the normal occur among the examples of Section 1.4, thus justifying our use of robust procedures.

The estimates, $\hat{\alpha}_k$, and their standard errors (Huber 1973) are shown in Table 1 for the three examples in Section 1.4. Each set of coefficients has been scaled by a power of 10 so that the coefficient that is largest in absolute value has one digit to the left of the decimal point. The coefficients for the messages series are high on weekdays and low on weekends, since a large portion of the series is business calls. The peak days are Tuesday and Wednesday, whereas Monday, Thursday, and Friday are reduced. Sunday is somewhat higher than Saturday due to an increase in residence calls on Sunday. For the housing-starts and airlines series, the coefficients are also higher on weekdays than weekends.

3.5 Calendar and Seasonal Adjustment

Since the α_k have now been estimated, the original aggregated monthly series can be calendar adjusted. First, the calendar-adjusted transformed series is

$$\hat{x}^{(p)}(m) = x^{(p)}(m) - \sum_{k=1}^{7} \hat{\alpha}_k d_k(m).$$

If seasonal adjustment is to be carried out also, then this series is decomposed into trend plus seasonal plus irregular, and the calendar and seasonally adjusted series, $\tilde{x}^{(p)}(m)$, is then the trend plus the irregular components. Now the adjusted monthly series for the mth month on the original, untransformed scale is $\tilde{x}(m) = [\tilde{x}^{(p)}(m)]^{(p)}$.

Table 1. Estimates of α_k for Three Series and Their Standard Errors

Day	Messages	Housing	Airline
Mon	2.24 ± .74	2.31 ± .78	.94 ± 1.31
Tue	4.91 ± .77	1.10 ± .77	1.20 ± 1.30
Wed	4.20 ± .76	−1.05 ± .79	3.14 ± 1.29
Thu	2.12 ± .74	1.23 ± .79	−.44 ± 1.29
Fri	1.69 ± .74	1.12 ± .80	−1.33 ± 1.29
Sat	−9.13 ± .73	−2.43 ± .79	−1.81 ± 1.35
Sun	−6.04 ± .71	−2.28 ± .78	−1.70 ± 1.32

4. CHECKING THE ADEQUACY OF THE FITTED MODEL

The model in (2.6) is a hypothesis whose validity must be checked in each application. Once the α_k have been estimated, the analyst can begin the diagnostic checking by making regression plots to check the validity of the regression model in (3.1) and by checking for residual calendar effects in the adjusted series using calendar-detection procedures (Cleveland and Devlin 1980).

4.1 Regression Plots

For (3.1) it is important to check the specification of

- the function of the explanatory variables, $\tilde{L}d_k(m)$
- the stochastic assumptions about $i(m)$.

The principal quantities for carrying out these checks are the fitted values $\sum_{k=1}^{7} \hat{\alpha}_k \tilde{L}d_k(m)$, and the residuals $Lx^{(p)}(m) - \sum_{k=1}^{7} \hat{\alpha}_k \tilde{L}d_k(m)$.

One diagnostic check for (3.1) is to plot residuals against fitted values (Anscombe and Tukey 1963); a nonlinear effect in the plot indicates a misspecification. Superimposing a smooth curve on the plot whose purpose is to describe the regression function of the residuals on the fitted values can greatly enhance the detection of effects. Figure 1 (upper left) shows such a plot for the airline series. The smooth curve, which has been computed using robust, locally weighted regression (Cleveland 1979), shows no dependence of the residuals on the fitted values. Another possible diagnostic for (3.1) is to plot residuals against each explanatory variable, $\tilde{L}d_k(m)$ (Draper and Smith 1966). But an even better procedure is to plot the residuals against the residuals from regressing the explanatory variable on all other explanatory variables in the model.

The assumptions made in the use of iterated weighted least squares are that the $Li(m)$ are independent and identically distributed with a symmetric distribution. These assumptions are considerably less restrictive than the usual normality assumptions made in regression. The symmetry of the residuals can be checked very effectively by the graphical display described by Wilk and Gnanadesikan (1968). A normal probability plot also gives information about symmetry along with information about the tails of the distribution of the residuals. Such a plot is shown in Figure 1 (lower right) for the airline series and demonstrates rather clearly the nonnormality of the errors. There are a number of outliers that distort the estimates and their standard errors when least squares is used. In the robust fit, the outliers receive small weights and thus do not adversely affect the fit.

The validity of the assumption of identical distributions can, in part, be assessed by checking for a nonconstant scale of the residuals. This can be done effectively by plotting the absolute residuals against the fitted values and smoothing them as in Figure 1 (upper right) for the airline series. The near constancy of the smooth curve indicates a nearly constant scale. Another check for identical distributions is to plot the residuals against time and smooth to see if a time trend is present. This has been done in Figure 1 (lower left) for the airline series.

The assumption of independence can be checked, in part, by computing estimates of the autocorrelations of the residuals. But again, because of outlier problems, we will not want to use the usual correlation coefficient but rather a robust estimate such as $r^*(SSD)$ (Devlin, Gnanadesikan, and Kettenring 1975).

The three examples in Section 1.4 pass the tests of this section. But there is more checking to be done.

4.2 Checking for Residual Calendar Effects

A second kind of diagnostic checking is to calendar adjust the series as described in Section 3.5, and then use calendar detection methods to check for residual calendar effects in the adjusted series. One procedure for doing this (Cleveland and Devlin 1980) is the following:

1. Decompose the calendar-adjusted $x^{(p)}(m)$ series into trend plus seasonal plus irregular components.
2. Adjust the outliers in the irregular component by "clipping."
3. Estimate the spectrum of the clipped irregular component and look for peaks at the two calendar frequencies, .348 cycles per month and .432 cycles per month.
4. Estimate the spectra of each of the 11 monthly subseries (excluding February) of the clipped irregular component, average the estimates, and look for peaks at .179 cycles per year and .357 cycles per year.

This spectrum detection procedure was used for the three examples described in Section 1.4. No residual effects were detected for the airline series or for the housing-starts series. The plots of the series spectrum estimate and the subseries spectrum estimate for the adjusted messages series are shown in Figure 2 and the calendar frequencies are indicated by dashed vertical lines. For this series, peaks occur in the subseries estimate at .179 and .357 cycles per year. Other information has suggested that the calendar effect in the messages series changes slightly with the time of year (e.g., a Monday in June has an effect different from that of a Monday in December). For a series with "seasonal" calendar effects that is adjusted assuming constant effects, the residual calendar effect will typically not appear in the series spectrum estimate but can appear in the subseries estimate. Thus the peaks at the calendar frequencies might be a result of changing calendar effects, which are discussed in Section 6.1.

5. THE CENSUS X-11 PROCEDURE

The most-used calendar adjustment procedure at the moment is that in Census X-11 (Shiskin, Young, and Musgrave 1967). For the additive version of X-11, the irregular component is regressed on $\tilde{d}_k(m)$ subject to the constraint that the coefficients sum to zero. For the mul-

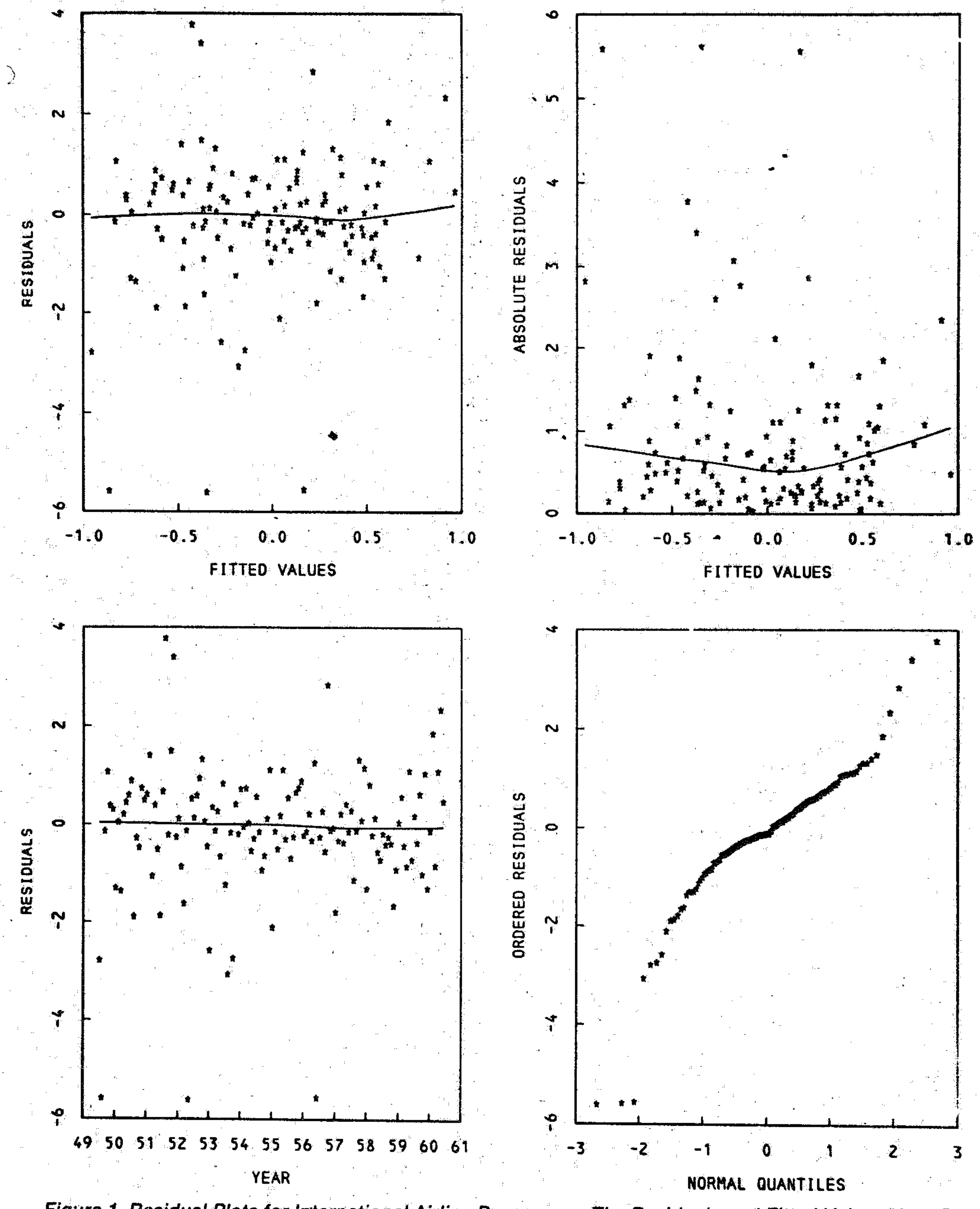


Figure 1. Residual Plots for International Airline Passengers. The Residuals and Fitted Values Have Been Multiplied by 1,000. Transformation power −.25

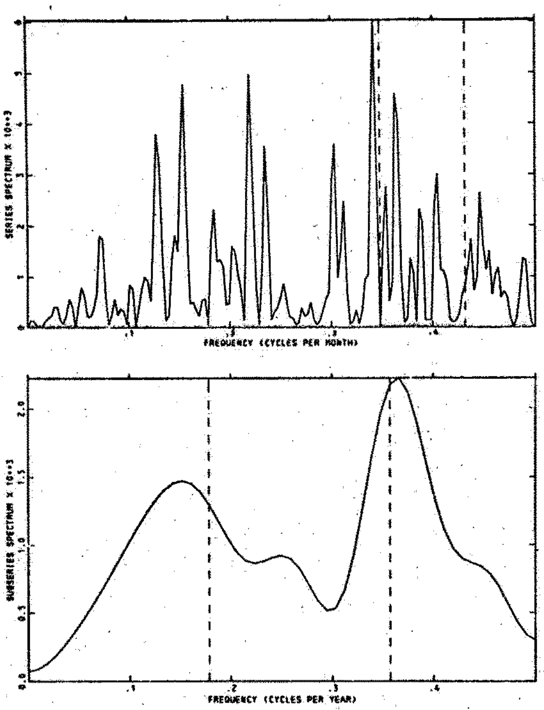


Figure 2. Spectrum Estimates of Clipped Irregular for Adjusted Interstate Toll Messages. Transformation power −.25

tiplicative model the irregular component minus one is regressed on the same variables with the same constraint.

There are a number of serious inadequacies in the X-11 procedure. First, there is no month-length correction for the additive version; this contributes to instability in the seasonal component and thus increases the complexity of the seasonal variation. In X-11, no matched processing is carried out on the explanatory variables. But the discussion in Section 3.3 shows that the unprocessed explanatory variables will not capture all of the calendar variation in the series.

In X-11 the user chooses between a multiplicative or an additive model, which is a much more limited set of choices than one gets by employing the transformations described in Section 3.2. When $p = 1$, the model is additive and when $p = 0$, the model is multiplicative, but for all other cases the data are neither purely additive nor purely multiplicative. For all of the examples in Section 1.4, the value of p is neither zero nor one.

The regression methodology in X-11 has procedures for dealing with outliers, but they are not adequate. During the 15 years since the development of X-11, a substantial amount of work, both theoretical and empirical, has been done on robust estimation. The bisquare procedure used in Section 3 is one that has been shown to be nearly optimal when the data are well behaved (e.g., have a normal distribution) but that protects against the distorting effect of outliers when they occur (Mosteller and Tukey 1977).

Finally, the X-11 package contains no procedures for diagnosing the performance of the calendar adjustment. As explained in Section 4 this is an important part of the whole adjustment process without which confirmation can be replaced only by hope.

Monte Carlo experimentation was used to investigate the properties of the X-11 calendar-coefficient estimation procedures and those described here. In the first experiment we generated 250 daily series of the form

$$X(D) = T(D) + S(D) + C(D) + I(D).$$

Both $T(D)$ and $S(D)$ were taken to be zero. The values of α_j that make up the calendar component, $C(D)$, were the following: $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = 1$, $\alpha_4 = 1$, $\alpha_5 = 0$, $\alpha_6 = -1.5$, and $\alpha_7 = -1.5$. $I(D)$ was a sequence of independent normal random variables, each with mean 0 and standard deviation .05. The length of each series was 15 years.

Each of the 250 daily series was aggregated to form a monthly series that was then processed by the additive X-11 procedures and by the procedures described here to get estimates of the seven calendar coefficients. Figure 3 shows the Monte Carlo biases and the square roots of

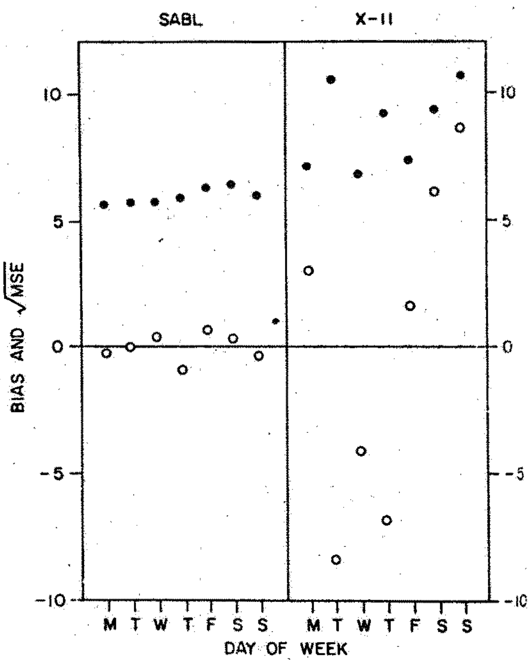


Figure 3. Biases (○) and Square Roots of Mean Squared Errors (•) for Normal Monte Carlo

the mean square errors of the estimates for the two procedures. The X-11 estimates (right panel) are severely biased, which greatly increases the mean square errors, whereas the biases of the procedures described here (left panel) are insignificant compared with their standard errors (not shown on the plot). The reason for the biases in X-11 is the lack of matched processing.

We ran a second Monte Carlo experiment, identical to the first except that $I(D)$ was replaced by .0229 times a slash random variable, to investigate the robustness properties of the X-11 calendar-coefficient estimation procedures and those described here. A slash variable, which is a ratio of a normal to a uniform, has longer tails than the normal and produces outliers; it is commonly used in Monte Carlo experiments that investigate the properties of estimators under long-tailed error distributions (e.g., Gross 1977). The constant .0229 was chosen so that the quartiles of the slash variable matched the quartiles of the normal variable used in the first Monte Carlo experiment. Figure 4 shows the square roots of the mean squared errors for the two procedures. X-11 has done very poorly owing to the lack of adequate robust estimation in the calendar regression.

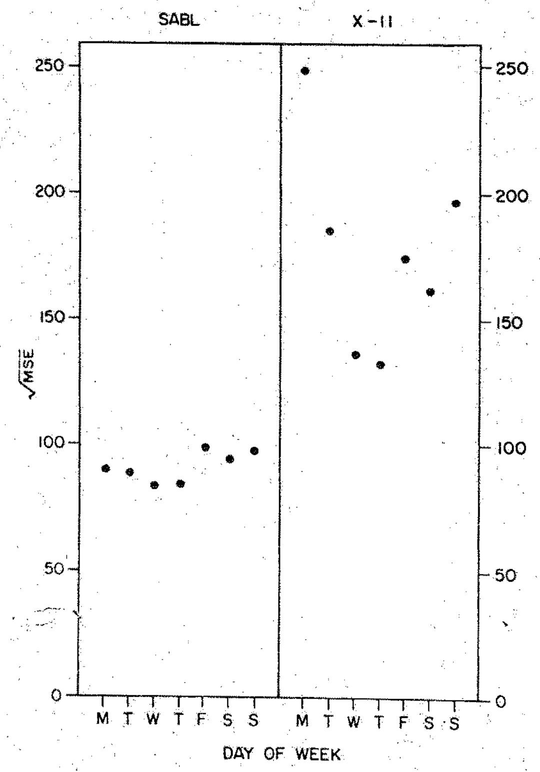


Figure 4. Square Roots of Mean Squared Errors for Slash Monte Carlo

6. SUMMARY AND DISCUSSION

6.1 Calendar Effects That Change Through Time

The modeling and adjustment procedures presented here are based on an assumption that the calendar effects do not change through time. But two types of changes that can occur are slow change over long periods (i.e., trend) and change with the month of the year (i.e., seasonal). For example, the subseries-spectrum plot for the adjusted messages series suggests a small amount of seasonal change in the calendar effects.

The procedures presented here can be modified in a straightforward way to account for changing calendar effects. One method is to allow the parameters to change through time. Also, trend in the calendar effects can be dealt with by limiting the time period used to estimate the parameters. For example, although the messages series extends much further back into time, data prior to 1964 were not used.

6.2 Outline of Procedures

Most time series that represent a total of some variable for each month contain calendar effects due to changing month length, day-of-the-week effects, and holidays. It is important to remove the calendar variation to allow an effective assessment of the variation due to important factors. The most-used calendar adjustment procedure is that contained in the Census X-11 package. But the X-11 procedure has a number of inadequacies (Sec. 5) that can result in residual calendar effects remaining in the adjusted series and biased estimates of calendar coefficients.

A plausible model for the daily data is used to derive a model for the monthly data (Sec. 2) in which the power-transformed, month-length corrected data is equal to trend plus seasonal plus calendar plus irregular components. The procedure for fitting the calendar component in the monthly model (Sec. 3) is the following:

1. Divide the aggregated monthly data by month length and multiply by 30.4375.
2. Choose a power transformation.
3. Remove the trend and seasonal components and process the day-of-the-week variables in a similar fashion.
4. Estimate the calendar parameters by robust regression.

Since the model is only a hypothesis, it is important to check its validity (Sec. 4). This can be done by using residual plots in the regression analysis and by using spectrum analysis to detect residual calendar effects in the adjusted series.

6.3 Computer Programs

The calendar estimation procedures described in this article have been incorporated into the SABL Computer Package. We cannot discuss computing considerations here in the interest of space, but those who would like to read a discussion of computer implementation can find it in Cleveland, Devlin, Schapira, and Terpenning (1981).

[*Received April 1980. Revised February 1982.*]

REFERENCES

ANDREWS, DAVID F. (1971), "A Robust Method for Multiple Linear Regression," *Technometrics*, 16, 523–531.

ANSCOMBE, F.J., and TUKEY, JOHN W. (1963), "The Examination and Analysis of Residuals," *Technometrics*, 5, 141–160.

BOX, G.E.P., and COX, D.R. (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society*, Ser. B, 26, 211–252.

BOX, GEORGE E.P., and JENKINS, GWILYM M. (1976), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day.

BUREAU OF ECONOMIC ANALYSIS (1976), *Business Statistics 1975*, Washington, D.C.: U.S. Department of Commerce.

CLEVELAND, WILLIAM S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74, 829–836.

CLEVELAND, WILLIAM S., and DEVLIN, SUSAN J. (1980), "Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods," *Journal of the American Statistical Association*, 75, 487–496.

CLEVELAND, WILLIAM S., DEVLIN, SUSAN J., SCHAPIRA, DORIS R., and TERPENNING, IRMA J. (1981), "The SABL Computer Package: Users' Manual," and "The SABL Computer Package: Specialized Usage," available from Computing Information Library, Bell Laboratories, 600 Mountain Ave., Murray Hill, N.J. 07974.

CLEVELAND, WILLIAM S., DEVLIN, SUSAN J., and TERPENNING, IRMA J. (1981), "The SABL Statistical and Graphical Methods," and "The Details of the SABL Transformation, Decomposition and Calendar Methods," available from Computing Information Library, Bell Laboratories, 600 Mountain Ave., Murray Hill, N.J. 07974.

DAGUM, ESTELA BEE (1978), "Modelling, Forecasting, and Seasonally Adjusting Economic Time Series With the X-11 ARIMA Method," *The Statistician*, 27, 203–216.

DEVLIN, S.J., GNANADESIKAN, R., and KETTENRING, J.R. (1975), "Robust Estimation and Outlier Detection With Correlation Coefficients," *Biometrika*, 62, 531–545.

DRAPER, N.R., and SMITH, H. (1966), *Applied Regression Analysis*, New York: John Wiley.

GROSS, ALAN M. (1977), "Confidence Intervals for Bisquare Regression Estimates," *Journal of the American Statistical Association*, 72, 341–354.

HAMPEL, FRANK R. (1973), "Robust Estimation: A Condensed Partial Survey," *Zeitschrift für Wahrscheinlichkeitstheorie*, 27, 87–104.

HUBER, PETER J. (1964), "Robust Estimation of a Location Parameter," *Annals of Mathematical Statistics*, 35, 73–101.

——— (1973), "Robust Regression: Asymptotics, Conjectures, and Monte Carlo," *Annals of Statistics*, 1, 799–821.

KUZNETS, SIMON (1933), "On Moving Correlation of Time Series," *Journal of the American Statistical Association* 48, 121–136.

MOSTELLER, FREDERICK, and TUKEY, JOHN W. (1977), *Data Analysis and Regression*, Reading, Mass.: Addison Wesley.

SHISKIN, JULIUS, YOUNG, ALLAN H., and MUSGRAVE, JOHN C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Bureau of the Census Technical Paper, U.S. Department of Commerce, Washington, D.C.

TUKEY, J.W. (1957), "On the Comparative Anatomy of Transformations," *Annals of Mathematical Statistics*, 28, 602–632.

WILK, M.B., and GNANADESIKAN, R. (1968), "Probability Plotting Methods for the Analysis of Data," *Biometrika*, 55, 1–17.

YOUNG, ALLAN (1965), "Estimating Trading-Day Variation in Monthly Economic Time Series," Technical Paper No. 12, Bureau of the Census, Washington, D.C.

Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods

WILLIAM S. CLEVELAND and SUSAN J. DEVLIN*

Many time series, particularly national business and economic series, which are reported monthly and represent a total of the series for each month, contain calendar effects due to changing month length, weekly periodicities, and holidays. It is important to *detect* and *remove* this spurious calendar variation to allow a better appreciation of the variation in the series due to important factors. This article discusses detection. Two sets of diagnostic methods for detecting calendar effects in monthly time series, spectrum analyses and time domain graphical displays, are described. These methods can be used in an initial analysis to decide if calendar adjustment is necessary and can be used on an adjusted series to determine if the adjustment has properly removed all of the calendar effects.

KEY WORDS: Calendar adjustment; Trading-day adjustment; Seasonal adjustment; Spectrum analysis; Graphics.

1. INTRODUCTION

1.1 Calendar Effects

Many time series, particularly national business and economic series, are reported monthly and represent a total, or time aggregation, of the series for each month. Because the month length is not the same for all months and because many of these series have strong weekly periodicities (i.e., the series for a particular day depends on the day of the week), the aggregated monthly series will vary, in part, due to the changing month length, the effect of the weekly periodicity, and the effect of holidays. We shall refer to this variation as calendar effects in the monthly series. Examples of economic series with strong calendar effects are manufacturing shipments, barrels of imported petroleum products, money supply, housing starts, telephone messages and revenues, and retail sales. To allow an effective interpretation of aggregated monthly series it is important to remove from the series that variation which is due to calendar effects. We shall refer to this as calendar adjustment.

1.2 The Investigation of Calendar Effects

Relatively little research on calendar effects has been carried out in the past, which contrasts with the large amount of work devoted to seasonal adjustment (e.g., Zellner 1979). Young (1965) and Eisenpress (1956) discuss calendar effects and present methods for carrying out an adjustment. The X-11 package (Shiskin, Young, and Musgrave 1967) contains a calendar adjustment option that is referred to as a "trading-day adjustment." Granger (1963) has studied series whose calendar effects depend on the number of work days in a month. The imbalance in research on calendar adjustment and seasonal adjustment would seem to be inappropriate since for many series the calendar variation is at least as important as the yearly seasonal variation not due to calendar effects.

In this article we present two sets of methods for detecting calendar effects in a monthly time series. These methods focus only on that part of the calendar variation due to weekly periodicity and changing month length, and do not explicitly consider holiday effects. The first set, which is described in Section 2, consists of spectrum analyses in which peaks in the spectrum at certain "calendar frequencies" indicate the presence of calendar effects. The second set, which is described in Section 3, consists of graphical displays in which the presence of certain patterns indicates the presence of calendar effects. In Sections 2 and 3 the techniques are illustrated by their application to an artificial "weekday" series, Bell System revenues from toll calls, an international airline passenger series, and manufacturers' shipments.

1.3 Detection and Modeling

There are two stages in the overall analysis of calendar effects in which these graphical and spectrum diagnostic techniques can be used. The first stage is one in which the techniques would be used to decide if calendar effects are present and are sufficiently important to warrant modeling and adjustment. (Techniques for such modeling and adjustment are discussed elsewhere (Cleveland and Devlin 1980).) The second stage occurs after an adjustment has been carried out. Here the techniques can be applied to check the adequacy of the adjustment by checking for the presence of remaining calendar effects in the adjusted series. Thus the use of the detection procedures is analogous to the use of the usual summed, lagged product estimates of autocorrelation in time series modeling (Box and Jenkins 1976). The autocorrelation function, which is relatively simple to estimate, is used initially to determine if autocorrelation is present. If so, a model is used to account for it, an adjusted series (residuals) is computed, and the auto-

* William S. Cleveland and Susan J. Devlin are Statisticians, Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974.

September 1980, Volume 75, Number 371
Applications Section

correlation function of the adjusted series is studied to determine if there is any residual autocorrelation.

2. SPECTRUM ANALYSIS FOR DETECTING CALENDAR EFFECTS

2.1 Overview

Suppose $c(t)$ is a periodic series with period l so that $c(t + l) = c(t)$. For l even, $c(t)$ can be written in the form

$$c(t) = \mu + \sum_{j=1}^{l/2} \left\{\alpha_j \cos\left(2\pi t \frac{j}{l}\right) + \beta_j \sin\left(2\pi t \frac{j}{l}\right)\right\} \tag{2.1}$$

where

$$\alpha_j = \frac{2}{l} \sum_{t=1}^{l} c(t) \cos\left(2\pi t \frac{j}{l}\right) ,$$

$$\beta_j = \frac{2}{l} \sum_{t=1}^{l} c(t) \sin\left(2\pi t \frac{j}{l}\right) ,$$

for $j = 1, \ldots, (l/2) - 1$,

$$\mu = \frac{1}{l} \sum_{t=1}^{l} c(t) ,$$

$$\alpha_{l/2} = \frac{1}{l} \sum_{t=1}^{l} c(t) \cos(\pi t) ,$$

and

$$\beta_{l/2} = 0 .$$

(A similar formula holds when l is odd.) The spectrum $c(t)$ at frequency $f_j = j/l$, for $j = 1, \ldots, l/2$, is the squared amplitude, $\alpha_j^2 + \beta_j^2$, of the harmonic component in (2.1) at frequency f_j. The spectrum at other frequencies is 0.

Let $x(t)$ be an aggregated monthly series and let $x_m(y)$ be the mth monthly subseries of $x(t)$ so that for the yth year,

$$x_m(y) = x(12(y - 1) + m) .$$

Calendar effects result in periodic components of the form (2.1) in $x(t)$ and $x_m(t)$ at certain known frequencies, which will be referred to as "calendar frequencies" and will be discussed further in Sections 2.2.1 and 2.2.2. The first spectrum procedure for identifying calendar effects is to estimate the spectrum of $x(t)$ and look for peaks at the calendar frequencies. The second spectrum procedure is to estimate the spectrum of $x_m(y)$ for each month (excluding February, which behaves differently from the other months), average the spectrum estimates across the 11 months for each frequency, and look for peaks at the calendar frequencies. For example, in Figures C through E the peaks in the spectrum estimates at the calendar frequencies shown by the vertical dotted lines are indications of calendar effects. The principal reason for introducing a second spectrum estimate is to have a corroborative procedure. There are times, however, when one procedure will yield a very clear indication of calendar effects and the other will not.

2.2 The Spectrum of Calendar Effects

2.2.1 The Calendar Frequencies. We shall think of the unaggregated series as a continuous parameter time series, $X(T)$, where the units of the parameter T are days. Let T_0 be the beginning of the first month and let T_t be the time at the end of the tth month. Then the aggregated monthly series is

$$x(t) = \int_{T_{t-1}}^{T_t} X(T)dT ,$$

for $t = 1, 2, \ldots$. Let $C(T)$ be a weekly periodicity in $X(T)$ (i.e., $C(T + 7) = C(T)$) whose integral over a period of 7 days is 0, and suppose $X(T) = B + C(T) + R(T)$, where B is a constant and $R(T)$ is the remaining variation in $X(T)$. Then $x(t) = b(t) + c(t) + r(t)$, where b, c, and r are the aggregates of B, C, and R, respectively. The calendar component in $x(t)$ induced through the aggregation process is $b(t) + c(t)$. The term $b(t)$ is B times the number of days in month t. As we shall see, $c(t)$ has a more complicated pattern.

All 30-day months that start on a particular day of the week, say Friday, will have the same value of $B + C(T)$ and therefore the same value of $b(t) + c(t)$. A similar statement holds for 31-, 29-, and 28-day months. If we neglect the fact that leap year is omitted every 400 years, the calendar has a period of 28 years. Thus $b(t) + c(t)$ has a period of 28 years = 336 months. Let $c_m(y)$ be the mth monthly subseries of $c(t)$ so that for the yth year $c_m(y) = c(12(y - 1) + m)$. Define $b_m(y)$ in a similar fashion. Then $b_m(y) + c_m(y)$ has a period of 28 years. Thus from (2.1), the spectrum of $b_m(y) + c_m(y)$ is concentrated at frequencies $j/28$, for $j = 1, \ldots, 14$, and the spectrum of $b(t) + c(t)$ is concentrated at the frequencies $j/336$, for $j = 1, \ldots, 168$. But knowledge of these "calendar frequencies" is not very informative since there are so many. It is necessary to know which of these frequencies are typically important for series in practice. In the next sections we shall use a line of reasoning that will lead to a small number of "important" calendar frequencies and then demonstrate the validity of this reasoning in practice by several examples.

2.2.2 The Important Calendar Frequencies. Since $B + C(T)$ is a periodic function with period equal to 7 days, we can write

$$B + C(T) = \sum_{k=0}^{\infty} \gamma_k \cos\left(2\pi \frac{kT}{7} + \phi_k\right)$$

where γ_k is the amplitude of the cosine at frequency $k/7$ cycles/day and ϕ_k is the phase. Thus for the aggregated calendar effects

$$b(t) + c(t) = \sum_{k=0}^{\infty} \gamma_k h_k(t) \tag{2.2}$$

where

$$h_k(t) = \int_{T_{t-1}}^{T_t} \cos\left(2\pi \frac{kT}{7} + \phi_k\right) dT . \tag{2.3}$$

1. Values of Spectrum of $h_k(t)$ Greater Than .1

Frequency	Spectrum	k
.083	.100	0
.220	.151	2
.250	.210	0
.304	.157	2
.333	.161	0
.348	2.649	1
.402	.119	1
.416	.653	0
.432	.473	1
.500	.210	0

For two reasons the contributions in (2.2) for small k are the most important ones and those for larger k have a negligible effect. The first is that the spectrum of $h_k(t)$, which is derived in Section A.1 of the Appendix, becomes small for large k. Table 1 shows the important calendar frequencies for $b(t) + c(t)$, which are defined to be frequencies at which the spectrum of some $h_k(t)$ is greater than .1. Only values of k less than or equal to 2 appear. The second reason is an empirical result; for most weekly patterns, γ_k, which depends on the shape of the weekly pattern, will be small except for small values of k. For example, suppose $B + C(T)$ is a weekday indicator series, which means $B + C(T)$ equals 1 on weekdays and equals 0 on weekends. If $T = 0$ is taken to be Wednesday at noon, then $B + C(T)$ is symmetric about 0, so that $\phi_k = 0$,

$$\gamma_0 = \frac{1}{7}\int_{-3.5}^{3.5} (B + C(T))dT = \frac{5}{7},$$

and, for $k \geq 1$,

$$\gamma_k = \frac{2}{7}\int_{-3.5}^{3.5} (B + C(T)) \cos\left(2\pi\frac{k}{7}T\right) dT = 2\,\frac{\sin\left(2\pi\frac{k}{7}2.5\right)}{\pi k}.$$

Thus γ_k decreases rapidly for the weekday indicator series.

Let $h_{km}(y)$ be the mth monthly subseries of $h_k(t)$ so that $h_{km}(y) = h_k(12(y - 1) + m)$ and $b_m(y) = c_m(y) = \sum_{k=0}^{\infty} \gamma_k h_{km}(y)$. For the same reasons given in the previous paragraph, the spectrum of $b_m(y) + c_m(y)$ typically will be determined by $h_{km}(y)$ for small k. (The derivation of the spectra of $h_{km}(y)$ is given in Section A.2 of the Appendix.) For a fixed k the spectra of the $h_{km}(y)$ for m corresponding to 30-day months are the same, and the spectra for m corresponding to 31-day months are the same. Furthermore, the values of these two sets of spectra are similar. February, however, has a spectrum that is quite different from the other two. Thus the procedure in practice will be to estimate 11 monthly subseries spectra (excluding February) and look at their average. The important calendar frequencies for the average are shown in Table 2. These are defined to be the frequencies at which, for some k, the average of the 11 spectra of $h_{km}(y)$ is greater than .075.

A heuristic explanation can be given for the importance of the calendar frequencies .348 cycles/month and .179 cycles/year, which have the largest spectrum values in Tables 1 and 2, respectively. Suppose the lengths of all months were equal to

$$\frac{365.25}{12}\, days = 30.4375\, days\,.$$

Suppose a cosine with a period of 7 days is sampled every month. Then the sampled series has a frequency of

$$\frac{cycle}{7\ days} = \frac{30.4375/7\ cycles}{month} = 4.348\ cycles/month$$

and the alias of this frequency is .348 cycles/month. If the sampled series is now further sampled once per year, the resulting series has a frequency of

$$\frac{.348\ cycles}{month} = \frac{4.179\ cycles}{year}\,,$$

which aliases to .179 cycles/year.

2.3 Estimating the Spectrum of $x(t)$ and $x_m(y)$

2.3.1 Removing Trend and Seasonal. Often the aggregated monthly time series, $x(t)$, will contain strong trend and yearly seasonal components. It is generally desirable to remove these components from $x(t)$ before estimating the spectra, since the components tend to have a substantial influence on the estimates and can obscure the effects at the calendar frequencies. One procedure for removing the trend and seasonal is to decompose $x(t)$ into trend plus seasonal plus irregular by using a procedure such as SABL (Cleveland, Dunn, and Terpenning 1978 and 1979) or Census X-11 (Shiskin, Young, and Musgrave 1967) and then to estimate the spectra by using the irregular. A second procedure is to difference the series; if U is the shift operator $U_l x(t) = x(t - l)$, this means computing spectrum estimates from $(1 - U_1)^{d_1}(1 - U_{12})^{d_2}x(t)$ where d_1 and d_2 are chosen to remove the trend and seasonal, respectively (see Box and Jenkins 1976). In the examples in Sections 2 and 3 the trend and seasonal removal is carried out by using SABL with the length of the trend smooth equal to 7 and the length of the seasonal smooth equal to 11.

2. Values of Average Spectrum of $h_{km}(y)$ Greater Than .075

Frequency	Spectrum	k
.071	.124	1
.107	.078	2
.179	3.854	1
.286	.080	3
.357	.446	2
.429	.079	1
.464	.137	3

2.3.2 Transforming the Data. In many cases the trend, seasonal, and irregular are nonadditive, and the seasonal component has oscillations whose amplitudes increase as the trend increases. In the case in which $x(t) > 0$, this nonadditivity can be removed by a power transformation of the form

$$u^{(p)} = u^p \quad \text{for} \quad p > 0$$
$$= \log u \quad \text{for} \quad p = 0$$
$$= -u^p \quad \text{for} \quad p < 0 \ .$$

A procedure for selecting the value of p is given in Cleveland, Dunn, and Terpenning (1978 and 1979). Transforming $x(t)$ to make the seasonal oscillations more nearly stable is desirable because it not only facilitates the trend and seasonal removal but also generally leads to simpler models to explain the variation, both calendar and otherwise, in the series. In particular the transformation tends to result in the additivity of the calendar component as assumed in Section 2.2.1.

2.3.3 Clipping Outliers. Outliers in a series whose spectrum is to be estimated can have a substantially distorting effect on the estimate (Kleiner, Martin, and Thomson 1979). When a series with outliers is decomposed into trend, seasonal, and irregular by a routine such as SABL, which robustly estimates the trend and the seasonal, the outliers become part of the irregular. To guard against the distorting effect of outliers in a spectrum estimate from an irregular component, the irregular will be modified by the procedure described in the next paragraph, called "clipping."

As we have discussed in Section 2.2.1, for the 30-day months there is one of seven values that $b(t) + c(t)$ can take, which is determined by the day of the week on which the month starts, and similarly for the 31-day months there are seven values that $b(t) + c(t)$ can take. Furthermore, the two sets of seven values will typically be very similar. Thus the irregular will be divided into eight groups. The first seven groups are formed from the irregular, with February excluded, according to the starting day of the month. The eighth group consists of the February values. Now for a set of numbers the midmean is defined to be the average of all values between the quartiles, the upper (lower) semimidmean is defined to be the midmean of all values above (below) the median, and (Tukey 1977) one step = 1.5 (upper semimidmean minus lower semimidmean). For each group the semimidmeans and one step are computed. Any value in a group that is greater than the quantity, upper semimidmean + 1.5 steps, is set equal to this quantity; any value in the group that is less than the quantity, lower semimidmean − 1.5 steps, is set equal to this quantity.

2.3.4 Computing the Spectrum Estimate. Let $z(\nu)$, for $v = 1, \ldots, n$, be a series whose spectrum is to be estimated. One method of estimation begins by first subtracting the mean, $\bar{z}$, from $z(v)$ and then multiplying the result by a data window (Bloomfield 1976), $w(v)$, such as

$$w(v) = .5 - .5 \cos(\pi v/(r+1)) \ , \quad v = 1, \ldots, r$$
$$= 1 \ , \quad v = r+1, \ldots, n-r$$
$$= .5 - .5 \cos(\pi(n+1-v)/(r+1)) \ ,$$
$$v = n-r+1, \ldots, n \ .$$

(For the spectrum estimates in the examples of this article, r will be taken to be 20 percent of n.) Multiplying by $w(v)$ reduces the chance of a strong peak at one frequency causing spurious peaks at other frequencies. The spectrum estimate at frequency f is the squared modulus of the Fourier transform of the tapered data,

$$s(f) = \frac{1}{w}\left|\sum_{v=1}^{n} (z(v) - \bar{z})w(v) \exp(2\pi i f v)\right|^2$$

where $w = \sum_{v=1}^{n} w^2(v)$. Since we are looking for very narrow-band peaks in the spectrum at particular frequencies, it is generally not necessary to smooth $s(f)$ as is frequently done in spectrum estimation. $s(f)$ does not estimate the spectrum as defined in Section 2.1, but rather estimates a constant times the spectrum. However, since the constant does not depend on frequency and therefore does not interfere with the identification of peaks, we shall not adjust the spectrum estimates in the examples in Section 3.

2.4 Peak Picking

The procedure for identifying calendar effects in the data is to plot both the spectrum estimate from $x(t)$ and the average spectrum estimate from $x_m(y)$ and check for peaks at the calendar frequencies described in Section 2.2.2. Let us first consider the "peak picking" procedure for the spectrum estimate from $x(t)$ suggested by the values in Table 1. Each of the frequencies in the table with $k = 0$ is one of the yearly seasonal frequencies of the form $j/12$ cycles/month, for $j = 1, \ldots, 6$. However, peaks in the spectrum estimate at these seasonal frequencies would not often provide us with diagnostic information about the existence of calendar effects, for often the series will have yearly seasonal behavior due to causes other than the calendar effects. The removal of the yearly seasonal from $x(t)$, which removes the peaks in the spectrum estimate at the seasonal frequencies, does not reduce the effectiveness of the estimate at other calendar frequencies for detecting calendar effects. The largest spectrum value in Table 1 occurs for $k = 1$ at frequency .348 cycles/month. A peak in an estimate of the spectrum of $x(t)$ at this frequency is an almost certain indication of effects. (Some other cause is, of course, possible, but is highly unlikely). The next largest nonseasonal peak occurs at .432 cycles/month. These two frequencies, as our examples will illustrate, are the most important for diagnosing calendar effects in the spectrum estimate from $x(t)$.

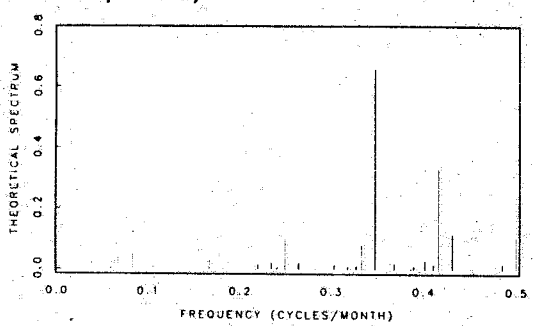
For the average spectrum estimate from $x_m(y)$, Table 2 shows that the two important frequencies for diagnosing calendar effects are .179 and .357 cycles/year. Thus peaks in the average spectrum estimate from $x_m(y)$ at these frequencies are an indication of a calendar effect.

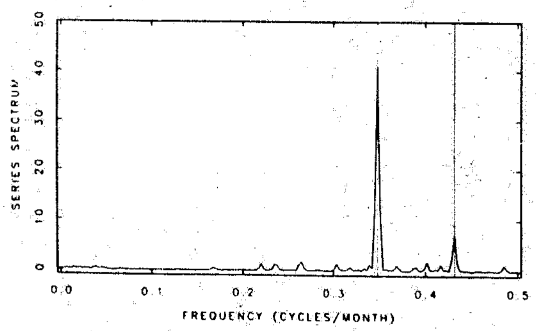
2.5 Examples

2.5.1 The Weekday Series. To illustrate the procedures of the previous sections, suppose $X(T)$ is the weekday indicator defined in Section 2.2.2. Thus $X(T)$ is a pure calendar series so that $X(T) = B + C(T)$ and $x(t)$ is simply the number of weekdays in a month. The spectrum of $x(t)$ is shown in the top panel of Figure A; the values of the spectrum at the yearly seasonal frequencies, $j/12$ cycles/month for $j = 1, \ldots, 6$, are shown by dotted lines. The average of the spectra of the 11 monthly subseries is shown in the top panel of Figure B.

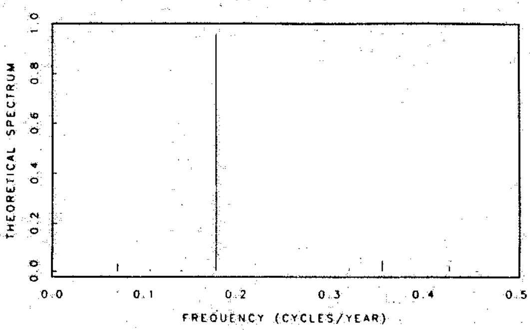
For January 1958 to December 1977, $x(t)$ was processed by the methods described in Section 2.3. First, $x(t)$ was decomposed into trend, seasonal, and irregular using the SABL decomposition procedure described in Section 2.3.1. (No transformation or clipping was used in this case.) The spectra were then estimated by the method described in Section 2.3.4 and plotted in the bottom panels of Figures A and B. Vertical lines have been drawn at the most important calendar frequencies (.348 and .432 cycles/month in Figure A; and .179 and .357 cycles/year in Figure B).

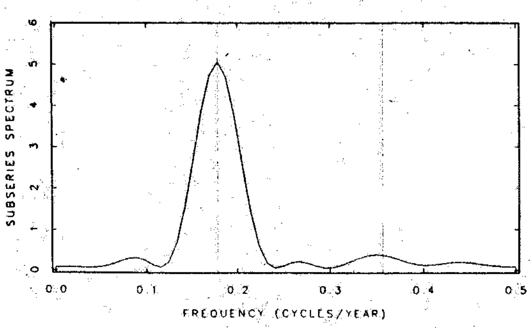
A. Spectrum of the Weekday Series (The top panel shows the theoretical spectrum of the series. The dotted lines are used for the values at the yearly seasonal frequencies. The bottom panel shows the spectrum estimate from the clipped irregular. The dotted vertical lines indicate the two important calendar frequencies.)





B. Spectrum of the Weekday Subseries (The top panel shows the average of the theoretical spectra. The bottom panel shows the estimate from the clipped irregular. The dotted vertical lines indicate the two important calendar frequencies.)

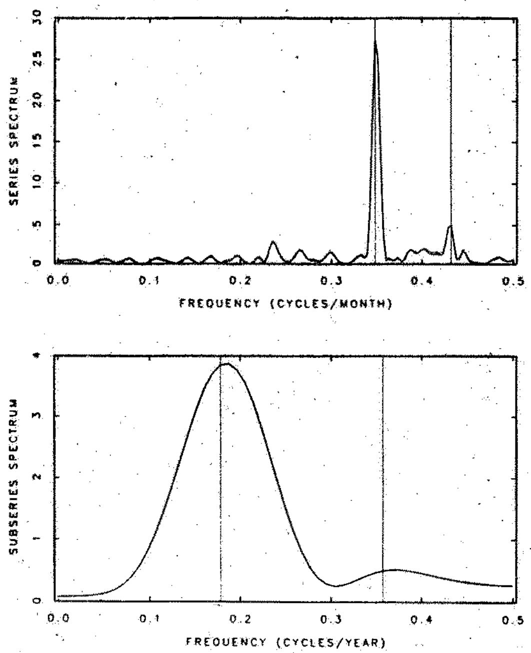




For the estimate from $x(t)$, peaks occur at the two important calendar frequencies. Were this a real series with an unknown structure it is clear that the estimate would serve to identify the calendar effects. The peaks at the seasonal frequencies are absent because the seasonal component has been removed. The estimate from $x_m(y)$ also clearly demonstrates the calendar effects since peaks occur at the two important calendar frequencies. This estimate, however, is not substantially affected by the seasonal component removal since a (nearly) periodic seasonal component in $x(t)$ contributes power in the individual subseries spectra at (low) zero frequency, which is removed by the subtraction of the mean.

2.5.2 Bell System Toll Revenues. The investigation of calendar effects in Bell System revenues from toll calls from January 1964 to October 1973 was begun by applying the transformation procedure described in Section 2.3.2. The value of p that most nearly stablizes the seasonal oscillations is .25. The transformed series was then decomposed into trend, seasonal, and irregular by using SABL. Figure C shows the spectrum estimates of

C. Spectrum Estimates From Clipped Irregular of Fourth Root Toll Revenues (The top panel shows the spectrum estimate from the series. The bottom panel shows the average spectrum estimate from the monthly subseries.)



the clipped irregular and the monthly subseries of the clipped irregular. Very large peaks occur at calendar frequencies, which indicates that a large fraction of the variation in the irregular is due to calendar effects. In fact, behavior at the calendar frequencies is quite similar to that of the spectrum estimates for the weekday series in Figures A and B.

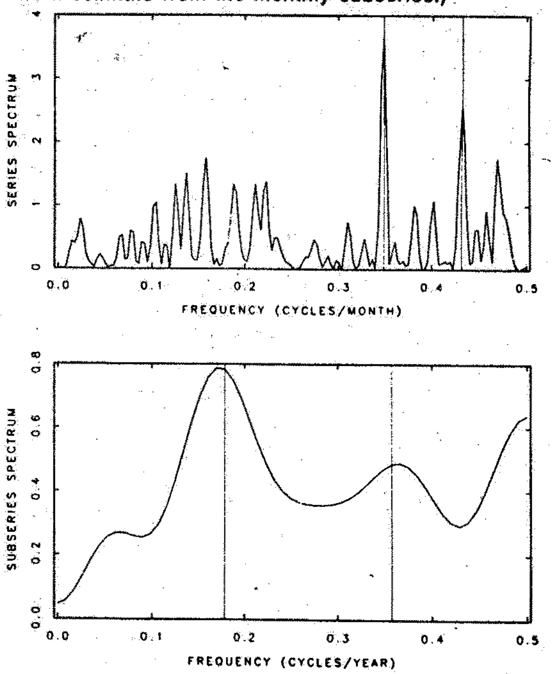
2.5.3 International Airline Series. The number of passengers on international flights from January 1949 to December 1960 has been used by Box and Jenkins (1976) and Brown (1962) to illustrate the use of forecasting methodology. The data were transformed to stabilize the seasonal oscillations by taking logarithms, and SABL was used to remove the seasonal and the trend. In Figure D the spectrum estimates from the clipped irregular and the 11 monthly subseries of the clipped irregular show calendar effects in the airline series, since peaks occur at all four of the important calendar frequencies. Thus the variation of the airline series is, in part, due to calendar effects, so that if these effects were incorporated into a model, they could be used to enhance forecasts because current and future values of the calendar effects are known.

2.5.4 Manufacturers' Shipments. The values of all shipments of manufactured durable goods in the United States are reported in Bureau of Economic Analysis (1976) and are calendar adjusted using X-11. Figure E shows the spectrum estimates of the logarithms of the calendar adjusted data from January 1960 to December 1974 using the clipped irregular from SABL. The peak at .348 cycles/month in the estimate of the top panel and the peaks at .179 and .357 cycles/year in the estimate of the bottom panel indicate there are residual calendar effects still in the series. Thus the X-11 calendar adjustment procedure does not appear to have successfully removed all of the calendar effects from the data.

3. TIME DOMAIN GRAPHICAL METHODS

Various time domain graphical displays of $x(t)$ and $x_m(y)$ can also be used for detection of calendar effects. The basic approach has been to design displays that allow the detection of the periodic behavior due to calendar effects, which we discussed in Section 2. The time domain displays are, in fact, not as powerful as the spectrum analyses for detecting the periodic behavior, but they are somewhat simpler to apply, and outliers and peculiar behavior at specific points in time can be detected. As with the spectrum analysis, we will generally not want to work with the aggregated monthly data but rather with the data after preanalysis by transformation and trend-seasonal removal.

D. Spectrum Estimates From Clipped Irregular of the Logarithms of International Airline Passengers (The top panel shows the spectrum estimate from the series. The bottom panel shows the average spectrum estimate from the monthly subseries.)



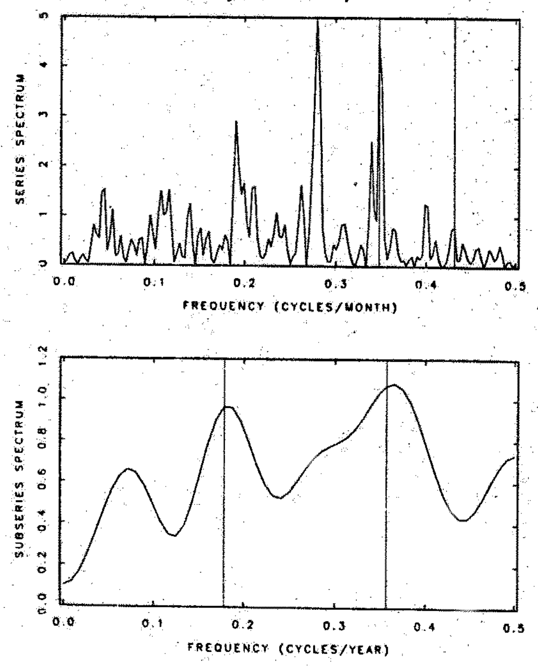
As discussed in Section 2.3.3, when a series with outliers is decomposed into trend, seasonal, and irregular by a routine such as SABL, which robustly estimates the trend and seasonal, the outliers become part of the irregular. Thus when an irregular component is displayed we do not want the outliers to cause a large increase in the scale of the plot and thereby substantially reduce the resolution of the calendar pattern. To prevent this, we shall modify the irregular component, before plotting, by the clipping procedure described in Section 2.3.3.

3.1 Seasonal Adjustment Plots

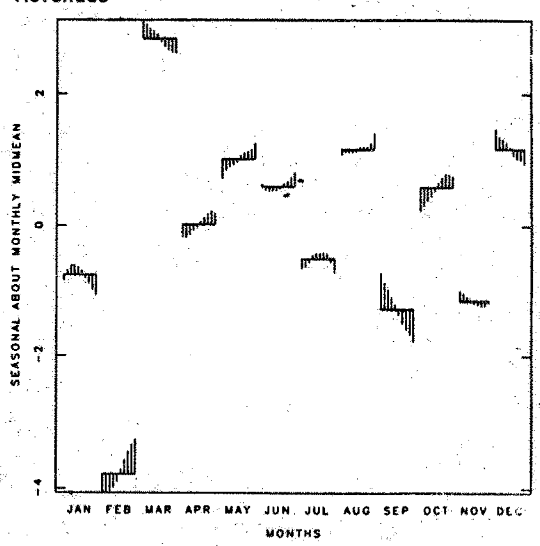
For situations in which the monthly data are decomposed into trend, seasonal, and irregular, certain plots that have already been proposed for enhancing seasonal adjustment procedures (Cleveland, Dunn, and Terpenning 1978 and 1979) can also serve to assist in detecting calendar effects. In this section we shall describe two of these plots.

3.1.1 Plots of the 12 Monthly Subseries of the Irregular. In Section 2.2.2 it was shown that a calendar effect typically results in oscillations with a period of 5.6 years in each of the monthly subseries of the irregular, excluding February. Thus plots of the monthly subseries of the clipped irregular can reveal the calendar effect if the effect accounts for a major part of the variation in the ir-

E. Spectrum Estimates From Clipped Irregular of the Logarithms of Manufacturers' Shipments (The top panel shows the spectrum estimate from the series. The bottom panel shows the average spectrum estimate from the monthly subseries.)



F. Seasonal-by-Month Plot for Fourth Root Toll Revenues



regular. For less pronounced effects, however, the calendar pattern in the irregular is more clearly revealed in the graphical display that will be described in Section 3.2.

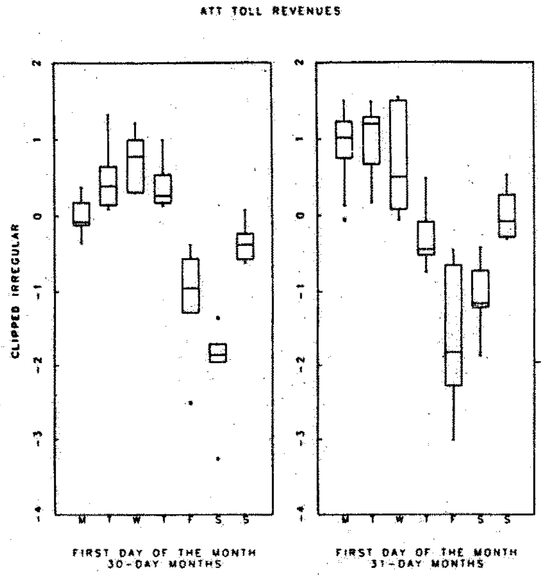
3.1.2 Seasonal-by-Month Plots. A second useful plot, called the seasonal-by-month plot, is one in which the seasonal for each month is plotted against year. An example is given in Figure F where the seasonal of the fourth root of Bell System toll revenues is displayed. The midmean for each month is portrayed by a horizontal line, and the seasonal component for that month is portrayed by vertical lines emanating from the horizontal line.

In Section 2.2.1 it was shown that the effect of aggregating the constant term, B, in $X(T)$ is to produce a component, $b(t)$, in $x(t)$ where $b(t)$ is the number of days in the month times B. Almost all of the variation in $b(t)$ is concentrated at the yearly seasonal frequencies, and therefore almost all of $b(t)$ is contained in the seasonal. The seasonal component will contain, of course, the yearly seasonal behavior not due to calendar causes; however, calendar effects can be detected in the seasonal-by-month plot by detecting a correlation between the level of each monthly subseries (as portrayed by the midmean) and the number of days in the month.

For example, in Figure F it is clear that the midmeans are highly correlated with month length so that almost all of the seasonal component of toll revenues consists of calendar effects. We have already seen in Section 2.5.2 that this is also true of the irregular. Thus apart from the overall trend the major factor in the variation of toll revenues is the calendar effect.

The task of detecting calendar effects is but one of many tasks that the seasonal-by-month plot allows one

G. Starting-Day-of-the-Month Plot for the Clipped Irregular of Fourth Root Toll Revenues



to carry out. (For example, evolution of the seasonal through time can be assessed.) If we were designing this display just for calendar detection, we would then quite likely use a different method (e.g., plotting midmeans versus monthly length). But in practice, because resources are finite, it is desirable to have multipurpose displays like the seasonal-by-month plot, which, while perhaps not the optimal display for detecting calendar effects in the seasonal, nevertheless appears to perform quite adequately for the examples on which we have tested it.

3.2 A Highly Tailored Calendar Plot

The special structure of the calendar effect can be exploited to design a plot that is quite powerful for identifying a calendar effect. As we have seen in Section 2.2.1, for the 30-day months there are seven values that $b(t) + c(t)$ can take, which are determined by the day of the week on which the month starts, and for the 31-day months there are seven values that $b(t) + c(t)$ can take. Thus a calendar effect can be detected by plotting, for each month (except February), the value of $x(t)$ against the day of the week on which the month starts. A dependence of the values on the day of the week indicates a calendar effect.

Since the patterns in the 30-day months are the same, the four plots can be combined into one. Similarly, the seven 31-day months can be combined into one. This is done in Figure G for the clipped irregular of the fourth roots of Bell System toll revenues. For each starting day of the month a schematic plot (Tukey 1977) is used to summarize the distribution of the values. The line inside the rectangle of the schematic plot is the midmean and the upper and lower edges of the rectangle are the upper semimidmean and the lower semimidmean, respectively. The top of the vertical line extending from the top of the box is the largest data point that is within one step of the upper semimidmean. All points beyond one step are plotted individually. A similar procedure is followed for the bottom of the diagram. (The midmean, semimidmeans, and step are defined in Section 2.3.3.) In Figure G it is quite clear that the distributions are very different for different days of the week, which indicates a strong calendar effect.

4. SUMMARY

Spectrum analyses and time domain graphical methods have been used to detect calendar effects in monthly aggregated time series. First, these methods can be used in an initial analysis to decide if calendar adjustment is necessary. For example, it was shown that revenues from toll calls and the number of international airline passengers have significant calendar effects. The methods also can be used on an adjusted series to determine if the adjustment has properly removed all of the calendar effects. For example, it was shown that the manufacturers' shipments series, which was adjusted by the X-11 procedure, has residual calendar effects.

Let $x(t)$ be the aggregated monthly subseries and let $x_m(y)$, for $m = 1, \ldots, 12$, be the 12 monthly subseries of $x(t)$. Two types of spectrum estimates are computed in the following manner:

1. Transform $x(t)$.
2. Remove the trend and seasonal from the transformed $x(t)$.
3. Clip outliers.
4. Compute the spectrum estimate of the resulting series in (3).
5. Compute the average spectrum estimate of the 11 monthly subseries (excluding February) of the resulting series in (3).

A peak in the spectrum estimate from (4) at .348 or .432 cycles/month, or a peak in the spectrum estimate from (5) at .179 or .357 cycles/month, indicates the presence of calendar effects.

Two time domain graphical methods are also useful for detecting effects. One is to plot each of the 12 monthly subseries of the seasonal component of $x(t)$ and look for a correlation between the levels of the subseries and the length of the months. The second graphical method is to plot each value of the clipped irregular (with February omitted) against the day of the week on which the month starts and look for a pattern in the locations of the seven sets of values. The time domain graphics are not so sensitive a tool for detecting calendar effects as the spectrum estimates, but the graphics, when they reveal the effects, do give information about the nature of the calendar component and are somewhat simpler to apply.

APPENDIX: SPECTRUM OF $h_k(t)$ AND $h_{km}(t)$

A.1 The Spectrum of $h_k(t)$

Let $T_0 = 0$ and let $d(t)$ be the number of days in the tth month so that

$$T_t = \sum_{j=1}^{t} d(j)$$

Let

$$\bar{d} = \frac{1}{48}\sum_{t=1}^{48} d(t) = 30.4375$$

be the average number of days per month. Since $d(t)$ has period 48 months (i.e., $d(t) = d(t+48)$),

$$r_t = T_t - \bar{d}t$$

has period 48 months.

For $k = 0$, we have from (2.3) that $h_0(t) = d(t)$. Since $d(t)$ has period 48 the spectrum is given by (2.1). For $k \geq 1$, we have from (2.3)

$$h_k(t) = \frac{\sin(\omega_k T_t + \phi_k) - \sin(\omega_k T_{t-1} + \phi_k)}{\omega_k} \quad \text{(A.1)}$$

where

$$\omega_k = \frac{2\pi k}{7}.$$

Let

$$g_k(t) = \exp(i\omega_k \bar{d}t)[\exp(i\omega_k r_t) - \exp(i\omega_k(r_{t-1} - \bar{d}))]$$

then

$$h_k(t) = (2\omega_k i)^{-1}[g_k(t)\exp(i\phi_k) - \bar{g}_k(t)\exp(-i\phi_k)] .$$

Since r_t has period 48 we can write

$$\exp(i\omega_k r_t) = \sum_{j=0}^{47} \xi_{kj} \exp(i\theta_j t)$$

where

$$\theta_j = \frac{2\pi j}{48}$$

and

$$\xi_{kj} = \frac{1}{48}\sum_{t=0}^{47} \exp(i\omega_k r_t)\exp(-i\theta_j t) .$$

Thus

$$g_k(t) = \sum_{j=0}^{47} \zeta_{kj} \exp(2\pi i f_k(j) t)$$

where $f_k(j)$ is the alias of

$$\frac{\omega_k \bar{d} + \theta_j}{2\pi}$$

and

$$\zeta_{kj} = \xi_{kj}(1 - \exp(-2\pi i f_k(j)))$$

Thus

$$h_k(t) = \frac{1}{\omega_k}\sum_{j=0}^{47} im(\zeta_{kj}\exp(i\phi_k))\cos 2\pi f_k(j)t + \text{re}(\zeta_{kj}\exp(i\phi_k))\sin 2\pi f_k(j)t .$$

Thus the spectrum of $h_k(t)$ at frequency $f_k(j)$ is

$$\frac{|\zeta_{kj}|^2}{\omega_k^2} .$$

A.2 The Spectrum of $h_{km}(y)$

For $k = 0$, $h_{0m}(y)$ is the number of days in the mth month of the yth year. Thus, except for February, $h_{0m}(y)$ is constant (as a function of y) so that the spectrum is concentrated at zero frequency. For February, $h_{0m}(y)$ is periodic with period 4.

Suppose $k \geq 1$. Let $b_m(y)$ be the time (in units of days) at the end of the mth month in the yth year and let $a_m(y)$ be the time at the beginning of the month. For $k \geq 1$, we have from (A.1)

$$h_{km}(y) = \frac{\sin(\omega_k b_m(y) + \phi_k) - \sin(\omega_k a_m(y) + \phi_k)}{\omega_k} . \quad \text{(A.2)}$$

Let $b'_m(y)$ be $b_m(y)$ reduced modulo 7 to one of the integers 0, ..., 6. Define $a'_m(y)$ similarly. Then $h_{km}(y)$ may be written as in (A.2) but with $a_m(y)$ and $b_m(y)$ replaced by $a'_m(y)$ and $b'_m(y)$, respectively. Since leap year occurs every fourth year, $a'_m(y)$ and $b'_m(y)$ consist of the integers (mod 7) with every fourth value removed. For example, if the month is March, if $b'_m(1) = 1$, and if year 4 is a leap year, then the sequence begins with 1, 2, 3, 5, 6, 0, 1, 3, 4.... Thus $a'_m(y)$ and $b'_m(y)$ are periodic with period 28 and, consequently, so is $h_{km}(y)$. If the month has 31 days, then $b'_m(y) = a'_m(y) + 3$ (mod 7), whereas if the month has 30 days, $b'_m(y) = a'_m(y) + 2$ (mod 7). Thus the spectra for all the 31-day monthly subseries are equal and the spectra for all the 30-day montly subseries are equal. February stands alone.

Since, for $k \geq 0$, $h_{km}(y)$ has period 28, we have from (2.1) that the spectrum is concentrated at the frequencies $j/28$, $j = 1, \ldots, 14$. The procedure of averaging the 11 spectrum estimates from the monthly subseries of $x(t)$, excluding February, is justifiable because the difference between the 30-day and 31-day theoretical spectra is not large. February has been excluded since its values are substantially different.

[*Received May 1979. Revised March 1980.*]

REFERENCES

Bloomfield, Peter (1976), *Fourier Analysis of Time Series: An Introduction*, New York: John Wiley & Sons.

Box, George E.P., and Jenkins, Gwilym M. (1976), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day.

Brown, R.G. (1962), *Smoothing, Forecasting and Prediction of Discrete Time Series*, Englewood Cliffs, N.J.: Prentice-Hall.

Bureau of Economic Analysis (1976), *Business Statistics 1975*, U.S. Dept. of Commerce, Washington, D.C.: U.S. Government Printing Office.

Cleveland, William S., and Devlin, Susan J. (1980), "Calendar Effects in Monthly Time Series: Modeling and Adjustment," Bell Laboratories memorandum.

Cleveland, William S., Dunn, Douglas M., and Terpenning, Irma J. (1978), "The SABL Seasonal Analysis Package—Statistical and Graphical Procedures," Bell Laboratories, Murray Hill, N.J.: Computing Information Service.

Cleveland, William S., Dunn, Douglas M., and Terpenning, Irma J. (1979), "SABL—A Resistant Seasonal Adjustment Procedure With Graphical Methods for Interpretation and Diagnosis," in *Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, U.S. Dept. of Commerce, Bureau of the Census.

Eisenpress, Harry (1956), "Regression Techniques Applied to Seasonal Corrections and Adjustments for Calendar Shifts," *Journal of the American Statistical Association*, 51, 615-620.

Granger, C.W.J. (1963), "The Effect of Varying Month-Length on the Analysis of Economic Time Series," *Milano*, 41-53.

Kleiner, B., Martin, R.D., and Thomson, D Estimation of Power Spectra," *Journal of Society*, Ser. B, 41, 313-351.

Shiskin, Julius, Young, Allan H., and Musgr "The X-11 Variant of the Census Method II Program," *Bureau of the Census Technical* Commerce.

Tukey, John W. (1977), *Exploratory Data Ana* Addison-Wesley Publishing Co.

Young, Allan (1965), "Estimating Tradin Monthly Economic Time Series," Technical of the Census.

Zellner, Arnold (ed.) (1979), *Seasonal Analy Series*, U.S. Dept. of Commerce, Bureau of

Notes* II

MANAGEMENT SCIENCE
Vol. 26, No. 1, January 1980
Printed in U.S.A.

ANALYSIS OF TIME SERIES WITH CALENDAR EFFECTS†

LON-MU LIU‡

Most national economic data and many marketing series are compiled monthly according to the Gregorian Calendar, but some of the ancient festivals or holidays, such as Easter, Jewish Passover, and Chinese New Year, are set by lunar calendar. Therefore the date of a holiday may vary between two adjacent months from year to year. Some product marketing and consumer behavior patterns are closely related to such holidays: the amount of monthly sales may change as the date of the holiday changes from one month to the other. This paper shows how calendar intervention influences the sample autocorrelation, both by a theoretical study and an actual example. Model estimation for a time series subject to calendar intervention is also discussed.
(FORECASTING ARIMA PROCESSES; FORECASTING TIME SERIES)

1. Introduction

Many marketing data recorded in the form of time series are subject to various interventions (such as policy changes, holidays, wars, or terrorist attacks) which may make the usual autoregressive-integrated moving average (ARIMA) models inadequate to describe the data. Box and Tiao [2] and Glass, Wilson and Gottman [4] propose models that seem to provide reasonably accurate descriptions of many important kinds of interventions. This article discusses one kind of intervention that cannot be described by the models reported. Most national economic data and many marketing series are compiled monthly according to the Gregorian Calendar, but some of the ancient festivals or holidays, such as Easter, Jewish Passover, and Chinese New Year, are set by lunar calendar. Therefore the date of a holiday may vary between two adjacent months from year to year. For example, Easter, which always falls on a Sunday following the fourteenth day of the Paschal Moon, may fall either in March or in April. Some product marketing and consumer behavior patterns are closely related to such holidays, hence the amount of monthly sales may change as the date of the holiday changes from one month to the other. We refer to this kind of change as calendar effects.

Below, we present a theoretical study of how the model identification of a time series is affected by calendar effects. We then study a set of actual data, the Taiwan highway traffic volume recorded by the Taiwan Highway Bureau. The highway traffic volume in Taiwan is greatly affected by the Chinese New Year which varies between January 21 and February 19 in the Gregorian calendar (it is always January 1 in the Chinese lunar calendar).

2. A Theoretical Study

The sample autocorrelation functions (SACF's) of a time series are very useful in identifying models for a time series. Here we study the expectation of the SACF's of a time series when calendar effects exist.

* All Notes are refereed.

† Accepted by Ambar G. Rao; received February 1, 1979. This paper has been with the author 2 months for 1 revision.

‡ University of California, Los Angeles

0025-1909/80/2601/0106$01.25

Consider the following general intervention model

$$z_t = f(\beta, \xi_t) + \frac{\theta(B)}{D(B)\phi(B)} a_t \qquad (Bz_t = z_{t-1}), \tag{1}$$

$$a_t \sim \text{iid } N(0, \sigma_a^2), \qquad t = 1, 2, \ldots, N,$$

where z_t is an appropriate transformation of the observed time series; $f(\beta, \xi_t)$ describes the calendar effects at time t in which ξ_t represents the calendar intervention and β is a vector of parameters reflecting the calendar effects; $\phi(B)$ is the autoregressive, $\theta(B)$ the moving-average, and $D(B)$ the difference operators (Box and Jenkins [1]). $\phi(B)$, $\theta(B)$, and $D(B)$ may be multiplicative operators.

The model in (1) can be rewritten as

$$w_t = \eta_t + \epsilon_t, \qquad t = 1, 2, \ldots, n,$$

where $w_t = D(B)z_t$, $\eta_t = D(B)f(\beta, \xi_t)$, $\epsilon_t = \theta(B)/\phi(B)a_t$, and $n = N -$ the highest order of B in $D(B)$.

The SACF of the kth lag of w_t is defined as

$$r_k = c_k / c_0, \qquad k = 0, 1, \ldots, K,$$

where

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (w_t - \bar{w})(w_{t+k} - \bar{w}),$$

$$\bar{w} = \frac{1}{n} \sum_{t=1}^{n} w_t.$$

The c_k can also be expressed as

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} w_t w_{t+k} - \bar{w}^2 - \frac{k}{n} \bar{w}^2 + \frac{\bar{w}}{n} \left(\sum_{t=1}^{k} w_t + \sum_{t=n-k+1}^{n} w_t \right).$$

The expectation of $w_t w_{t+k}$ is

$$E(w_t w_{t+k}) = \eta_t \eta_{t+k} + \rho_k, \qquad k = 0, 1, 2, \ldots,$$

where ρ_k is the kth autocorrelation of the ARIMA process ϵ_t.

Therefore, we have

$$\frac{1}{n} E \sum_{t=1}^{n-k} w_t w_{t+k} = \frac{1}{n} \sum_{t=1}^{n-k} \eta_t \eta_{t+k} + \left(1 - \frac{k}{n}\right) \rho_k,$$

$$E\bar{w}^2 = \bar{\eta}^2 + \frac{\rho_0}{n^2} + \frac{2}{n} \sum_{j=1}^{n-1} \rho_j - \frac{2}{n^2} \sum_{j=1}^{n-1} j\rho_j, \qquad \bar{\eta} = \frac{1}{n} \sum_{t=1}^{n} \eta_t,$$

$$E\bar{w}\left(\sum_{t=1}^{k} w_t \right) = \bar{\eta}\left(\sum_{t=1}^{k} \eta_t \right) + \frac{1}{n} \sum_{j=1}^{k} \sum_{l=1}^{n} \rho_{|l-j|},$$

$$E\bar{w}\left(\sum_{t=n-k+1}^{n} w_t \right) = \bar{\eta}\left(\sum_{t=n-k+1}^{n} \eta_t \right) + \frac{1}{n} \sum_{j=n-k+1}^{n} \sum_{l=1}^{n} \rho_{|l-j|}.$$

Hence

$$Ec_k = \frac{1}{n}\sum_{t=1}^{n-k}(\eta_t - \bar{\eta})(\eta_{t+k} - \bar{\eta}) + \rho_k - \frac{k}{n}\rho_k - \frac{2}{n}\sum_{j=1}^{n-1}\rho_j - \frac{\rho_0}{n^2} + \frac{2}{n^2}\sum_{j=1}^{n-1} j\rho_j$$
$$+ \frac{1}{n^2}\sum_{j=1}^{k}\sum_{l=1}^{n}\rho_{|l-j|} + \frac{1}{n^2}\sum_{j=n-k+1}^{n}\sum_{l=1}^{n}\rho_{|l-j|} - \frac{2k}{n^2}\sum_{j=1}^{n-1}\rho_j$$
$$- \frac{k}{n^3}\rho_0 + \frac{2k}{n^3}\sum_{j=1}^{n-1} j\rho_j.$$

Since the series ϵ_t is stationary, the above expression can be simplified to

$$Ec_k \doteq \frac{1}{n}\sum_{t=1}^{n-k}(\eta_t - \bar{\eta})(\eta_{t+k} - \bar{\eta}) + \rho_k, \qquad k = 0, 1, 2, \ldots,$$

if n is relatively larger than k.

When the sample size n is large, the expectation of r_k can be roughly approximated by $E(c_k)/E(c_0)$.

As an example, we consider a special case of (1)

$$z_t = \beta\xi_t + \frac{(1-\theta B)(1-\Theta B^{12})}{(1-B)(1-B^{12})}a_t. \tag{2}$$

The ARIMA model in (2) is denoted as $(0, 1, 1) \times (0, 1, 1)_{12}$ in Box and Jenkins [1]. This ARIMA model, which provides a useful basis for a popular forecasting technique, exponential smoothing (Brown [3] and Box and Jenkins [1]), is particularly useful in economic and market forecasting.

The expectation of c_k, $k = 0, 1, 2, \ldots$, for the model in (2) can be expressed as

$$\begin{aligned} E(c_0) &\doteq \beta^2\lambda_0 + (1+\theta^2)(1+\Theta^2)\sigma_a^2, \\ E(c_1) &\doteq \beta^2\lambda_1 - \theta(1+\Theta^2)\sigma_a^2, \\ E(c_{11}) &\doteq \beta^2\lambda_{11} + \theta\Theta\sigma_a^2, \\ E(c_{12}) &\doteq \beta^2\lambda_{12} - \Theta(1+\theta^2)\sigma_a^2, \\ E(c_{13}) &\doteq \beta^2\lambda_{13} + \theta\Theta\sigma_a^2, \end{aligned} \tag{3}$$

and

$$E(c_k) \doteq \beta^2\lambda_k \quad \text{for } k = 2, 3, \ldots, 10, 14, 15, \ldots$$

where

$$\lambda_k = \frac{1}{n}\sum_{t=1}^{n-k}(\zeta_t - \bar{\zeta})(\zeta_{t+k} - \bar{\zeta}), \quad \bar{\zeta} = \frac{1}{n}\sum_{t=1}^{n}\zeta_t \quad \text{and}$$
$$\zeta_{t-13} = (1-B)(1-B^{12})\xi_t.$$

In most cases a holiday varies between two adjacent months. Hence except for λ_{12s}, λ_{12s+1}, λ_{12s+2}, λ_{12s+10}, λ_{12s+11}, $s = 0, 1, 2, \ldots$, the rest of λ_k's are all equal to zero. As a result we may find that the corresponding SACF's are significant if calendar effects exist. However if calendar effects do not exist, only r_1, r_{11}, r_{12} and r_{13} may be significant. From (3), we can readily see that if β^2 is dominated by σ_a^2, the SACF's of the series may behave like the SACF's of an ARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$ model. On the other hand, if σ_a^2 is dominated by β^2, the SACF patterns will be determined by ξ_t.

The above discussion concentrates on studying the SACF's of $(1-B)(1-B^{12})z_t$. For the SACF's of z_t, $(1-B)z_t$, and $(1-B^{12})z_t$ the ARIMA component has a more pronounced influence on the behavior of the SACF's than the $(1-B)(1-B^{12})z_t$ series, since the models for the first three situations are nonstationary. Thus the SACF patterns are dominated by the ARIMA component (unless the calendar effects are very large), but they may still be seriously disturbed by calendar effects.

3. The Taiwan Highway Bureau Data

The data we study in this paper are the monthly highway traffic volume in Taiwan from 1963 to 1976 (listed in Table 1). By examining the data, we readily see that the Chinese New Year causes the highest peak of monthly traffic volume each year. The SACF's of z_t, $(1-B)z_t$, $(1-B^{12})z_t$, and $(1-B)(1-B^{12})z_t$, plotted in Figure 1, suggest that this series may be stationary after a regular difference. However, the pattern of the SACF's of $(1-B)\dot{z}_t$ is puzzling: it has significant SACF's at lags 1, 11 and 36 and shows no indication of a twelve month seasonality. It is clear that the pattern of the SACF's in the Taiwan Highway Bureau (THB) series is several disturbed by the calendar effects. We next eliminate the calendar effects by employ the traffic pattern during the new year period and then reexamine the SACF's of the adjusted THB series $\dot{z}_t$. The SACF's of $\dot{z}_t$, $(1-B)\dot{z}_t$, $(1-B^{12})\dot{z}_t$, and $(1-B)(1-B^{12})\dot{z}_t$ are shown in Figure 2. It is clear that the adjusted THB series follows an ARIMA $(0,1,1)\times(0,1,1)_{12}$ model (Box and Jenkins [1]).

TABLE 1

Monthly Highway Traffic Volume in Taiwan, 1963–1976
(Unit = 1000 Passenger-kms)

Year	Jan.	Feb.(*)	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1963	197702	154489	168969	162248	156900	149465	173652	183419	187842	199369	189815	205281
1964	195572	230525	219845	212869	211875	196445	203394	209030	225356	220589	225318	225991
1965	229962	258089	235446	230609	234150	212465	220186	227071	247999	246561	235427	247279
1966	296647	235722	251079	249428	243214	231728	240803	245398	257509	268385	253727	259684
1967	249532	288670	269473	263403	256899	239869	252670	247215	262563	256164	252619	257952
1968	292426	264250	282427	279957	282792	270086	261567	261599	267941	289789	281790	289934
1969	281933	321661	301941	296187	297104	284177	274190	279816	279447	282332	285270	292253
1970	289827	327579	305289	303868	301853	288149	287871	291529	309490	307599	303736	306289
1971	354077	305131	335761	323854	334469	298160	306387	326190	322680	349472	328692	341173
1972	334944	368312	365549	361528	350110	339488	346461	346287	377658	385497	361051	388486
1973	363823	428905	406316	407777	393267	397980	418055	418397	450295	423975	436052	454548
1974	522497	429506	481111	464936	451695	439111	457346	470782	496502	472891	464752	480253
1975	466108	503289	500633	476999	493942	475436	495321	498209	511880	515475	509336	520425
1976	530991	573337	552652	536038	544801	522590	550301	569753	579638	567229	550223	566217

(*) The traffic volume for February in a leap year has been adjusted to 28 days for the month.

TABLE 2

Results of Estimation for THB Series
(Standard errors shown in parentheses)

Estimates \ Models	$\epsilon_t = \dfrac{(1-\theta B)(1-\Theta B^{12})}{(1-B)(1-B^{12})}a_t$	
	$z_t = \epsilon_t$	$z_t = \beta_1\xi_{1t} + \beta_2\xi_{2t} + \epsilon_t$
β_1		44264.4 (5904.1)
β_2		1486.9 (692.4)
θ	0.7099 (0.054)	0.5346 (0.068)
Θ	0.9040 (0.021)	0.8551 (0.035)
S. S. of residuals	3.5581×10^{10}	1.5817×10^{10}
d. f.	153	151
Standard error	15249.7	10234.5

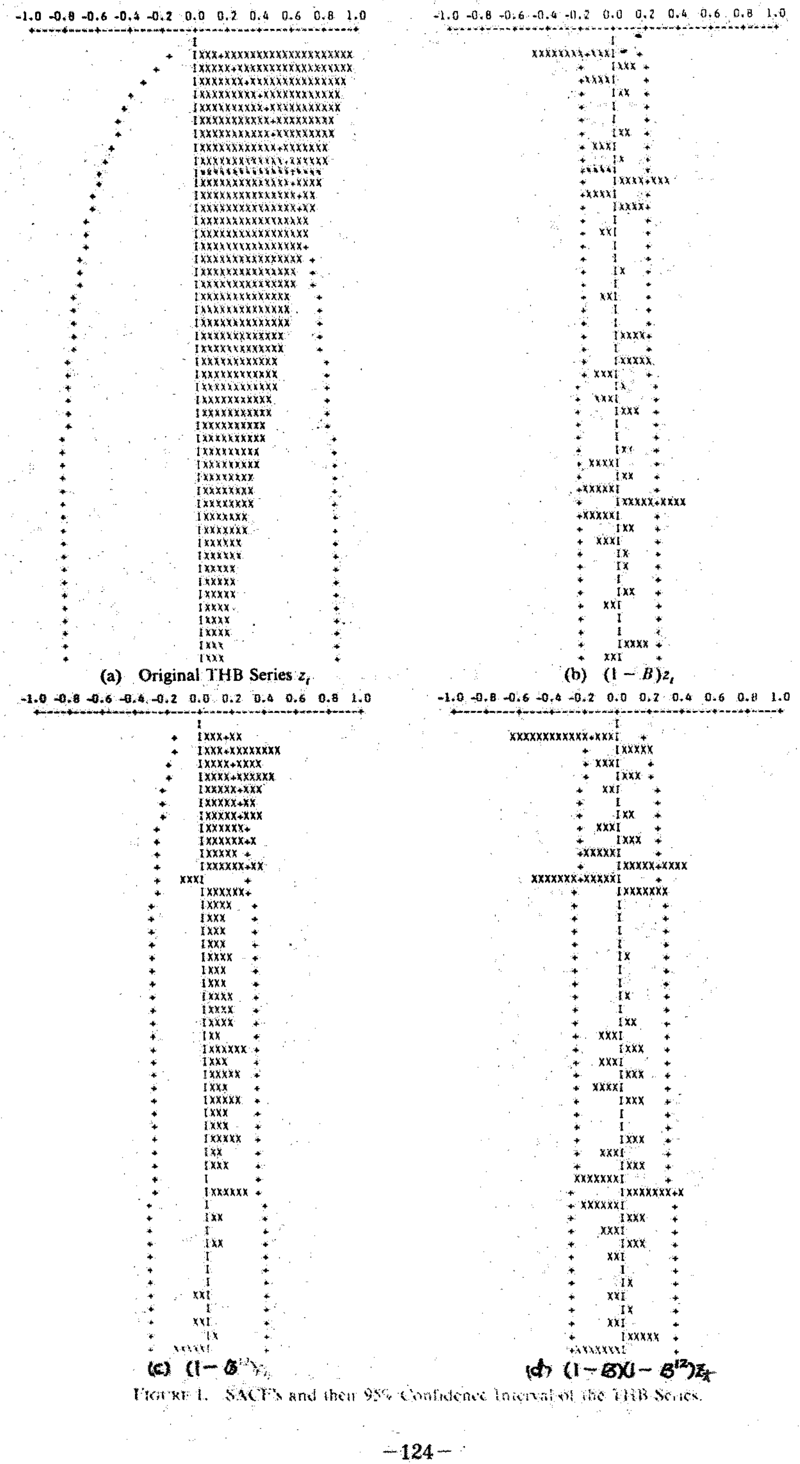


FIGURE 1. SACF's and their 95% Confidence Interval of the THB Series.

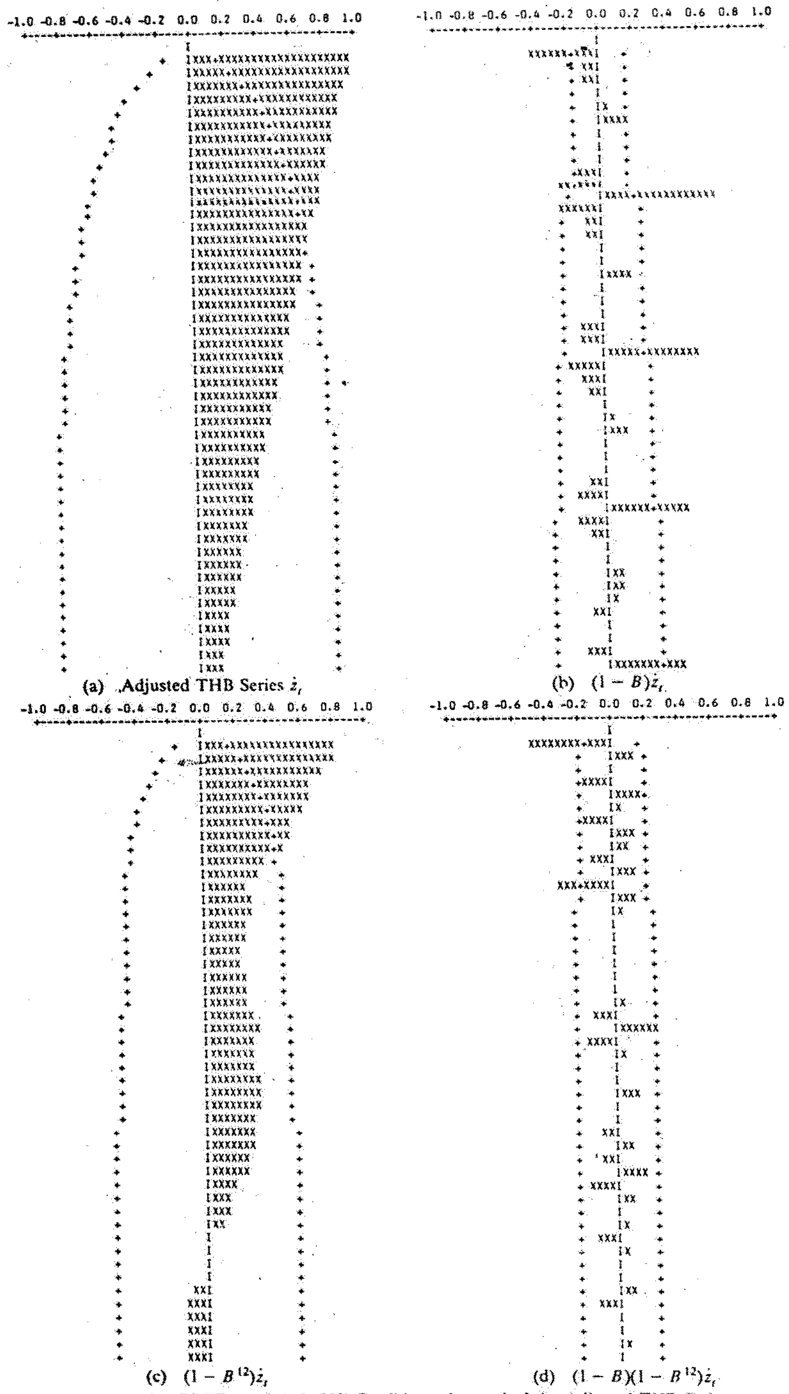


FIGURE 2. SACF's and their 95% Confidence Interval of the Adjusted THB Series.

An examination of the approximate calendar effects reveals that the traffic volume due to the Chinese New Year increases linearly. The following model may represent the calendar effects adequately:

$$f(\beta, \xi_t) = \beta_1 \xi_{1t} + \beta_2 \xi_{2t} \qquad (4)$$

where $\xi_{1t} = \xi_t$, $\xi_{2t} = \xi_t \times (I\text{-}1962)$ with I representing the year and ξ_t the proportion of the new year period in the tth month. Using the calendar effect model in (4) and the ARIMA model in (2), we obtain the set of parameter estimates listed in Table 2. For the purposes of comparison we may assume $\beta_1 = \beta_2 = 0$ and study the reduced model. The result of this model is also listed in Table 2. From these results, we find that the calendar effects component reduces 56% of the sum of squares of the residuals.

4. Concluding Remarks

In building a time series model, it is usually assumed that interventions have little effect on the overall SACF patterns of the series. In this article we show that calendar intervention may be significant enough to completely disturb the SACF patterns. When calendar intervention is present, a preliminary adjustment of the series is necessary before the identification of a model. This paper presents a comprehensive way to identify and estimate the model of a series that is subject to calendar intervention.[1]

[1] The author wishes to thank Professor George C. Tiao and referees for helpful comments and suggestions and Lyda Boyer for editorial assistance on this article. This research was supported in part by National Science Foundation Grant MCS78-16827.

References

1. Box, G. E. P. and Jenkins, G. M., *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, Calif., 1970.
2. ——— and Tiao, G. C., "Intervention Analysis with Application to Economic and Environmental Problems," *J. Amer. Statist. Assoc.*, Vol. 70 (1975), pp. 70–79.
3. Brown, R. G., *Smoothing, Forecasting and Prediction*, Prentice Hall, Englewood Cliffs, N.J., 1963.
4. Glass, G. V., Wilson, V. L. and Gottman, J. M., *Design and Analysis of Time Series Experiments*, Colorado Associated University Press, 1975.

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

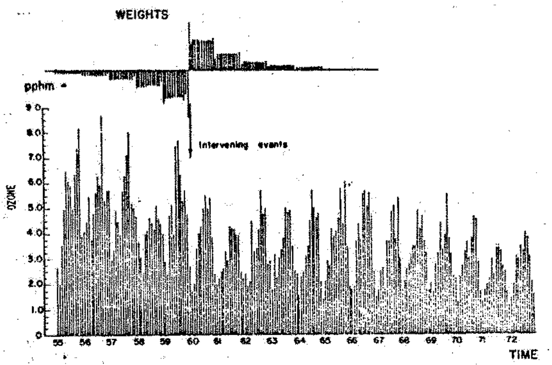
Data of potential value in the formulation of public and private policy frequently occur in the form of time series. Questions of the following kind often arise: "Given a known intervention,[1] is there evidence that change in the series of the kind expected actually occurred, and, if so, what can be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

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

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

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;
$f(\kappa, \xi, t)$ can allow for deterministic effects of time, t, the effects of exogenous variables, ξ, and in particular, interventions;
N_t represents stochastic background variation or noise;
κ 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 $N_t = y_t - f(\kappa, \xi, t)$ 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

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

where the roots of $\phi(B)$ 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 $\varphi_2(B^s) = (1 - B^s)^D\phi_2(B^s)$ to allow for seasonal nonstationarity.

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

$$\phi_1(B)\phi_2(B^s)(1 - B)^d(1 - B^s)^D N_t = \theta_1(B)\theta_2(B^s)a_t \tag{2.3}$$

where the polynomials $\phi_1(B)$, $\phi_2(B^s)$, $\theta_1(B)$, $\theta_2(B^s)$ are of degrees p_1, p_2, q_1, q_2, respectively.

3. A DYNAMIC MODEL FOR INTERVENTION

Frequently the effects of exogenous variables ξ can be represented by a dynamic model of the form

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

where:

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

$\delta_j(B) = 1 - \delta_{1j}B - \cdots - \delta_{r_j j}B^{r_j}$ 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

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

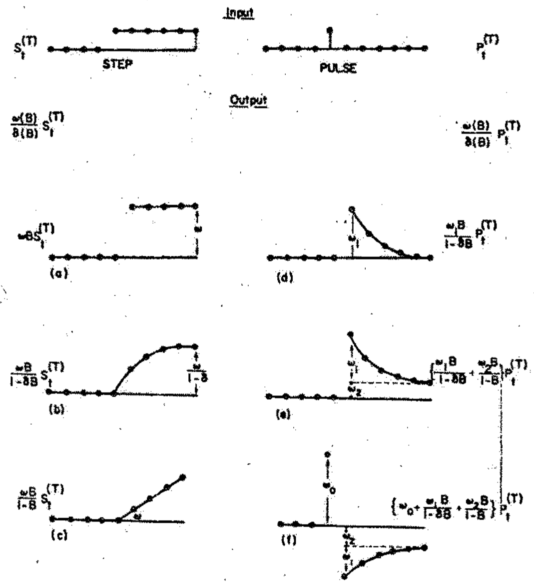
might do this (Figure B(d)) with ω_1 indicating the initial increase in sales immediately following the intervention and δ representing the rate of decay of this increase.

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), (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

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

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,

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

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;

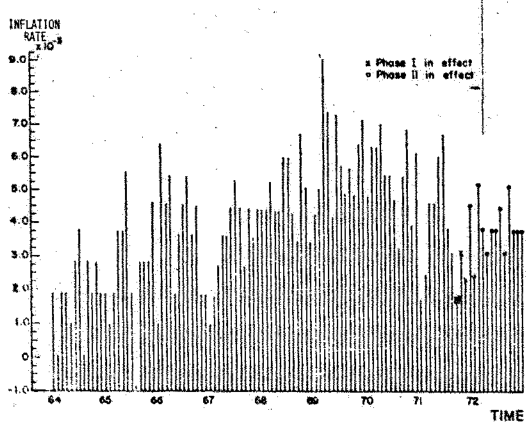
(ii) associated with I_2 there is a progressive reduction in O_3. Over the period studied, there is a yearly increment of approximately $\hat{\omega}_{02} = -.25$ in the summer months, but the increment (if any) in the winter is slight.

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

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



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.

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^2	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 $\quad (6.4)$

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

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

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

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

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

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

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

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

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

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

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). $\quad \omega(B)/\delta(B) = \beta B/(1 - B)$. $\quad$ (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

$$Q(B) = (1 - B)R(B)F \tag{6.14}$$

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

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

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

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 inte-vening 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

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

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,

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

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

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 $j = 12n + m$, we find

$$\pi_{12n+m} = (1 - \theta_1)^{-1}(\phi - \theta_2)^{-1}[(\theta_1)^{m+1}\{(1 - \phi)\phi^n - (1 - \theta_2)(\theta_2)^n\} + (\phi - \theta_2)(\theta_2)^n] \ ,$$

$$m = 0, \cdots, 11; n = 0, \cdots, \infty \tag{6.18}$$

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

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

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

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

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

$$(\varphi(B)/\theta(B))(\omega_j(B)/\delta_j(B)) = \beta_j R_j(B) \tag{6.20}$$

we can write

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

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

$$\mathbf{A}\hat{\boldsymbol{\beta}} = \mathbf{b} \qquad (6.22)$$

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,

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

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(c) 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 $|\mathbf{A}| \neq 0$,

$$\left.\begin{aligned}\hat{\beta}_1 &= |\mathbf{A}|^{-1}\{a_{22}b_1 - a_{12}b_2\} ,\\ \hat{\beta}_2 &= |\mathbf{A}|^{-1}\{a_{11}b_2 - a_{12}b_1\}\end{aligned}\right\} \qquad (6.24)$$

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

$$\sum_{l=0}^{\infty} \alpha_{1l} = \sum_{l=0}^{\infty} \alpha_{2l} = 1 , \qquad (6.25)$$

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

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

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 $\mathrm{Var}(\hat{\beta}_2)/\mathrm{Var}(\hat{\beta})$ where it is understood that the denominator corresponds to the model in (6.13). Now

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

where

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

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

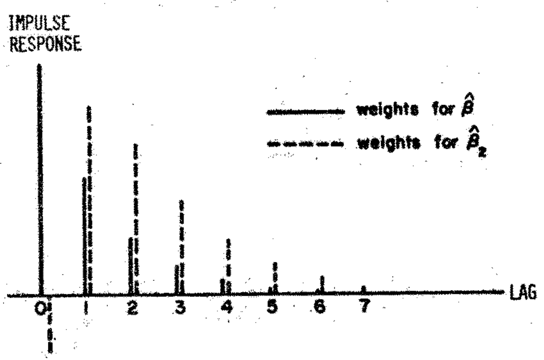
$$\eta(B) = (1-\delta B)^{-1}, \quad \varphi(B) = 1 - B \text{ and } \theta(B) = 1 - \theta B.$$

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

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

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 $t \le 0$. 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).

O.D. Anderson & M.R. Perryman, eds., TIME SERIES ANALYSIS

NONSTATIONARY FIRST-ORDER MOVING AVERAGE PROCESSES: THE MODEL-BUILDING PROBLEM

Marc Hallin
Université Libre de Bruxelles

Necessary and sufficient conditions are given under which a sequence $\rho_t = \text{corr}(z_t, z_{t-1})$, $t \in \mathbb{Z}$ is the sequence of autocorrelations of some MA(1) model with time-dependent coefficient $z_t = \varepsilon_t + a_t \varepsilon_{t-1}$. The corresponding model-building problem, i.e. the problem of finding all the models with a given (time-dependent) autocorrelation sequence is solved. It appears that the solution to this problem generally consists in a one-parameter family of models. The invertibility and Granger-Andersen invertibility of these models is studied, and it is shown that at most one of them is invertible, while one at most is Granger-Andersen *non* invertible. The so-called *generalized unit-circle* class of processes (i.e. processes z_t such that the innovation at time t cannot be expressed as a linear combination of the z_t's), extending the couple of stationary processes defined by $z_t = \varepsilon_t \pm \varepsilon_{t-1}$, is also characterized.

A. THE PROBLEM

1. Time-dependent and time-independent first-order moving average processes.

Let $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ denote a real univariate white noise, with mean zero and variance σ^2. A first-order moving average (MA(1)) model is characterized by a sequence of real numbers $a_t \neq 0$, $t \in \mathbb{Z}$, and associates to any white noise ε_t the process $\{z_t,\ t \in \mathbb{Z}\}$ defined by

$$z_t = \varepsilon_t + a_t \varepsilon_{t-1} \qquad t \in \mathbb{Z} \quad . \tag{1}$$

This paper being devoted to first-order moving average time series models only, by the word "model", we always mean a model of that particular class. The a_t's will be called the *coefficients of the model*. Indeed, the most general form of a MA(1) model is

$$y_t = a_{to}\, \varepsilon_t + a_{t1}\, \varepsilon_{t-1} \qquad t \in \mathbb{Z},$$

with $a_{to} > 0$ and $a_{t1} \neq 0$; however, setting $z_t = y_t/a_{to}$, z_t has the same autocorrelations as y_t, and is generated by an equation of type (1).

A model is said to be *time-dependent* or *-independent* according as its coefficients are time-dependent or constant. By a *one-parameter family of models*, we shall mean a set of models with coefficients $a_t = a_t(\alpha)$, $t \in \mathbb{Z}$, depending on one parameter Θ belonging to some real interval. The process $\{z_t;\ t \in \mathbb{Z}\}$ generated by a model for some given input white noise ε_t is called a first-order moving average (MA(1)) process; of course, time-independent models generate (second-order) stationary processes only. But, as we shall see, processes with time-independent autocorrelations can also be modelled into time-dependent models.

2. Autocorrelation sequences and the model-building problem.

Denoting by ρ_t the correlation coefficient corr (z_t, z_{t-1}), where z_t is generated by (1) for some white noise ε_t, we have

$$\rho_t = a_t \Big/ \sqrt{(1 + a_t^2)(1 + a_{t-1}^2)} \quad ; \tag{2}$$

higher order autocorrelations of course vanish.

The sequence ρ_t, $t \in \mathbb{Z}$ will be called the *autocorrelation sequence* associated with model (1); its values only depend on the coefficients *of the model*, whatever the input white noise ε_t. Time-independent models give rise to constant autocorrelation sequences ($\rho_t = \rho\ \forall\ t$).

Now, it is of some interest to ask, ρ_t $(t \in \mathbb{Z})$ being some given sequence of real numbers, which models, if any, ρ_t is the autocorrelation sequence of. Indeed, this is the essence of the model-building problem, and breaks up into two questions : *(i)* does (2), considered as a non-homogeneous non-linear difference equation (a_t unknown), admit real solutions ? *(ii)* if it does, what are these solutions ? Can (2) be solved for the coefficients a_t in terms of the autocorrelations ρ_t ?

If the problem is restricted to the case of time-independent models (coefficients $a_t = a\ \forall\ t$), and thus to constant autocorrelation sequences, (2) reduces to a second-degree equation, and the well-known answers to the above questions are

(i) a constant sequence $\rho_t = \rho$ is the autocorrelation sequence of some stationary model if and only if $\rho^2 \leq 1/4$;

(ii)-if $\rho^2 < 1/4$, $\rho_t = \rho$ is the autocorrelation sequence of two distinct stationary models, with coefficients

$$a_t^- = a^- = (1 - \sqrt{1 - 4\rho^2})/2\rho$$
and
$$a_t^+ = a^+ = (1 + \sqrt{1 - 4\rho^2})/2\rho = 1/a^- \ , \tag{3}$$

respectively, the model with coefficient a^- being the only *invertible* one.

-if $\rho^2 = \frac{1}{4}$ (*unit-circle case*), $\rho_t = \pm\ 1/2$ is the autocorrelation sequence of one, non-invertible, time independent model only, with coefficients

$$a_t = a = \begin{cases} 1 & \text{if } \rho_t = 1/2 \\ -1 & \text{if } \rho_t = -1/2 \end{cases} \qquad (4)$$

Our purpose is to give these two questions an answer in the general case of time-dependent models. As a first step, section B investigates the particular case of constant autocorrelation sequences, involving non-linear difference equations methods. In section C, we study the general case by means of continued fractions methods.

B. NONSTATIONARY PROCESSES WITH CONSTANT AUTOCORRELATION SEQUENCES

In this section, we consider the constant sequence $\rho_t = \rho$; the case $\rho = 0$ leading to the degenerate (and improper) model with coefficients $a_t = 0$, we assume $0 < \rho^2 \leq 1$.

Keeping in mind that, in equation (2), a_t and ρ_t have the same signs, (2) can be written

$$a_t^2 = \rho^2(1 + a_{t-1}^2)/(1 - \rho^2(1 + a_{t-1}^2)) \qquad t \in \mathbb{Z} \qquad (5)$$

or

$$a_t^2 = (a_{t+1}^2 \,/\, \rho^2(1 + a_{t+1}^2)) - 1 \qquad t \in \mathbb{Z} \;. \qquad (6)$$

Both these difference equations can be solved by means of classical methods (cf. [7], Chapter 5). The results can be recapitulated as follows.

Theorem 1.

$\rho_t = \rho \neq 0$, $t \in \mathbb{Z}$, is an autocorrelation sequence if and only if $\rho^2 \leq 1/4$.

Theorem 2.

(i) If $0 < \rho^2 < 1/4$ (assume $\rho > 0$; the case $\rho < 0$ follows immediatly),

(i1) $\rho_t = \rho$ is the autocorrelation sequence of a one-parameter family $\mathcal{F}$ of time-dependent models with coefficients a_t ranging over the interval $[a^-, a^+]$, where a^- and a^+ are the coefficients of the two time-independent models mentioned in (3).

(i2) Let $t_0 \in \mathbb{Z}$ and $a_{t0} \in [a^-; a^+]$: $\mathcal{F}$ contains one and only one model with coefficients $a_t(a_{t_0})$ such that $a_{t_0}(a_{t_0}) = a_{t_0}$.

(i3) $a_t(a_{t_0})$ is a continuous monotonically increasing function mapping $[a^-, a^+]$ on itself; of course, $a_t(a^\pm) = a^\pm$, $\forall\, t$.

(i4) Moreover, whatever the values of ρ, t_0 and $a_{t_0} \neq a^\pm$, $a_t(a_{t_0})$ is a strictly decreasing function of time, and

$$\lim_{t=-\infty} a_t(a_{t_0}) = a^+ , \quad \lim_{t=\infty} a_t(a_{t_0}) = a^- .$$

(i5) Hence, the only invertible model is the time-independent model with coefficients $a_t = a^-$, and the only non-invertible one in the sense of Granger and Andersen [3] is the time-independent model with coefficients $a_t = a^+$.

(ii) If $\rho^2 = 1/4$ (*unit-circle case*), the only model admitting $\rho_t = \rho$ as its autocorrelation sequence is the non-invertible time-independent model $z_t = \epsilon_t + \text{sgn}(\rho)\, \epsilon_{t-1}$, (denoting by sgn($\rho$) the sign of ρ).

Proof of theorems 1 and 2.

Suppose first $0 < \rho < 1/2$; consider equation (5) and the corresponding graphic description (fig.1).

The curve crosses the straight line $a_t^2 = a_{t-1}^2$ twice : in A^- (with slope less than unity) and in A^+ (with slope greater than unity). Since a_t must be real and positive, it follows from (5) and (6) that a_t^2 cannot take values outside of $(0, (1-\rho^2)/\rho^2)$, $\forall$ t.

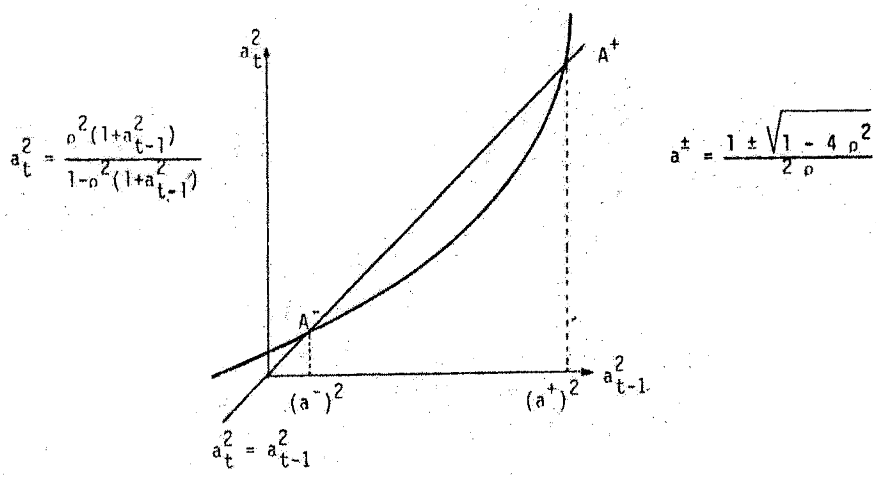


fig.1

Suppose $a_t^2 < (a^-)^2$: a_{t-1}^2, a_{t-2}^2 , ... can be obtained graphically as follows (fig.2).

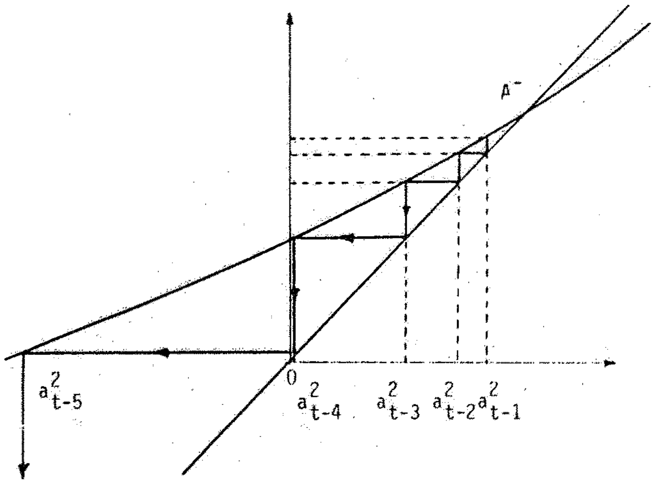


fig. 2

Clearly, a^2_t, a^2_{t-1}, a^2_{t-2}, ... is a strictly decreasing unbounded sequence, which is not consistent with the condition $a^2_t \geq 0 \; \forall \; t$. Similarly, if $a^2_{t-1} > (a^+)^2$, a^2_{t-1}, a^2_t, a^2_{t+1}, ... forms a strictly increasing unbounded sequence, contradicting $a^2_t < (1-\rho^2)/\rho^2 \; \forall \; t$. Now, if $a^2_{t-1} \in ((a^-)^2, (a^+)^2)$, it is easy to see (fig. 3) that a^2_t, a^2_{t-1}, a^2_{t-2}, ... forms a strictly decreasing sequence, converging to $(a^-)^2$, while a^2_t, a^2_{t+1}, a^2_{t+2}, ... forms a strictly increasing sequence, converging to $(a^+)^2$ (fig. 3). $(a^-)^2$ and $(a^+)^2$ are, obviously, fixed points.

When $\rho^2 = 1/4$, $(a^+)^2 = (a^-)^2 = 1$ is the only admissible solution to (5). Finally, if $\rho^2 > 1/4$, the curve lies everywhere above the straight line $a^2_t = a^2_{t-1}$, and, for any value of a^2_t, $\lim_{k=-\infty} a^2_{t+k} = -\infty$ and $\lim_{k=\infty} a^2_{t+k} = \infty$; equation (5) has thus no admissible solution.

Since a time-dependent model with coefficients $a_t > 0$ has an autocorrelation sequence $\rho_t = \rho$ iff $0 < a^2_t < (1-\rho^2)/\rho^2$ is a solution of (5), Theorem 1 and parts (*i*1), (*i*2), (*i*4) and (*ii*) of Theorem 2 follow immediatly from the preceding remarks on equation (5). Obviously, the function $a^2_{t_0+1}$ $(a^2_{t_0})$ is continuous and differentiable on $((a^-)^2, (a^+)^2)$, and its derivative is $1/(1-\rho^2(1+a^2_{t_0}))^2 > 0$, which completes the proof of (*i*3).

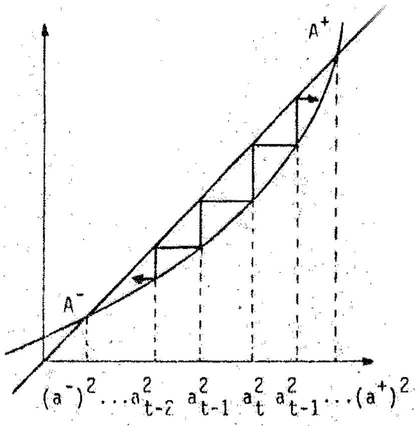


fig. 3

A first-order moving average model with coefficients a_t is invertible iff $\lim_{s=\infty} a_t \cdot a_{t-1} \cdot \ldots \cdot a_{t-s} = 0 \ \forall\ t$, and Granger-Andersen invertible iff $\lim_{s=\infty} a_t \cdot a_{t+1} \cdot \ldots \cdot a_{t+s} = 0 \ \forall\ t$. (*i*5) thus follows from

$$\lim_{s=\infty} a_t \cdot a_{t-1}(a_t) \cdot \ldots \cdot a_{t-s}(a_t) = \begin{cases} \lim_{s=\infty} (a^-)^s = 0 & \text{if } a_t = a^- \\ \lim_{s=\infty} (a^+)^s = \infty & \text{if } a_t \neq a^- \end{cases},$$

and

$$\lim_{s=\infty} a_t \cdot a_{t+1}(a_t) \cdot \ldots \cdot a_{t+s}(a_t) = \begin{cases} \lim_{s=\infty} (a^+)^s = \infty & \text{if } a_t = a^+ \\ \lim_{s=\infty} (a^-)^s = 0 & \text{if } a_t \neq a^+ \end{cases}. \quad \square$$

C. TIME-DEPENDENT AUTOCORRELATION SEQUENCES

1. Three characterizations of autocorrelation sequences.

As we have seen in the preceding section, a necessary and sufficient condition for $\rho_t = \rho$ to be a constant autocorrelation sequence is $\rho_t^2 \leq 1/4$. In the general case of time-dependent autocorrelations, we cannot expect such a simple formulation. The following examples illustrate two somewhat surprising (compare with [1] and [2]) properties of autocorrelation sequences.

Example 1.

ρ_t can reach values that are arbitrarily close to 1 or -1.

Consider the model ($\varepsilon > 0$ arbitrarily small)

$$z_t = \begin{cases} \varepsilon_t + \sqrt{\frac{\varepsilon}{1-\varepsilon}}\,\varepsilon_{t-1} & t < 0 \\ \varepsilon_t + \sqrt{\frac{1-\varepsilon}{\varepsilon}}\,\varepsilon_{t-1} & t \geq 0 \; ; \end{cases}$$

its autocorrelation sequence is

$$\rho_t = \begin{cases} 1 - \varepsilon & t = 0 \\ \sqrt{\varepsilon(1-\varepsilon)} & t \neq 0 \; . \end{cases}$$

(In order to obtain $\rho_0 = -1 + \varepsilon$, just modify $z_0 = \varepsilon_0 + \sqrt{.}\,\varepsilon_{-1}$ into $z_0 = \varepsilon_0 - \sqrt{.}\,\varepsilon_{-1}$).

Example 2.

ρ_t^2 *can be greater than* 1/4 *for an infinite number of values of* t.

Consider the model

$$z_t = \begin{cases} \varepsilon_t + 3\,\varepsilon_{t-1} & t = 0, \pm 2, \pm 4, \ldots \\ \varepsilon_t + \frac{1}{3}\,\varepsilon_{t-1} & t = \pm 1, \pm 3, \ldots \; ; \end{cases}$$

its autocorrelation sequence is

$$\rho_t = \begin{cases} 1/4 & t = 0, \pm 2, \pm 4, \ldots \\ 3/4 & t = \pm 1, \pm 3, \ldots \; . \end{cases}$$

We now give three necessary and sufficient conditions for a sequence ρ_t to be an autocorrelation sequence; these conditions are closely linked with the concept of *positive definite continued fractions*. Also, a fundamental role will be played by the so-called *chain sequences*; a *chain sequence* is a sequence of the form

$$(1-g_0)\,g_1,\; (1-g_1)\,g_2,\; (1-g_2)\,g_3,\; \ldots \quad ,$$

where the numbers $0 \leq g_p \leq 1$, $p = 0,1,2, \ldots$, are called *parameters* of the chain sequence. For other definitions about continued fractions, we refer the reader to [8].

Theorem 3.

The sequence ρ_t, $t \in \mathbb{Z}$, is an autocorrelation sequence if and only if one of the three (equivalent) conditions (*i*1), (*i*2), (*i*3) *and* condition (*ii*) hold for some $t \in \mathbb{Z}$:

(*i*1) the continued fractions

$$\cfrac{1}{1+z_1 - \cfrac{\rho_{t-1}^2}{1+z_2 - \cfrac{\rho_{t-2}^2}{1+z_3 - \ddots}}} \quad \text{and} \quad \cfrac{1}{1+z_1 - \cfrac{\rho_{t+1}^2}{1+z_2 - \cfrac{\rho_{t+2}^2}{1+z_3 - \ddots}}} \tag{7}$$

(z_p, $p = 1,2,3, \ldots$ are complex variables) are positive definite.

(*i2*) the quadratic forms

$$\sum_{r=1}^{p} \xi_r^2 - 2 \sum_{r=1}^{p-1} \rho_{t-r}\, \xi_r\, \xi_{r+1} \quad \text{and} \quad \sum_{r=1}^{p} \xi_r^2 - 2 \sum_{r=1}^{p-1} \rho_{t+r}\, \xi_r\, \xi_{r+1} \tag{8}$$

are non-negative for $p = 1,2,3, \ldots$ and for all real values of ξ_1, ξ_2, $\xi_3, \ldots$.

(*i3*) the sequences

$$\rho_{t-1}^2, \rho_{t-2}^2, \rho_{t-3}^2, \ldots \quad \text{and} \quad \rho_{t+1}^2, \rho_{t+2}^2, \rho_{t+3}^2, \ldots \tag{9}$$

are *chain sequences*.

(*ii*) the following inequality holds :

$$\cfrac{\rho_t^2}{1 - \cfrac{\rho_{t-1}^2}{1 - \cfrac{\rho_{t-2}^2}{\ddots}}} + \cfrac{\rho_{t+1}^2}{1 - \cfrac{\rho_{t+2}^2}{1 - \cfrac{\rho_{t+3}^2}{\ddots}}} \leq 1 \tag{10}$$

(if (*i1*) or (*i2*) or (*i3*) is satisfied, the two continued fractions in the left-hand member of (10) converge).

Proof.

(*i1*) and (*i2*) are equivalent, by definition. The equivalence between (*i2*) and (*i3*) follows from a theorem by Wall and Wetzel ([8], theorems 16.2 and 20.1). It is thus sufficient to prove that (*i3*) and (*ii*) characterize autocorrelation sequences.

A sequence ρ_t, $t \in \mathbb{Z}$, such that $0 < \rho_t^2 \leq 1$, is an autocorrelation sequence iff there exist real numbers a_t such that

$$\rho_t^2 = a_t^2 \,/\, (1 + a_t^2)(1 + a_{t-1}^2) \qquad t \in \mathbb{Z} \; ; \tag{11}$$

equivalent forms of this equation are

$$\rho_{t+1}^2\,(1 + a_t^2)(1 - \rho_t^2(1 + a_{t-1}^2)) = \rho_{t+1}^2 \tag{12}$$

and

$$\rho_t^2 = \frac{1}{1+a_{t-1}^2}\left(1 - \frac{1}{1+a_t^2}\right) . \tag{13}$$

Thus (11) holds iff there exists some $t_o \in \mathbb{Z}$ such that (12) holds for t_o, t_o+1, t_o+2, ... and (13) for t_o-1, t_o-2, ...

Putting

$$\alpha_t = 1 + a_t^2 \quad , \quad g_t = \rho_t^2 \, \alpha_{t-1} \quad , \quad f_t = \frac{1}{\alpha_{t-1}} \quad , \tag{14}$$

(12) and (13) take the forms

$$\rho_{t+1}^2 = g_{t+1}\,(1-g_t) \tag{12'}$$

and

$$\rho_t^2 = f_t\,(1 - f_{t+1}) . \tag{13'}$$

Hence, ρ_t is an autocorrelation sequence iff, simultaneously, ρ_{t+1}^2, ρ_{t+2}^2, ... is a chain sequence with parameters $g_p = \rho_{t+p}^2(1+a_{t+p-1}^2)$ $(p = 0,1, \ldots)$ *and* ρ_{t-1}^2, ρ_{t-2}^2, ... a chain sequence with parameters $f_p = 1/(1+a_{t-p-1}^2)$ $(p = 0,1, \ldots)$, for some $t \in \mathbb{Z}$. It remains to show, now, that this is possible provided condition (*ii*) is satisfied.

A chain sequence does not, in general, determine its parameters uniquely (as a consequence, an autocorrelation sequence will not, in general, determine an unique model) : a given chain sequence has minimal parameters m_p and maximal parameters M_p $(p = 0,1,2, \ldots)$ and if, for some $p_o \in \mathbb{Z}$, a value of the parameter $h_{p_o} = h$ is fixed, the values of the parameters are uniquely determined for every $p \in \mathbb{Z}$ ([8], theorem 19.2). Now, for the chains ρ_{t+1}^2, ρ_{t+2}^2, ... and ρ_{t-1}^2, ρ_{t-2}^2, ..., denote by m_p and M_p, n_p and N_p, respectively, these minimal and maximal parameters; m_o, n_o, M_o and N_o are given by

$$m_o = n_o = 0$$

$$M_o = 1 - \cfrac{\rho_{t+1}^2}{1 - \cfrac{\rho_{t+2}^2}{1 - \ddots}} \qquad\qquad N_o = 1 - \cfrac{\rho_{t-1}^2}{1 - \cfrac{\rho_{t-2}^2}{1 - \ddots}} \quad , \tag{15}$$

the latter continued fractions being convergent. Since $g_o = \rho_t^2\,\alpha_{t-1} = \rho_t^2/f_o$, we must have

$$\frac{\rho_t^2}{N_o} \leq g_o \leq M_o \quad ,$$

which is possible iff

$$\frac{\rho_t^2}{1-\frac{\rho_{t-1}^2}{1-\ddots}} \qquad 1-\frac{\rho_{t+1}^2}{1-\frac{\rho_{t+2}^2}{1-\ddots}} \tag{16}$$

i.e. iff condition (*ii*) is satisfied □.

Of course, if (*i*1), (*i*2), (*i*3) and (*ii*) hold for some $t \in \mathbb{Z}$, they hold for any $t \in \mathbb{Z}$. It is easy to verify, for example, that if inequality (16) is satisfied for t, it is also satisfied for t+1 :

$$\frac{\rho_{t+1}^2}{1-\frac{\rho_t^2}{1-\ddots}} \leqslant \frac{\rho_{t+1}^2}{1-\left(1-\frac{\rho_{t+1}^2}{1-\frac{\rho_{t+2}^2}{1-\ddots}}\right)} = 1-\frac{\rho_{t+2}^2}{1-\frac{\rho_{t+3}^2}{1-\ddots}}$$

In the case of a constant term sequence, Theorem 3 reduces to Theorem 1; indeed, a constant term sequence $\rho^2, \rho^2, \rho^2, \ldots$ is a chain sequence iff $\rho^2 \leq 1/4$ ([8], theorem 19.1); moreover,

$$\frac{\rho^2}{1-\frac{\rho^2}{1-\ddots}} = \frac{1}{2}\left(1-\sqrt{1-4\rho^2}\right) \leq \frac{1}{2},$$

and condition (*ii*) of Theorem 3 can thus be skipped.

We may also state the following theorem :

Theorem 4.

If ρ'_t is an autocorrelation sequence, and if ρ''_t is such that $\rho''^2_t \leq \rho'^2_t$, $t \in \mathbb{Z}$, then ρ''_t is also an autocorrelation sequence.

Corollary

Any sequence ρ_t, $t \in \mathbb{Z}$, such that $0 \neq \rho_t^2 \leq 1/4$, is an autocorrelation sequence.

Proof.

From (8), it is easy to see that, if $\rho'^2_{t-1}, \rho'^2_{t-2}, \ldots$ and $\rho'^2_{t+1}, \rho'^2_{t+2}, \ldots$ are chain sequences, then $\rho''^2_{t-1}, \rho''^2_{t-2}, \ldots$ and $\rho''^2_{t+1}, \rho''^2_{t+2}, \ldots$ are also chain sequences. Now, since $\rho''^2_t \leq \rho'^2_t$, we have, for some sequence of real numbers x_i

such that $x_i^2 \leq 1$ $(i = 0,1,\ldots)$,

$$\cfrac{\rho_t''^2}{1 - \cfrac{\rho_{t-1}''^2}{1 - \ddots}} = \cfrac{\rho_t'^2 x_0}{1 - \cfrac{\rho_{t-1}'^2 x_1}{1 - \ddots}} \leq \cfrac{\rho_t'^2}{1 - \cfrac{\rho_{t-1}'^2}{1 - \ddots}}$$

(the inequality follows from [8], Theorem 11.1). Hence, if the sequence ρ_t' satisfies condition (10), the sequence ρ_t'' also does. □

2. The general model-building problem.

Let ρ_t, $t \in \mathbb{Z}$ be an autocorrelation sequence.

Theorem 5.

(i) If (16) is a strict inequality, ρ_t is the autocorrelation sequence of a one-parameter family of time-dependent models with coefficients a_t such that $(a_t^-)^2 \leq a_t^2 \leq (a_t^+)^2$, where a_t^-, a_t^+ (and a_t) have the same sign as ρ_t,

$$(a_t^-)^2 = \cfrac{1}{1 - \cfrac{\rho_t^2}{1 - \cfrac{\rho_{t-1}^2}{1 - \ddots}}} - 1 \tag{17}$$

and

$$(a_t^+)^2 = \cfrac{1}{\cfrac{\rho_{t+1}^2}{1 - \cfrac{\rho_{t+2}^2}{1 - \ddots}}} - 1 \quad , \tag{18}$$

(17) and (18) are also equivalent to

$$(a_t^-)^2 = \frac{1}{1 + \text{g.l.b.}\left[\sum_{r=1}^{\infty} \xi_r^2 - 2 \sum_{r=1}^{\infty} \rho_{t-r}\, \xi_r\, \xi_{r+1} - 2\rho_t \xi_1\right]} - 1 \tag{17'}$$

and

$$(a_t^+)^2 = \frac{1}{\rho_{t+1}^2}(1+\text{g.l.b.}[\sum_{r=1}^{\infty} \xi_r^2 - 2\sum_{r=1}^{\infty} \rho_{t+r+2}\xi_r\xi_{r+1} - 2\rho_{t+2}\xi_1]) - 1, \qquad (18')$$

the greatest lower bounds (g.l.b.) being taken with respect to all real ξ_r for which the infinite series converge.

(ii) If (16) is an equality, then $a_t = a_t^- = a_t^+$, $t \in \mathbb{Z}$, are the coefficients of the *unique* model with autocorrelation sequence ρ_t.

(iii) Let $t_0 \in \mathbb{Z}$ and $a_{t_0}^2 \in [(a_{t_0}^-)^2, (a_{t_0}^+)^2]$ (with the correct sign) : there exists one and only one model, in the above mentioned family, with coefficients $a_t(a_{t_0})$ such that $a_{t_0}(a_{t_0}) = a_{t_0}$. The function $a_t^2(a_{t_0}^2)$ is a continuous, derivable, monotonically increasing function mapping $[(a_{t_0}^-)^2, (a_{t_0}^+)^2]$ into $[(a_t^-)^2, (a_t^+)^2]$, and $a_t(a_{t_0}^-) = a_t^-$, $a_t(a_{t_0}^+) = a_t^+$. The models with coefficients a_t^- and a_t^+ can thus be considered, respectively, as a *minimal* and a *maximal* model for the given autocorrelation sequence.

Proof.

Denote by g_t^- and g_t^+, respectively, the two (convergent) continued fractions appearing in the left and right members of (16). *(i)* is an immediate consequence of $g_t^- < g_t^+$ and $g_t^- \le g_t = \rho_t^2(1 + a_{t-1}^2) \le g_t^+$. As soon as a suitable value of a_{t_0} is fixed, the other values of the coefficients a_t can be determined, recursively, by means of (2), defining the function $a_t(a_{t_0})$; also, a simple computation shows that a_t^- and a_t^+, defined in (17) and (18), satisfy (2). (17') and (18') are obtained by expressing the maximal parameters of the chain sequences $\rho_{t+1}^2, \rho_{t+2}^2, \ldots$ and $\rho_{t-1}^2, \rho_{t-2}^2, \ldots$ in terms of the quadratic forms appearing in (8) (cf. [8] Theorem 20.2).

(ii) If (16) is an equality, $a_t^- = a_t^+$, since $g_t^- = g_t^+$ are the only admissible parameters for the chain sequences $\rho_{t+1}^2, \rho_{t+2}^2, \ldots, t \in \mathbb{Z}$.

Now *(iii)*, $a_t(a_{t_0})$ is obviously derivable, and

$$\frac{d}{d(a_{t_0}^2)} a_{t_0-1}^2 = \frac{1}{\rho_t^2(1+a_{t_0}^2)^2} = \frac{1 + a_{t_0-1}^2}{a_{t_0}^2(1+a_{t_0}^2)} > 0 . \quad \square \qquad (19)$$

3. Invertibility and Granger-Andersen invertibility.

In the classical model-building problem for moving average processes, a considerable role is played by the invertibility condition. The practical reason for this is that forecasting is impossible with non-invertible models; in addition, invertibility provides the model-building problem with unicity, since, among the various (more or less invertible; cf. [5]) models generating the same autocorrelation function (up to $(2q)!/(q!)^2$ distinct ones), one at most is an invertible model. As a matter of fact, the importance of the invertibility condition originates in the existence of a privileged model, with underlying causal implications : the Wold-Cramer decomposition. Hence, it is by no means surprising that these results still hold in the time-dependent case. As we shall see, there exists at most one invertible model for a given autocorrelation sequence (thus the Wold-Cramer decomposition of the corresponding process); the case when no invertible model exists is what we call the *generalized unit-circle case*.

Theorem 6.

Let ρ_t, $t \in \mathbb{Z}$ be an autocorrelation sequence, and denote by Π_t^- the infinite product

$$\Pi_t^- = \frac{(1 - g_t^-) \cdot (1 - g_{t-1}^-) \cdot (1 - g_{t-2}^-) \cdot \dots}{g_t^- \cdot g_{t-1}^- \cdot g_{t-2}^- \cdot \dots} , \tag{20}$$

where

$$g_t^- = \cfrac{\rho_t^2}{1 - \cfrac{\rho_{t-1}^2}{1 - \ddots}} . \tag{21}$$

(i) If $\Pi_t^- = \infty$ for some $t \in \mathbb{Z}$ (thus for any $t \in \mathbb{Z}$), the minimal model $z_t = \varepsilon_t + a_t^- \varepsilon_{t-1}$ where a_t^- is given by (17) or (17')) is the only invertible model with autocorrelation sequence ρ_t; $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ is thus the innovation process of $\{z_t;\ t \in \mathbb{Z}\}$. Moreover, for any other (non-minimal) model $z_t = \varepsilon_t + a_t \varepsilon_{t-1}$,

$$\lim_{t=-\infty} (a_t - a_t^+) = 0$$

(a_t^+ is given by (18) or (18')).

(ii) *(generalized unit-circle case)* If $\Pi_t^- \neq \infty$ for some $t \in \mathbb{Z}$ (thus for any $t \in \mathbb{Z}$), a process z_t with autocorrelation sequence ρ_t has no invertible representation.

Proof.

(i) A model with coefficients a_t is invertible iff

$$\lim_{s=\infty} a_t \cdot a_{t-1} \cdot \dots \cdot a_{t-s} \ (= \lim_{s=-\infty} G(t, t-s)) = 0 , \qquad t \in \mathbb{Z}$$

($G(t,s)$ being the one-sided Green's function associated with the model, considered as a first-order linear difference operator); with the notations of (14), $a_t^2 = g_{t-1}/(1 - g_{t-1})$, and the above condition can be written

$$\lim_{s=\infty} \frac{g_t \cdot g_{t-1} \cdot \ldots \cdot g_{t-s}}{(1-g_t) \cdot (1-g_{t-1}) \cdot \ldots \cdot (1-g_{t-s})} = 0 , \quad t \in \mathbb{Z} \tag{22}$$

Since $(a_t^-)^2 = g_{t-1}^-/(1 - g_{t-1}^-)$, the minimal model (with coefficients a_t^-) is invertible iff $H_t^- = \infty$.

It remains to show that the minimal model is the *only* invertible one. Consider a model with coefficients a_t such that $(a_t^-)^2 < a_t^2 \leq (a_t^+)^2$, and let $g_t = \rho_t^2 (1 + a_{t-1}^2)$; we have $g_t^- < g_t$, and

$$g_t^- = \cfrac{\rho_t^2}{1 - \cfrac{\rho_{t-1}^2}{1 - \ddots}} = g_t \cdot \cfrac{(1 - g_{t-1})}{1 - \cfrac{\rho_{t-1}^2}{1 - \ddots}} = g_t \cdot F ;$$

thus

$$F = \cfrac{1 - g_{t-1}}{1 - \cfrac{\rho_{t-1}^2}{1 - \ddots}} < 1 .$$

But $\rho_{t-1}^2, \rho_{t-2}^2, \ldots$ is a chain sequence with parameters (cf. (13'))
$f_t = 1/(1 + a_{t-1}^2) = (\rho_t^2 / g_t) = 1 - g_{t-1}$. Hence, F is a continued fraction of the type considered in theorem 11.1 in [8], from which we may conclude that $F = 1 - 1/S$, with

$$S = 1 + \sum_{s=0}^{\infty} \frac{f_t \cdot f_{t-1} \cdot \ldots \cdot f_{t-s}}{(1-f_t) \cdot (1-f_{t-1}) \cdot \ldots \cdot (1-f_{t-s})}$$

$$= 1 + \sum_{s=0}^{\infty} \frac{(1-g_{t-1}) \cdot (1-g_{t-2}) \cdot \ldots \cdot (1-g_{t-s-1})}{g_{t-1} \cdot g_{t-2} \cdot \ldots \cdot g_{t-s-1}} .$$

Therefore, $F < 1$ implies $S < \infty$,

$$\lim_{s=\infty} \frac{(1-g_t) \cdot \ldots \cdot (1-g_{t-s})}{g_t \cdot \ldots \cdot g_{t-s}} = 0 ,$$

and thus, owing to (22), that any model with coefficients $a_t \neq a_t^-$ is a non-invertible one.

We have seen (Theorem 5) that a_t^2 ($a_{t_0}^2$) is a derivable, monotonically increasing function from $[(a_{t_0}^-)^2, (a_{t_0}^+)^2]$ to $[(a_t^-)^2, (a_t^+)^2]$, with $a_t^2((a_{t_0}^-)^2) = (a_t^-)^2$ and $a_t^2((a_{t_0}^+)^2) = (a_t^+)^2$, respectively, as minimal and maximal values. By successive derivations,

$$\frac{d(a_{t-s-1}^2)}{d(a_t^2)} = \frac{\rho_{t+1}^2}{\rho_{t-s}^2} \frac{d}{dg_{t+1}} g_{t-s} = \frac{\rho_{t+1}^2}{\rho_{t-s}^2} \frac{(1-g_t) \cdot \ldots \cdot (1-g_{t-s})}{g_t \cdot \ldots \cdot g_{t-s}} . \tag{23}$$

The value in $a_t = a_t^-$ of this derivative becomes arbitrarily large as $s \to \infty$; hence, $\lim_{s=\infty} (a_{t-s} - a_{t-s}^+) = 0$. *(ii)* If $\Pi_t^- \neq \infty$, condition (22) is not satisfied, even for the minimal model; thus, there exists no invertible model with autocorrelation sequence ρ_t. □

An analog but highly different theorem can be established for the Granger-Andersen invertibility, strongly contrasting the two concepts. Because it relies on the same proof, we also mention here a property *(iii)* about $\lim_{t=\infty} a_t$.

Theorem 7.

Let ρ_t, $t \in \mathbb{Z}$ be an autocorrelation sequence, and denote by Π_t^+ the infinite product

$$\Pi_t^+ = \frac{(1-g_t^+) \cdot (1-g_{t+1}^+) \cdot (1-g_{t+2}^+) \cdot \ldots}{g_t^+ \cdot g_{t+1}^+ \cdot g_{t+2}^+ \cdot \ldots} , \tag{24}$$

where

$$g_t^+ = 1 - \cfrac{\rho_{t+1}^2}{1 - \cfrac{\rho_{t+2}^2}{1 - \ddots}} \tag{25}$$

(i) If $\Pi_t^+ = \infty$ for some $t \in \mathbb{Z}$ (thus for any $t \in \mathbb{Z}$), any model with autocorrelation sequence ρ_t is Granger-Andersen invertible.

(ii) $\Pi_t^+ \neq \infty$ for some $t \in \mathbb{Z}$ (thus for any $t \in \mathbb{Z}$), the only model with autocorrelation sequence ρ_t that is not Granger-Andersen invertible is the maximal model $z_t = \varepsilon_t + a_t^+ \varepsilon_{t-1}$ (a_t^+ is given by (18) or (18')).

(iii) Whatever the values of Π_t^+, and provided that $\lim \inf_{t=-\infty} \rho_t^2 > 0$,

$$\lim_{t=\infty} (a_t - a_t^-) = 0$$

for any (non maximal) model with parameters $a_t \neq a_t^+$.

Proof.

A model with coefficients a_t is Granger-Andersen invertible iff

$$\lim_{s=\infty} a_t \cdot a_{t+1} \cdot \ldots \cdot a_{t+s} \ (= \lim_{s=\infty} G(t+s,t)) = 0 \ , \qquad t \in \mathbb{Z},$$

thus iff

$$\lim_{s=\infty} \frac{g_t \cdot g_{t+1} \cdot \ldots \cdot g_{t+s}}{(1-g_t) \cdot (1-g_{t+1}) \cdot \ldots \cdot (1-g_{t+s})} = 0 \ , \qquad t \in \mathbb{Z} \ . \tag{26}$$

(i) Consider the parameters g_t^+, and denote by g_t the set of parameters associated with coefficients $a_t \neq a_t^+$:

$$1 - g_t^+ = \cfrac{\rho_{t+1}^2}{1 - \cfrac{\rho_{t+2}^2}{1 - \ddots}} = (1-g_t) \cfrac{g_{t+1}}{1 - \cfrac{\rho_{t+2}^2}{1 - \ddots}} = (1-g_t) \cdot H(g_t) \ ,$$

where $g_t \in [g_t^-, g_t^+)$; therefore, $H(g_t) < 1$. But $H(g_t)$ is a continued fraction of the type considered in theorem 11.1 in [8]; thus $H(g_t) = 1 - 1/S(g_t)$, with

$$S(g_t) = 1 + \sum_{s=1}^{\infty} \frac{g_{t+1} \cdot g_{t+2} \cdot \ldots \cdot g_{t+s}}{(1-g_{t+1}) \cdot (1-g_{t+2}) \cdot \ldots \cdot (1-g_{t+s})} \ ;$$

$H(g_t) < 1$ implies that $S(g_t)$ converge and that condition (26) be satisfied.

(ii) If $\Pi_t^+ = \infty$, condition (26) is also satisfied for the maximal model.

(iii) If $\liminf_{s=-\infty} \rho_{t-s}^2 > 0$, the convergence of $S(g_t)$ implies that

$$\lim_{s=\infty} \frac{d(a_{t+s}^2)}{d(a_{t-1}^2)} = \lim_{s=\infty} \frac{\rho_t^2}{\rho_{t+s+1}^2} \ \frac{g_t \cdot g_{t+1} \cdot \ldots \cdot g_{t+s}}{(1-g_t) \cdot (1-g_{t+1}) \cdot \ldots \cdot (1-g_{t+s})} = 0$$

for $a_t \neq a_t^+$. □

These two theorems contain parts (*i*4) and (*i*5) of Theorem 2 as particular cases : if $\rho^2 < \frac{1}{4}$, $g_t^-/(1-g_t^-) = a^-$ is always smaller than unity, and $g_t^+/(1-g_t^+) = a^+$ is always greater than unity; thus $\Pi_t^+ = \infty = \Pi_t^-$.

As an illustration of Theorems 6 and 7, let us consider Examples 1 and 2 above.

Example 1.

Clearly, if ε is small,

$$\lim_{s=\infty} a_{-1} \cdot a_{-2} \cdot \ldots \cdot a_{-s} = \lim_{s=\infty} (\sqrt{\varepsilon/(1-\varepsilon)})^{s} = 0 ,$$

and

$$\lim_{s=\infty} a_1 \cdot a_2 \cdot \ldots \cdot a_s = \lim (\sqrt{(1-\varepsilon)/\varepsilon})^s = \infty ;$$

hence the given model is invertible but not Granger-Andersen invertible. According to Theorem 7, we must have

$$a_t^- = a_t = a_t^+ \qquad t \in \mathbb{Z} .$$

But

$$g_0^+ = 1 - \cfrac{\varepsilon(1-\varepsilon)}{1 - \cfrac{\varepsilon(1-\varepsilon)}{1 - \ddots}} = \frac{1 + \sqrt{1-4\varepsilon(1-\varepsilon)}}{2} = 1 - \varepsilon$$

and

$$g_0^- = \cfrac{(1-\varepsilon)^2}{1 - \cfrac{\varepsilon(1-\varepsilon)}{1 - \ddots}} = \frac{2(1-\varepsilon)^2}{1 + \sqrt{1-4\varepsilon(1-\varepsilon)}} = 1-\varepsilon .$$

Thus the given autocorrelation sequence determines an unique model; it is easy to verify that

$$\Pi_1^+ = \frac{(1-g_1) \cdot (1-g_2) \cdot \ldots}{g_1 \cdot g_2 \cdot \ldots} = \frac{\varepsilon \cdot \varepsilon \cdot \ldots}{(1-\varepsilon) \cdot (1-\varepsilon) \cdot \ldots} = 0 ,$$

which confirms the non-invertibility in the sense of Granger and Andersen. This example also confirms that the Granger-Andersen invertibility is not a generalized invertibility (cf. [4] and [6]).

Example 2.

In this example, neither $\lim_{s=\infty} a_{-1} \cdot a_{-2} \cdot \ldots \cdot a_{-s}$, nor $\lim_{s=\infty} a_1 \cdot a_2 \cdot \ldots \cdot a_s$ exist.

Thus, the given model is neither classically, nor Granger-Andersen invertible. As a consequence, this autocorrelation sequence is a generalized unit-circle sequence and determines an unique model.

REFERENCES

[1] Anderson, O.D., On the individual Moving Average Inequality, Metrika 25 (1978) 241-245.

[2] Davies, N., Pate, M.B., and Frost, M.G., Maximum autocorrelations for moving average processes, Biometrika 61 (1974) 199-200.

[3] Granger, C . W.J. and Andersen, A., On the invertibility of time series models, Stoch. Proc. and Appl. 8 (1978) 87-92.

[4] Hallin, M., Invertibility and generalized invertibility of time series models, J.R.S.S. (B) (1980) 210-212.

[5] Hallin, M., Modèles non inversibles de séries chronologiques, Cahiers du C.E.R.O. 22 (1980) 363-368.

[6] Hallin, M., Addendum to "Invertibility and generalized invertibility of time series models", J.R.S.S. (B) 43 (1981) (to appear, n°1).

[7] Levy, H. and Lessman, F., Finite Difference Equations (Pitman, London, 1961).

[8] Wall, H.S., Analytic theory of continued fractions (Chelsea Publishing Company, Bronx, N.Y., 1948).

O.D. Anderson & M.R. Perryman, eds., TIME SERIES ANALYSIS

THE EFFECT OF OUTLIERS ON ARIMA MODELS FOR THE X-11-ARIMA

Guy Huot

Statistics Canada
Ottawa, Canada

This article is devoted to an improved estimation of the seasonal factors in the X-11-ARIMA method given the presence of outliers in the original unadjusted series. A Monte Carlo study is performed to assess the gain obtained by utilizing the X-11-ARIMA prior adjustment option where the replacement of the outliers is made to the original unadjusted series. Series are simulated according to ARIMA models and contaminated with outliers. The series are then modelled by an ARIMA process and the outliers are identified relative to the fitted values of the model. They are then replaced by the corresponding function values.

INTRODUCTION

It is well known that the X-11-ARIMA method (Dagum, 1980) offers decision-makers much more reliable current seasonally adjusted observations, as is instanced in Kuiper (1976) and Pierce (1978). To do so, the X-11-ARIMA proceeds as follows: (1) univariate ARIMA models are fitted to the monthly or quarterly series to be seasonally adjusted; (2) these series are extrapolated (forecasted or forecasted and backcasted) on a twelve-month or four-quarter basis; and (3), provided the generated extrapolations are acceptable, the X-11 method is then applied to the series thus extended.

Acceptable extrapolations have to satisfy three criteria (Dagum, 1979). First, the fitting of the ARIMA models to the series is controlled the Box and Pierce "portemanteau" χ^2 test of fit (Box and Pierce, 1970 and Prothero and Wallis, 1976) based on the residual autocorrelation function. The null hypothesis is tested at a 10% level of significance. In the second place, the mean absolute error between the observed values and the corresponding extrapolations must be smaller than 12% for forecasting and less than 18% for backcasting. Finally, the three ARIMA models of the automated version are checked for overdifferencing, which would result in the cancellation of parameters.

Outliers or extreme values frequently lead to the rejection of good ARIMA models because they will often increase the mean absolute extrapolation error, distort the χ^2 values or induce overdifferencing. It is frequently the case that the replacement of outliers decreases the mean absolute error and corrects the low χ^2 values.

This article is devoted to an improved estimation of the seasonal factors by the X-11-ARIMA method given the presence of outliers in the original unadjusted series. To this end, the interplay between the rejection and replacement of outliers, the extension of the series with acceptable extrapolations and the sensitivity of the seasonal factors are analyzed.

MONTE CARLO PROCEDURE

A Monte Carlo study was performed to assess the gain obtained by utilizing the X-11-ARIMA *prior-replacement-of-outliers option*, where the replacement of outliers is made to the original unadjusted series. A total of 100 series were simulated according to ARIMA models. The simulation procedure was (broadly) similar to the Wasim 1 simulation procedure for ARIMA models designed by McLeod and Hipel (1978). In this way, the ARIMA models represent a stochastic description of series with known properties. The simple $(0,1,1)(0,1,1)_{12}$ and the more complex $(2,1,2)(0,1,1)_{12}$ ARIMA models of the automated version of X-11-ARIMA have each been utilized for simulating 50 ten-year seasonal synthetic series. The white noise terms or residuals were normally independently distributed with mean 0 and variance 30. Only the white noise input varied from one series to the other. The relative contribution of the stochastic stable seasonal component and the irregular component to the stationary portion of the variance in the original series were approximately 70% and 14% respectively.

Each simulated series was subsequently contaminated with the same four pairs of outliers (opposite in sign and about equal in size) placed in the same randomly chosen positions. For the $(2,1,2)(0,1,1)_{12}$ model the identification of the outliers is related to their position in time and a *second test* was run with the same outliers placed in new positions.

There are different ways of handling outliers, as discussed in Barnett (1978). The rejection and replacement are one method among many. It is a valid approach when the true ARIMA model for the series is known or when a reasonable fitted model is available. The outliers are then identified relative to the fitted ARIMA model. For, the vector of univariate residuals between the data and the fit can be tested.

The automatic X-11-ARIMA prior-replacement-of-outliers option is as follows (Dagum, 1980, p. 94): "To extend the series with one extra year of forecasts, three ARIMA models are automatically fitted to the original series (or if applicable, to the original series modified by the prior monthly adjustment factors or the prior trading-day adjustment factors) and extreme values are replaced by the corresponding function values of the ARIMA model chosen in a first iteration. Thus, a prior treatment of extreme values consists of testing the residuals from the fitted ARIMA model of the first iteration that fulfills the criteria of acceptance against $\pm 2.5\sigma$. The values that fall outside this interval are replaced by the corresponding function values. The same ARIMA model is then fitted to the modified data to produce the forecasts. No model is automatically selected and no forecasts are made without the fulfillment of the criteria of acceptance". The experiment was, however, run utilizing the user-supplied ARIMA model option in order to apply the true (generating) ARIMA models to the contaminated series. The only difference with the automatic option is that the replacement of the outliers and the forecasts are then made without regard to the acceptance criteria.

The main purpose of the exercise is to minimize the influence of the outliers on the estimated seasonal factors. Note however that the outliers may contain valid information such as a turning-point. Therefore, they are rejected for the estimation of the seasonal factors but reintroduced in the final seasonally adjusted series and in the estimation of the final trend-cycle component.

Finally, the series were seasonally adjusted with the additive version of the X-11-ARIMA method and the ARIMA model was used to generate only forecasts. The prior-replacement-of-outliers option prints the percentage error in forecasts (calculated from the last three years of within-sample extrapolations) and the χ^2 values for both the first fit made to the original data with outliers and the second fit made to the data modified for the outliers. Part of the analysis was carried out comparing the percentage error in forecasts and the χ^2 values, calculated with and without the outliers. As further criteria in assessing the gain associated with the prior-replacement-of-outliers option, the estimated parameter values were checked for both overdifferencing and convergence to the true values.

The seasonal adjustment of the current observation is made by using the current seasonal factor (which might be a forecast seasonal factor or, occasionally, last year's factor). Thus, the objective of the experiment was also to see if the replacement of the outliers would reduce significantly the revisions to the seasonal factor forecasts (and the current seasonal factors). Revisions occur when additional years of observations are added to the raw series and seasonally adjusted by the X-11-ARIMA method. In order to measure such an influence of the outliers, the monthly total errors[1] in the seasonal factor forecasts and the current seasonal factors were estimated from both the ARIMA-extended and the non-extended series. The monthly total errors were summarized in the mean absolute error statistic which measures the dispersion of the forecasts of current factors relative to the final seasonal factors (Dagum, 1975). This analysis was based on (1) computing the ratio of the mean absolute error statistics for the non-extended and the extended series for both the forecasts and the current seasonal factors, and (2) comparing the ratios of the series contaminated with outliers with the ratios of the series where the outliers have been replaced. The Wilcoxon Signed Rank test was applied to test the hypothesis of no difference between the monthly absolute errors of the non-extended to the extended series (with and without replacement), e.g. the 10% singificance level being 23.

3. EVALUATION OF THE PRIOR-REPLACEMENT-OF-OUTLIERS OPTION

The results of applying the prior-replacement-of-outliers option to the simulated series are analyzed in this section.

3.1. Series generated and fitted with the $(0,1,1)(0,1,1)_{12}$ model— The χ^2 test of fit and the extrapolation acceptance criteria are presented in Table 1 for both the series contaminated with outliers and the *modified* series, where the outliers have been replaced. All of the former series failed either one or both tests. On the other hand, 95% of the modified series passed the tests because the use of the prior-replacement-of-outliers option has decreased the average percentage error in forecasts to 3.89% from 13.01% and brought up the average χ^2 probability value to 57.40% from 0.43%.

Table 1 Results obtained from the $(0,1,1)(0,1,1)_{12}$ ARIMA model for the extrapolation and fitting statistics

	Series with outliers			Modified series		
	average error in forecasts	average χ^2 values	percentage of rejection	average error in forecasts	average χ^2 value	percentage of rejection
all series†	13.01%	0.43%	100%	3.89%	57.40%	5%
accepted series‡	n.a.	n.a.	n.a.	3.91%	60.07%	0%

† All series without regard to the acceptance criteria ‡Series that passed the acceptance criteria

Table 2 shows in what way the outliers have influenced the parameter values of the fitted model. The average parameter values estimated from all the contaminated series are $\hat{\theta} = 0.91$ and $\hat{\Theta} = 0.89$ with a rate of rejection of 70% for overdifferencing while the values estimated from the modified series are $\hat{\theta} = 0.67$ and $\hat{\Theta} = 0.79$ with a rate of rejection of only 10%. For the $(0,1,1)(0,1,1)_{12}$ model it is assumed that there are evidences of overdifferencing if $\hat{\theta}$ or $\hat{\Theta}$ is greater than 0.90 (Dagum, 1979). Furthermore, the parameter values of the modified series relative to the ones of the contaminated series converge to the true values which are 0.40 and 0.60 respectively.

Table 2 Percentages of overdifferencing and average estimated parameter values of the $(0,1,1)(0,1,1)_{12}$ ARIMA model with true parameter values $\theta = 0.40$ and $\Theta = 0.60$

	series with outliers			modified series		
	$\hat{\theta}$	$\hat{\Theta}$	rejection for overdifferencing	$\hat{\theta}$	$\hat{\Theta}$	rejection for overdifferencing
all series	0.91	0.89	70%	0.67	0.79	10%
accepted series	0.84	0.84	0%	0.65	0.78	0%

The spectral analysis of the asymmetric filters in the X-11-ARIMA method are influenced by the ARIMA model and, given the model, by its parameter values (Dagum, 1980). In this way, the ARIMA forecast function affects significantly the spectrum of the seasonally adjusted series which reflects the implicit hypotheses of the ARIMA model and its parameter values. Thus, the convergence property is important since the adjusted series will then reflect more closely the hypotheses of the generating or true ARIMA model.

The next two tables show the effect of the prior-replacement-of-outliers option on the stability of the seasonal factors. Considering Table 3, there is a loss in stability of the seasonal adjustment of the contaminated extended series relative to the contaminated non-extended series for both the current and the forecast seasonal factors. That is, the ARIMA option (series extension) has increased the revisions to the seasonal factors. In contrast, Table 4 shows gains of 42.00% and 17.00% respectively for the current and forecast seasonal factors. That improvement is associated with the seasonal adjustment of the modified extended series relative to the contaminated non-extended series. That is to say, the ARIMA option used in conjunction with the prior-replacement-of-outliers option has significantly reduced the revisions to the seasonal factors. As noted, the gain is greater for the current seasonal factors than the forecast seasonal factors. In fact, the presence of the outliers in the extended series has much less impact on the forecast factors because the weight system assigns less weight to the forecast values as they depart from the last observation.

Note that on average 87% of the outliers have been identified and properly replaced.

Table 3 Comparison of X-11-ARIMA versus X-11 for the revisions of the seasonal factors for series with outliers—model $(0,1,1)(0,1,1)_{12}$

	current seasonal factors	average Wilcoxon value	forecast seasonal factors	average Wilcoxon value
all series	-27.24%†	55	-18.22%†	47
accepted series	n.a.	n.a.	n.a.	n.a.

† Negative values indicate that the X-11 seasonal factors are more stable

Table 4 Comparison of X-11-ARIMA versus X-11 for the revisions of the seasonal factors for series modified by the prior-replacement-of-outliers option—model $(0,1,1)(0,1,1)_{12}$

	current seasonal factors	average Wilcoxon value	forecast seasonal factors	average Wilcoxon value
all series	42.00%†	17	17.00%†	34
accepted series (85%)	46.00%†	15	20.00%†	32

† Positive values indicate that the X-11-ARIMA seasonal factors are more stable

3.2. Series generated and fitted with the $(2,1,2)(0,1,1)_{12}$ model— Section 3.1 has shown the net positive impact of the replacement of the outliers on the extrapolation, the χ^2 and the overdifferencing statistics on the one hand; and, on the revisions to the seasonal factors on the other hand. The two effects have moved in tandem. In contrast, this is seen not always to hold true for the more complex $(2,1,2)(0,1,1)_{12}$ ARIMA model. As shown for the first test in Tables 5, 6 and 7, there is a significant reduction in the forecast error but no gain in the revisions to the seasonal factors. The reason for this is the fact that on average only 4 outliers have been identified and replaced. The other 4 have been incorporated into the systematic part of the series. That is, the second-order autoregressive process with complex zeroes has exhibited a cyclical behaviour and the position of the outliers in time became very important.

In the second test, the prior-replacement-of-outliers option was applied to series with the same outliers as above placed in new randomly chosen positions. A very good performance was obtained. The average percentage error in forecasts was deflated to 5.34% from 15.62% (Table 5, second test) and table 7 (second test) shows gains of 36.36% and 28.02% respectively for the current and the forecast seasonal factors. There was, however, no evidence of convergence to the true parameters in both cases and approximately 20% of the contaminated and the modified series were rejected for overdifferencing.

Note that, in both sections, there was no evidence of residual seasonality in the final seasonally adjusted series at the 1% level.

Table 5 Results obtained from the $(2,1,2)(0,1,1)_{12}$ ARIMA model for the extrapolation and fitting statistics

all series	series with outliers: average error in forecasts	average χ^2 value	percentage of rejection	modified series: average error in forecasts	average χ^2 value	percentage of rejection
first test	13.40%	26.69%	90%	6.97%	43.87%	13%
second test†	15.62%	23.36%	100%	5.34%	91.03%	0%

† Series with outliers placed in new randomly chosen positions

Table 6 Comparision of the X-11-ARIMA versus X-11 for the revisions of the seasonal factors for series with outliers—$(2,1,2)(0,1,1)_{12}$ model

all series	current seasonal factors	average Wilcoxon value	forecast seasonal factors	average Wilcoxon value
first test	-46.52%	60	-30.33%	50
second test	-19.58%	56	-9.68%	43

Table 7 Comparison of the X-11-ARIMA versus X-11 for the revisions of the seasonal factors for series modified by the prior-replacement-of-outliers option—model $(2,1,2)(0,1,1)_{12}$

all series	current seasonal factors	average Wilcoxon value	forecast seasonal factors	average Wilcoxon value
First test	-11.76%	44	-4.30%	42
Second test	36.36%	14	28.02%	23

4. CONCLUSION

The X-11-ARIMA prior-replacement-of-outliers option applicable to the original unadjusted series has been analysed in this study. The series is modelled by an ARIMA process and the outliers are identified relative to the fitted values of the model. They are then replaced by the corresponding function values. The replacement of the outliers improves significantly both the performance of the same ARIMA model fitted to the modified series and the revisions to the final seasonal factors.

The $(0,1,1)(0,1,1)_{12}$ and the $(2,1,2)(0,1,1)_{12}$ ARIMA models of the automated version of X-11-ARIMA have each been tested using 50 ten-year ARIMA simulated series contaminated with outliers. For the $(0,1,1)(0,1,1)_{12}$ model, 87% of the outliers on average have been identified and properly replaced. Consequently, the percentage error in the ARIMA forecasts decreased to 3.89% from 13.01% and the average χ^2 probability value was brought up to 57.40% from 0.43%. Of most importance is the minimization of the revisions to the final seasonal factors. The gains were 42.00% and 17.00% respectively for the current and the forecast seasonal factors.

On the other hand, the more complex $(2,1,2)(0,1,1)_{12}$ model is rather sensitive to the positions of the outliers in time. That is, some outliers might be incorporated into the function values. However, when the outliers are properly identified, gains similar to those of the $(0,1,1)(0,1,1)_{12}$ model can be expected.

ACKNOWLEDGEMENT

I am very grateful to Mr. Kim Chiu for writing the computer programs needed for the simulation and the comparisons.

FOOTNOTE

[1] For example, the total forecast error is defined as the difference between the seasonal factor forecasts and the seasonal factors for the same month when at least three more years are added to the series.

REFERENCES

[1] Barnett, V., The study of outliers: purpose and model, Applied Statistics, 3 (1972) 242-250.

[2] Box, G.E.P. and Pierce, D.A., Distribution of residual autocorrelations in autoregressive integrated moving average time series models, J.A.S.A., 65 (1970) 1509-1526.

[3] Dagum, E.B., Seasonal factor forecasts from ARIMA models, Proceedings of the International Statistical Institute 40th Session at Warsaw, 3 (1975).

[4] Dagum, E.B., Further modifications in the selection of ARIMA models for the X-11-ARIMA seasonal ajustment method (Statistics Canada, Research Paper, Time Series and Seasonal Adjustment Staff, 1979).

[5] Dagum, E.B., The X-11-ARIMA seasonal adjustment method, (Statistics Canada, Ottawa, Catalogue No. 12-564E, 1980).

[6] Dagum, E.B., Properties of the linear filters of the X-II-ARIMA method for current seasonal adjustment, (Statistics Canada, Research Paper, Time Series and Seasonal Adjustment Staff, 1980).

[7] Kuiper, J., A survey and comparative analysis of various methods of seasonal adjustment, Proceedings of the NBER, Bureau of the Census Conference on Seasonal Analysis of Economic Time Series, Washington, D.C. (September 1976).

[8] McLeod, A.I. and Hipel, K.W., Simulation procedures for Box-Jenkins models, Water Resources Research, 5 (1978) 969-975.

[9] Pierce, D.A., Data revisions with moving average seasonal adjustment procedures, Federal Reserve Board (1978).

[10] Prothero, D.L. and Wallis, K.F., Modelling macroeconomic time series (with discussion), J.R.S.S., A 139 (1976) 468-500.

O.D. Anderson & M.R. Perryman, eds., TIME SERIES ANALYSIS

DIAGNOSTIC CHECKS FOR THE ARIMA MODELS OF THE X-11-ARIMA SEASONAL ADJUSTMENT METHOD

Estela Bee Dagum

Seasonal Adjustment and Time Series Staff
Statistics Canada
Ottawa, Ontario
CANADA

This article introduces a set of criteria of fitting and extrapolation to be fulfilled by any of the three built-in ARIMA models of the X-11-ARIMA for being used in the seasonal adjustment of the series in question. The adequacy of any of the three program-supplied models for fitting the data is verified by: (1) testing the randomness of the residuals and (2) checking up on the estimated parameter values for evidence of overdifferencing. The adequacy of any of the three models for data extrapolations into the past (backcast) and into the future (forecast) is measured by the mean absolute extrapolation error, in percentage of the level of the series, (MAPE) over the first and last three years respectively. The upper bounds of acceptance of both the fitting and extrapolation tests are based on empirical and theoretical considerations. The article also discusses the effect of the logarithmic transformation on series that follow an additive decomposition model. It is shown that the logarithmic transformation destroys the underlying properties of the series and that the ARIMA models adequate for the non-transformed data are no longer applicable.

INTRODUCTION

The X-11-ARIMA Seasonal Adjustment Method developed by Dagum (1975, 1980) basically consists of:

(1) Modelling the original series by autoregressive integrated moving averages processes (ARIMA models) of the Box and Jenkins (1970) types.

(2) Extrapolating one year (or two) of unadjusted data at each end of the series from the ARIMA models used in (1). This operation called "forecasting" and "backcasting" is designed to extend the observed series at both ends.

(3) Seasonally adjusting the extended series with various linear filters (moving averages). For central observations, these linear filters are symmetric and fixed. They are equivalent to those of the Method II - X-11 variant (Shiskin, Young and Musgrave, 1967) if the series have 7 or more years of data. For end observations, the linear filters are flexible and change with the type of ARIMA model fitted to the series and its parameter values.

The X-11-ARIMA method minimizes the revisions of the estimated seasonal factor forecasts in MSE as shown theoretically by Geweke (1978) and Pierce (1980) and offers a model for the observed series.

The use of the ARIMA option by default automatically tests three built-in ARIMA models against built-in criteria of fitting and extrapolation. The program then chooses the one that gives the smallest absolute average forecasting error for the last three years among those models that fulfilled the above criteria. The forecasts and backcasts obtained from the chosen model are used to extend the original series at both ends. For series highly contaminated by outliers, an automatic replacement of extreme values can be made before the selection of a final built-in ARIMA model for forecasting and backcasting. The series can be extended with forecasts only or with forecasts and backcasts simultaneously.

If none of the built-in models passes the built-in criteria of fitting and extrapolation, the user can identify and then submit his/her own model or resubmit the built-in models that marginally failed the criteria of acceptance.

The user-supplied ARIMA model options overrides all the built-in criteria of fitting and extrapolation and it is the user who decides whether the model is or is not acceptable.

The three built-in models using the Box and Jenkins (1970) symbolic notation $(p,d,q)\ (P,D,Q)_s$ are:

(1) $(0,1,1)(0,1,1)_s$ (2) $(0,2,2)(0,1,1)_s$ (3) $(2,1,2)(0,1,1)_s$

Where s, the periodicity of the seasonality is equal to 12 for monthly series and to 4 for quarterly series. The first two models are applied to log transformed data when the decomposition model assumes multiplicative or log additive relationships among the trend-cycle, the seasonals and the irregulars.

The criteria of fitting and extrapolation that must be fulfilled by any of the three program-supplied ARIMA models to be used for seasonal adjustment given by Dagum (1978) and Lothian and Morry (1978), are here modified. The changes introduced result from further theoretical studies and empirical evidences and are analyzed in the following sections. Furthermore, the effect of the logarithmic transformation for the identification of ARIMA models of series that follow an additive decomposition model is discussed in the last section.

2. THE ARIMA MODEL FITTING CRITERIA

The adequacy of any of the three automatically fitted ARIMA models by the X-11-ARIMA method is verified by: (1) testing the randomness of the residuals with a general or "portmanteau" test of fit developed by Box and Pierce (1970) and (2) checking up on the estimated values of the parameters for evidences of overdifferencing.

2.1. The "Portmanteau" Test of Fit

The general, or "portmanteau" test verifies the adequacy of a fitted ARIMA model by taking as a whole the autocorrelations of the residuals for various lag, say, k=1, 2, ..., m. For large samples, the portmanteau statistics

$$(2.1.1) \qquad Q=n\sum_{k=1}^{m} \hat{r}_k^2$$

is asymptotically distributed as χ^2 with (m-p-P-q-Q) degrees of freedom, where the $\hat{r}_k$ are the autocorrelations of the estimated residuals $\hat{a}_t$ distributed independently in $N(0,n^{-1})$, n=N-d-D is the number of observations (left after differencing) used to fit the model. Since the assumption that $\hat{r}$ has variance n^{-1} is a good approximation to $var(r_k)=(n-k)/n(n+2)$ only for values of k which are not small relatively to n, the (2.1.1) is likely to be inadequate for relatively small samples. Prothero and Wallis (1976) and Ljung and Box (1978) introduced the following modified statistic,

$$(2.1.2) \qquad Q' = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\rho_k^2$$

This statistic is incorporated into the X-11-ARIMA method to test the randomness of the residuals. Given the seriousness of committing a type II error when testing the null hypothesis of non-autocorrelation of the residuals, the critical value of the test is checked at a 10% probability level for the three automatically fitted models. This probability level also reduces the importance of the objections made by Davies, Triggs and Newbold (1977) to the modified equation (2.1.2). According to these authors, the variance of the(2.1.2) can differ much from χ^2 and thus, the true significance levels are likely to be much lower than predicted by asymptotic theory.

Although the option for the automatic selection of the ARIMA model rejects any model for which the χ^2 probability value of the portmanteau test is smaller than 10%, the user can by-pass this constraint with the option for user's supplied model. In fact, the X-11-ARIMA program calculates and prints the p values of the χ^2 distribution corresponding to the portmanteau test and the analyst can then decide whether to accept or reject the null hypothesis for any particular level of significance.

Furthermore, it is important to assure that the ARIMA model has been fitted to data without outliers, for these can easily distort the value of the portmanteau statistic and thus, falsely increase or decrease the p value of the test.

2.2. Overdifferencing

The other criterion recently incorporated into the X-11-ARIMA computer program to assess the goodness of fit of an ARIMA model is overdifferencing.(1) Excessive application of the difference operator, to generate a process stationary in the differences, induces a non-invertible moving-average process in the residuals. When overdifferencing is present, the first order autocorrelation of the residuals is very high but not often is detected by the portmanteau test for this is applied to the residual autocorrelations of various lags, taken as a whole. For that reason, even though the automatically fitted models pass the criterion of randomness of the residuals using the portmanteau test, they are further checked for evidences of overdifferencing.

The correct approach to detect overdifferencing is to test whether the roots associated with the moving average polynomial are significantly different to one. For a simple $(0,1,1)(0,1,1)_s$ model, this is equivalent to test whether the estimated $|\hat{\theta}|$ and $|\hat{\Theta}|$ are significantly different to one. Unfortunately, the distribution of estimators of non-invertible moving average processes is unknown and the use of the student-t or Normal distribution to test whether $|\theta|$ or $|\Theta|$ is equal to one can lead to very misleading results as shown by Plosser and Schwert (1977) with Monte Carlo experiments. Furthermore, these authors showed that nonlinear estimation procedures of moving average processes which constrains the MA parameter to $|\hat{\theta}| \leq 1$ are bound to introduce a downward bias that decreases as the sample size increases. For 1000 replications of samples of 50,100, and 200 observations, they found that the average values of $\hat{\theta}$ were .9205, .9353 and .9508 respectively, when in fact the true value was one.

In the context of the X-11-ARIMA seasonal adjustment program, a $(0,1,1)(0,1,1)_s$ model is rejected because of overdifferencing if $\hat{\theta}$ or $\hat{\Theta}$ is greater than .90. For higher order moving average processes, there are evidences of overdifferencing if the sum of the ordinary moving average parameters or of the seasonal moving average parameters is greater than .90. For the built-in $(0,2,2)(0,1,1)_s$ model

$$(2.2.1) \qquad (1-B)^2\,(1-B^s)\,z_t = (1-\hat{\theta}_1 B - \hat{\theta}_2 B^2)(1-\hat{\Theta}B^s)a_t$$

given the conditions of invertibility,

$$(2.2.2) \qquad \hat{\theta}_1+\hat{\theta}_2<1\ ; \quad \hat{\theta}_2-\hat{\theta}_1<1\ ; \quad |\hat{\theta}_2|<1, |\hat{\Theta}|<1$$

overdifferencing is present if: (1) $\hat{\Theta}>0.90$ or (2) $\hat{\theta}_1+\hat{\theta}_2>0.90$.

In the latter case, the ordinary second order moving average term can be factored approximately into a first difference times a first order moving average term under the two following forms:

$$(2.2.3) \qquad 1-\hat{\theta}_1 B-\hat{\theta}_2 B^2 \simeq 1-(1+\hat{\lambda})B+\lambda B^2=(1-B)(1-\hat{\lambda}B)$$

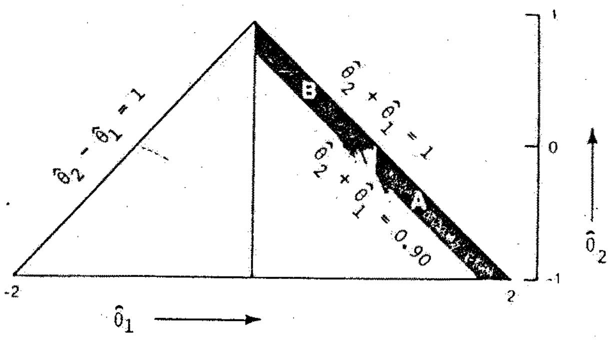
and

$$(2.2.4) \qquad 1-\hat{\theta}_1 B-\hat{\theta}_2 B^2 \simeq 1-(1-\hat{\lambda})B-\lambda B^2 = (1-B)(1+\hat{\lambda}B)$$

For the (2.2.3) $\hat{\theta}_1 \simeq 1+\hat{\lambda}>1$ and $\hat{\theta}_2 \simeq -\hat{\lambda}<0$ and, for the (2.2.4) $0<(\hat{\theta}_1\simeq 1-\hat{\lambda})<1$ and $\hat{\theta}_2\simeq\hat{\lambda}>0$ for the factorization to be feasible and, at the same time, satisfying the invertibility conditions. The Figure I shows the sets of feasible values of $\hat{\theta}_1$ and $\hat{\theta}_2$ given the invertibility conditions where, A corresponds to (2.2.3) and B to (2.2.4).

The (1-B) factor of (2.2.3) and (2.2.4) can be simplified with the second order difference of the autoregressive part of (2.2.1) and the $(0,2,2)(0,1,1)_s$ model is thus reduced to a $(0,1,1)(0,1,1)_s$ model. However, if in the reduced (0,1,1) $(0,1,1)_s$ model $\hat{\lambda}$ differs much from $\hat{\theta}_2$ (very roughly $|\hat{\theta}_2-\hat{\lambda}|$ is greater than .15) the cancellation of the parameters will not lead to the adequate model. Experimentation with a large number of series suggests that, in such cases, one difference be replaced by increasing the order of the moving average polynomial by one. This means that instead of being a $(0,1,1)(0,1,1)_s$ the adequate model for the overdifferenced (0,2,2)(0,1,1) model, it would be a $(0,1,2)(0,1,1)_s$, if the factorization is not exact.

For the $(2,1,2)(0,1,1)_s$ process, with the same conditions of invertibility of the $(0,2,2)(0,1,1)_s$ model plus the conditions of stationarity of the second order AR(2) process, namely, $\hat{\phi}_2+\hat{\phi}_1<1$, $\hat{\phi}_2-\hat{\phi}_1<1$, $-1<\hat{\phi}_2<1$, the presence of overdifferencing due to the second order moving average term is equivalent to the previous case. Concerning the AR(2) process, it can be identically factored as (2.2.3) and (2.2.4). When this is the case, the (1-B) factor of the AR(2) term indicates either: (1) the need of an extra difference, that is, there is underdifferencing, or (2) evidences of overfitting if the (1-B) factor can be cancelled with another (1-B) factor corresponding to the MA(2) member.



$A = \{(\hat{\theta}_1,\hat{\theta}_2)\,|\,\hat{\theta}_1>1,\ \hat{\theta}_2<0,\ 0.90<\hat{\theta}_2+\hat{\theta}_1<1\}$ $B = \{(\hat{\theta}_1,\hat{\theta}_2)\,|\,0<\hat{\theta}_1<1,\ \hat{\theta}_2>0,\ 0.90<\hat{\theta}_2+\hat{\theta}_1<1\}$

FIGURE 1.

ADMISSIBLE REGIONS FOR $\hat{\theta}_1$ AND $\hat{\theta}_2$ FOR PRESENCE OF OVERDIFFERENCING IN AN INVERTIBLE MA(2) PROCESS OF THE X-11-ARIMA METHOD.

3. THE ARIMA MODEL EXTRAPOLATION CRITERION

3.1. The extrapolation function and evaluation of the extrapolation values

The extrapolation procedure of ARIMA models described in Box and Jenkins (1970) expresses Z_t (or its log) directly in terms of earlier observed values of Z_t, earlier forecasts of Z_t, and current and previous values of the residuals a_t.

The general multiplicative seasonal model applied by the X-11-ARIMA seasonal adjustment method is,

(3.1.1) $\phi_p(B)\Phi_P(B^s)\Delta^d\Delta^D Z_t = \theta_q(B)\Theta_Q(B^s)a_t$

where p and P give the order of the autoregressive operators, q and Q give the order of the moving average operators and d and D give the order of the difference operators. The difference equation form of the corresponding extrapolation is,

(3.1.2) $Z_{t+\ell} = \alpha_1 Z_{t+\ell-1} + \dots + \alpha_m Z_{t+\ell-m} + a_{t+\ell} - \beta_1 a_{t+\ell-1} - \dots - \beta_n a_{t+\ell-n}$

$= \alpha(B)Z_t + \beta(B)a_t$

where $m = p + s.P+d+s.D$ and $n = q + s.Q$; $\alpha(B) = \phi_p(B)\Phi_P(B^s)\Delta^d\Delta^D$ and $\beta(B) = \theta_q(B)\Theta_Q(B)$ are the general autoregressive and the general moving average operators of (3.1.1) respectively.

The extrapolated value $\hat{Z}_t(\ell)$ of $Z_{t+\ell}$ made at origin t for time t+ℓ, is the conditional mathematical expectation of $Z_{t+\ell}$ given knowledge of all the Z's up to time t. If ℓ>0 the extrapolated values are forecasts and if ℓ<0, the extrapolated values are backcasts.

The accuracy of the extrapolations can be evaluated by expressing the (3.1.2) under the form of an infinite convergent sum of current and previous values of a_t,

$$(3.1.3) \quad Z_{t+\ell} = \sum_{j=-\infty}^{t+\ell} \Psi_{t+\ell-j} a_j = \sum_{j=0}^{\infty} \Psi_j \, a_{t+\ell-j}$$

where $\Psi_0 = 1$.

The extrapolated error $e_t(\ell)$ at time t+ℓ is

$$(3.1.4) \quad e_t(\ell) = Z_{t+\ell} - \hat{Z}_t(\ell) = \sum_{j=0}^{\ell-1} \Psi_j a_{t+\ell-j}$$

The expected value of the extrapolation error is zero and, thus, the extrapolation $\hat{Z}_t(\ell)$ is unbiased.

The variance of $e_t(\ell)$ is

$$(3.1.5) \quad \text{var}\,(e_t(\ell)) = (1 + \sum_{j=1}^{\ell-1} \Psi_j^2)\sigma_a^2$$

The forecast (backcast) $\hat{Z}_t(\ell)$ generated by an ARIMA model is optimal in the mean square error sense for it is well known in linear prediction theory that the conditional mathematical expectation minimizes the mean square error of a linear normal process [Wold, (1938), Kolgomorov (1939), Wiener (1949), Whittle (1963), Box and Jenkins (1979)].

The criterion applied by X-11-ARIMA to evaluate the forecasting (backcasting, if applicable) performance of the three built-in models and user-supplied models is the mean-absolute percentage error, (MAPE) estimated by:

$$(3.1.6) \quad D_e = \frac{1}{N} \sum_{\ell=1}^{N} \left| \frac{Z_{t+\ell} - \hat{Z}_t(\ell)}{Z_{t+\ell}} \right| \times 100$$

This linear cost function generally gives a good approximation to the non symmetric cost functions $C(e) \neq C(-e)$ that are likely to arise in the seasonal adjustment of economic time series. For the majority of these series, to generate forecasts that overestimate or underestimate a cyclical turning point that signals a recession will not have the same cost for the policy makers. Furthermore, the position of the forecasting error in the time is of high importance for the forecasts enter with different weights in the seasonal adjustment of the series. Finally, we should note that under the assumption of normality implicit in the ARIMA models, the extrapolations $\hat{z}_t(\ell)$ are also optimal in the mean absolute error (MAE) sense.

The mean absolute error in (3.1.6) is expressed in percentage of the level of the unadjusted series to enable the inclusion, into the automatic model selection option, of an upper bound of acceptance valid for a large class of series.

3.2. <u>The upper bound of ARIMA extrapolation errors for the X-11-ARIMA</u>

In the X-11-ARIMA program, the mean absolute percentage error of the forecasts and backcasts (if applicable) are calculated for each one of the first and last three years of the series under study and, then, averaged. The extrapolations are produced from different points of origin for one year at a time and compared with the corresponding observed values to determine the magnitude of the error. By looking only at the last three years, the forecasting performance of the model is mainly evaluated by its ability to trace as close as possible the recent movements of the original series.

The built-in ARIMA models pass the extrapolation criterion if the mean absolute forecasting error for the last three years is smaller than 12%. The model giving the smallest average forecasting error is chosen to backcast if the mean absolute backcasting error for the first three years is smaller than 18%. These upper limits of acceptance are based on the results obtained from the experiment described below, performed with a sample of 92 irregular or volatile series that passed the fitting criteria. The series were obtained from the labour force sector, external trade, retail trade, construction, transportation, manufacturer shipments and inventories. The average relative contribution of the variance of the irregulars to the total variance of the series (for one month span) as measured in table F2B of the X-11-ARIMA program is 17% for the whole sample, and the minimum and maximum values are 10% and 38% respectively. These ratios provide a good approximation to the degree of predictability of the series if the irregulars and the other systematic components are non-correlated and their estimation is optimal. A similar ratio between the variance of the forecasting error and the variance of the original series is analyzed by Granger and Newbold (1977) as a criterion for evaluating the forecasts and a measure of predictability of the series.

The design of the experiment with the 92 series consists of:

(1) Choosing the built-in ARIMA model that provides the smallest mean absolute forecasting error for every series ending in 1973, 1974 and 1975; and of extending every series with one year of forecasts;

(2) Separately seasonally adjusting every series with and without the ARIMA option, thus obtaining two sets of seasonal factor forecasts. The seasonal factor forecasts were generated for year 1974, 1975 and 1976;

(3) Calculating the mean absolute percentage error (MAPE) of the seasonal factor forecasts for all the months over the three years. The error is defined as the difference between the forecast and the corresponding "final" seasonal factor when the series ends four years later, in 1977, 1978 and 1979 respectively. Six values missing for 1979 are replaced with extrapolated values.

(4) Comparing the mean absolute errors of the seasonal factor forecasts generated with and without the ARIMA option.

The conclusions are summarized in table 1 which shows that: (a) if the mean absolute forecasting error of the ARIMA model is smaller than or equal to 12%, there is a 15% reduction in the mean absolute error of the seasonal factor forecasts generated with the ARIMA option; (b) if the mean absolute forecasting error of the ARIMA model falls between 12% and 16% there is a 5% reduction, (c) if the mean absolute forecasting error of the ARIMA model falls between 16% and 25% there is a

1% reduction, and (d) if the mean absolute forecasting error falls between 25% and 30%, there is no reduction. The nine series with a mean absolute forecasting error greater than 30%, were not seasonally adjusted with the ARIMA option because of their high forecasting error.

Table 1

Comparisons of mean absolute percentage error of seasonal factor forecasts generated with and without the ARIMA option for various range of mean absolute ARIMA forecasting percentage error with a sample of 92 highly irregular series

Number of series	MAPE of the forecasts from a built-in ARIMA model in X-11-ARIMA	Ratio of the MAPE of the seasonal factor forecasts with and without the ARIMA option MAPE(ARIMA)/MAPE(NO-ARIMA)
40	$D_e \leq 12\%$	.85
26	$12\% < D_e \leq 16\%$	.95
10	$16\% < D_e \leq 25\%$	.99
7	$25\% < D_e \leq 30\%$	1
9	$30\% < D_e$	N/A

Since the experimentation is based on highly irregular or volatile series, the 12% mean absolute forecasting error is the upper bound of acceptance for the built-in models of theX-11-ARIMA program. The upper bound of 18% for the backcasts is obtained by comparing the mean absolute error of the forecasts for the last three years with the mean absolute error of the backcasts for the first three years of every series. The ratio of the two MAPEis 1.5 indicating that on average, the backcasting error is 50% higher.

The above guidelines should certainly not be followed blindly and to avoid this, the X-11-ARIMA method provides an option for users-supplied models where the fixed bounds of extrapolation error and level of significance for the test of fit can be violated. The analyst should use all valuable information available to decide what should be the levels of acceptance of the extrapolation error and χ^2 probability value for the particular series under study. For the built-in models, a rather conservative approach has been adopted.

The program will choose the model with the lowest mean absolute forecasting error, but if the difference with respect to the other two models is small, it is recommendable to select the model with less parameters. One would like to determine, where possible, whether the difference is significant or not under the usual criteria of statistical significance. It is tempting in this case to apply a "t" test for the difference between two means but this would be inappropriate for two reasons. First, the errors generated by one model are likely to be correlated with those produced by the other model. Second, for forecasts of more than one lead time, the errors are autocorrelated. In fact, the ℓ-step ahead forecast error constitutes in general a moving average process of order ℓ-1. Granger and Newbold

(1977) discuss a test that can be applied when the errors for the two models are correlated but they should not be autocorrelated and, thus, it is applicable to one-lead time forecasts only.

4. SELECTION OF ARIMA MODELS FOR POSITIVE SERIES WITH ADDITIVE DECOMPOSITION MODELS

In an earlier version of the X-11-ARIMA program, the built-in models were log $(0,1,1)(0,1,1)_s$, log $(0,2,2)(0,1,1)_s$ and $(2,1,2)(0,1,1)_s$ so that the two first models were applied to previously log transformed data. It was often observed, however, that if the decomposition model for the seasonal adjustment of a series was additive (observed series equal to trend-cycle plus seasonals plus irregulars) the two log models generated residuals with a very low p value for the χ^2 distribution, whereas the third non log model passed well the criteria of fitting. We suspected that the logarithmic transformation was somehow destroying the properties of the original series and that probably the same simple models applied to non log transformed data would be adequate. 120 American and Canadian series that followed an additive decomposition model were fitted with the $(0,1,1)(0,1,1)_s$ and $(0,1,1)_s$ models before and after the logarithmic transformation. The results confirmed our suspicions and indicated, furthermore, that also the mean absolute percentage error of the forecasts was often deteriorated for the log models. Because of this and the theoretical reasons discussed in the next section, the three built-in ARIMA models of the current version of X-11-ARIMA are applied to non log transformed data whenever the series follows an additive decomposition model.

4.1. The effect of the log transformation for positive series with additive decomposition models

Let Z_t be an observed time series of positive values that can be additively decomposed into a systematic part X_t and a random part U_t which is covariance stationary and non-correlated with X_t.

$$Z_t = X_t + U_t \tag{4.1.1}$$

The systematic part X_t (trend, cycle, seasonals) can be represented either by well defined functions of time, in which case, is said to be deterministic, or, by a stochastic process. The case that interests us is this latter and then, under the conditions of normality and stationarity, the series Z_t can be modelled by an ARIMA process. Under these conditions, the ARIMA model is a minimum mean square (absolute) linear predictor of Z_t. Because these conditions are so important, it is often necessary to transform economic time series to fulfill them. If the process is non-stationary only in the mean, it can possibly be made stationary by applying the difference operator of a given order. There are other cases in which the series Z_t is non stationary in the mean and in the variance, and, then, other types of transformation should be applied to remove both kinds of non-stationarity. The family of Box and Cox (1964) power transformations:

$$Z^\lambda = \begin{cases} \dfrac{Z^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log Z & \text{for } \lambda = 0 \end{cases} \tag{4.1.2}$$

where $Z>0$ and $0<\lambda<1$, is often applied, particularly, the logarithmic and the square root transformations corresponding to $\lambda = 0$ and $\lambda = 1/2$, respectively. The logarithmic transformation is appropriate for processes that follow a multiplicative relationship,

$$(4.1.3) \quad Z_t = F(t)Y_t$$

where F(t) is a positive function of time and Y is a stationary process. The (4.1.3) can be made stationary by applying first the log transformation and then, removing the trend in mean by proper differencing. Economic time series have often similar trends in mean and variance but for the logarithmic transformation to be adequate, the trends must be proportional. If this is not the case, then other power transformations should be applied.

If the logarithmic transformation is applied to (4.1.1), we have,

$$(4.1.4) \quad Y_t = \log Z_t = \log (X_t+U_t) = \log X_t+\log (1+\frac{U_t}{X_t})$$

Expanding by Taylor (the subindex t is suppressed to simplify the notation)

$$(4.1.5) \quad \log (1+\frac{U}{X}) = \frac{U}{X} - \frac{1}{2}\frac{U^2}{X^2} + \frac{1}{3}\frac{U^3}{X^3} \ldots \quad \text{for all } \left|\frac{U}{X}\right| <1$$

and hence,

$$(4.1.6) \quad Y \approx \log X + \frac{U}{X}$$

The log transformed data can no longer follow an additive decomposition and simultaneously preserve the original statistical properties, for in (4.1.6) the random part is correlated with the systematic part. The estimation of Y_t by an ARIMA model will likely show autocorrelation of the residuals. Only in the case that X_t follows an exponential function, the influence of $\frac{U_t}{X_t}$ is negligible.

The mean, variance and autocovariance of the transformed series $Y_t = \log Z_t$ can be calculated by means of the Hermite polynomials using an approach similar to the one followed by Granger and Newbold (1976) to analyse the properties of forecasts for transformed series.

If Z_t is already a normal stationary stochastic process with mean μ, variance σ^2 and autocorrelation $\rho(\tau)$ then, the standardized variable

$$(4.1.7) \quad W_t = \frac{Z_t - \mu}{\sigma} \qquad \text{for all } t$$

is such that W_t, $W_{t-\tau}$ are jointly distributed as bivariate Normal with mean zero, unit variance and correlation $\rho(\tau)$.

The instantaneous transformation

$$(4.1.8) \quad Y_t = \log Z_t = \log (\sigma W_t+\mu)$$

can be expanded in a series of Hermite polynomials $H_n(W_t)$ (Erdélyi, et als, (1953) defined as, (4.1.9) $$H_n(W_t) = \frac{(-1)^n \phi^n(x)}{\phi(x)}$$

where $\phi(x)$ is the standard Normal probability density function. Because these polynomials are orthogonal with respect to the standard Normal p.d.f. and $H_0(W)=1$, the

(4.1.10) $$E(H_n(W)) = 0 \qquad \text{for all } n>0$$

and

(4.1.11) $$E[H_n(W)\,H_k(W)] = \begin{cases} 0 & \text{if } n \neq k \\ n! & \text{if } n = k \end{cases}$$

The (4.1.8) can be expanded as (suppressing the subindex t):

(4.1.12) $$\log Z = \log(\mu+\sigma W) = \log\mu + \log\left(1+\frac{\sigma W}{\mu}\right) = \log\mu + \frac{\sigma W}{\mu} - \frac{\sigma^2 W^2}{2\mu^2} + \frac{\sigma^3 W^3}{3\mu^3} - \ldots \quad \text{for all } \left|\frac{\sigma W}{\mu}\right| < 1$$

then using the first two terms of the approximation, the mean of the log transformed series can be approximated as:

(4.1.13) $$E(\log Z) \simeq E(\log\mu + \sigma W) = \log\mu - \frac{\sigma^2}{2\mu^2}$$

The variance of (4.1.12) is approximately equal to:

(4.1.14) $$E[\log Z - E(\log Z)]^2 \simeq E\left[\frac{\sigma W}{\mu} - \frac{\sigma}{\mu}(W^2-1)\right]^2 =$$

$$E\left[\frac{\sigma^2 W^2}{\mu^2} + \frac{\sigma^4}{4\mu^4}(W^2-1)^2 - \frac{\sigma^3}{\mu} W(W^2-1)\right] = \frac{\sigma^2}{\mu^2}\left(1+\frac{\sigma^2}{2\mu^2}\right)$$

and finally the covariance between $\log Z_t$ and $\log Z_{t+\tau}$ is

(4.1.15) $$\text{cov}(\log Z_t, \log Z_{t+\tau}) \simeq$$

$$E\left[\left(\frac{\sigma^2}{2\mu^2} + \frac{\sigma}{\mu} W_{t+\tau} - \frac{\sigma^2}{\mu} W_{t+\tau}^2\right)\left(\frac{\sigma^2}{2\mu^2} + \frac{\sigma}{\mu} W_t - \frac{\sigma^2}{2\mu^2} W_t^2\right)\right] =$$

$$\frac{\sigma}{\mu^2}\rho(\tau) - \frac{\sigma^3}{2\mu^3} E(W_{t+\tau} W_t^2) - \frac{\sigma^4}{4\mu^4} - \frac{\sigma^3}{2\mu^3} E(W_t W_{t+\tau}^2) + \frac{\sigma^4}{4\mu^4} E(W_t^2 W_{t+\tau}^2)$$

FOOTNOTES:

(1) I am grateful to Sandra McKenzie who was the first to bring to my attention the existence of overdifferencing in some models automatically selected by an earlier version of the X-11-ARIMA when analysing external trade series at the U.S. Bureau of the Census. (For this latter, see Dagum, Mayes and McKenzie, 1979).

REFERENCES:

[1] Box, G.E.P. and Cox, D.R. An Analysis of Transformations, J. Royal Statistical Society, B.26, (1964) 211-243

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

[3] Box, G.E.P. and Pierce, D.A., Distribution of residual autocorrelations in autogressive-integrated moving average time series models, J.A.S.A. 65, (1970) 509-26.

[4] Dagum, Estela Bee, Seasonal Factor Forecasts from ARIMA Models, Proceedings of the International Statistical Institute 40th Session at Warsaw part 3, (1975).

[5] Dagum, Estela Bee, Modelling, Forecasting and Seasonally Adjusting Economic Time Series with the X-11-ARIMA Method, The Statistician, 27, Nos. 3 and 4 (1978) 203-216.

[6] Dagum, Estela Bee, The X-11-ARIMA Seasonal Adjustment Method, Statistics Canada Publication, Catalogue No. 12-564E (1980).

[7] Dagum, E.B., Mayes, F.J., and McKenzie, S., ARIMA Modeling and Seasonal Adjustment, Applications in U.S. Foreign Trade Series, Proceedings of the Business and Economic Statistics Section of the American Statistical Association Annual Meeting, Washington, D.C. August (1979).

[8] Davies, N., Triggs, C.M. and Newbold, P., The Significance Levels of the Box-Pierce portmanteau statistic in finite samples, Biometrika, 64, 3, (1977) 517-22.

[9] Erdelèyi, A. et als, Higher Transcendental Functions, Bateman Manuscript Project, Vol. II, New York: McGraw-Hill, (1953).

[10] Geweke, John, An Efficient Method for Revising Seasonally Adjusted Time Series, Social Sciences Research Institute Paper No. 7822, Madison Wisconsin University (1978).

[11] Granger, C.W.J. and Newbold, P., Forecasting Transformed Series, J. Royal Statistical Society, B 38 (1976) 189-203.

[12] Granger, C.W.J. and Newbold, P., Forecasting Economic Time Series, London: Academic Press (1977)

[13] Kolgomorov, A., Sur l'interpolation et extrapolation des suites stationnaires Comptes Rendu Acadèmie de Sciences, Paris 208 (1939) 2043-2045.

[14] Ljung, G.M. and Box, G.E.P., On a measure of lack of fit in time series models Biometrika 65 (1978) 297-307.

[15] Lothian, J. and Morry M., Selection of Models for the Automated X-11-ARIMA Seasonal Adjustment Program, Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada (1978).

[16] Pierce, David, A. Data Revisions with Moving Averages Procedures, Journal of Econometrics (forthcoming) (1980).

[17] Plosser, C.I. and Schwert, G.W., Estimation of a non-invertible moving average process, Journal of Econometrics 6 (1977) 199-224.

[18] Prothero, D.L. and Wallis, K.F., Modelling macroeconomic time series (with discussion, J. Royal Statistical Society, A 139 (1976) 468-500.

[19] Whittle, P., Prediction and Regulation by Linear Least Squares Methods, London: English Universities Press (1963).

[20] Wiener, N., The Extrapolation, Interpolation and Smoothing of Stationary Time Series, Boston: M.I.T. Press (1949).

[21] Wold, Herman, A Study in the Analysis of Stationary Time Series, Uppsala: Almquist and Wicksell (1938), (second edition, 1954).

2. AN ARIMA MODEL BASED APPROACH TO SEASONAL ADJUSTMENT

In this section we summarize the theory behind the ARIMA model based seasonal adjustment method developed in Hillmer and Tiao (1982). A similar treatment can be found in Burman (1980). We assume that an observable time series or some appropriate power transformation of the series, Z_t, can be represented as:

$$Z_t = S_t + N_t \tag{2.1}$$

where S_t and N_t are mutually independent seasonal and nonseasonal components. If desired, N_t can be further decomposed into trend and noise components; however, because the usual practice in the U.S. is to publish estimates of N_t, we shall only consider decomposition into two components.

If in (2.1) the stochastic structures of S_t and N_t are known, then minimum mean squared error estimates of S_t and N_t using the observed values of Z_t are readily obtained from the theory of signal extraction given in Whittle (1963) and Cleveland and Tiao (1976). In particular, suppose the models for the components are:

$$\begin{aligned} \phi_S(B)S_t &= \eta_S(B)b_t \\ \phi_N(B)N_t &= \eta_N(B)c_t \end{aligned} \tag{2.2}$$

where the pairs of polynomials in the backshift operator B, $\{\phi_S(B), \eta_S(B)\}$, $\{\phi_N(B), \eta_N(B)\}$, and $\{\phi_S(B), \phi_N(B)\}$ have no common zeros, and b_t and c_t are mutually independent, i.i.d. $N(0, \sigma_b^2)$ and $N(0, \sigma_a^2)$, respectively. It follows from (2.1) that the model for Z_t is

$$\phi(B)Z_t = \theta(B)a_t \tag{2.3}$$

where (Cleveland, 1972) $\phi(B) = \phi_s(B)\phi_N(B)$, and $\theta(B)$ and σ_a^2 are determined from the equation:

$$\sigma_a^2 \frac{\theta(B)\theta(F)}{\phi(B)\phi(F)} = \sigma_b^2 \frac{\eta_s(B)\eta_s(F)}{\phi_s(B)\phi_s(F)} + \sigma_c^2 \frac{\eta_N(B)\eta_N(F)}{\phi_N(B)\phi_N(F)} \quad (2.4)$$

with $F = B^{-1}$. Furthermore, when all the zeros of $\phi_s(B)$ and $\phi_N(B)$ are on or outside the unit circle, the estimated seasonal and nonseasonal components are:

$$\hat{S}_t = W_s(B)Z_t \quad \text{and} \quad \hat{N}_t = W_N(B)Z_t \qquad (2.5)$$

where

$$W_s(B) = \frac{\sigma_b^2\ \phi(B)\phi(F)\eta_s(B)\eta_s(F)}{\sigma_a^2\ \theta(B)\theta(F)\phi_s(B)\phi_s(F)} = \frac{\sigma_b^2\ \eta_s(B)\eta_s(F)}{\sigma_a^2\theta(B)\theta(F)}\phi_N(B)\phi_N(F)$$

and

$$W_N(B) = \frac{\sigma_c^2\ \phi(B)\phi(F)\eta_N(B)\eta_N(F)}{\sigma_a^2\ \theta(B)\theta(F)\phi_N(B)\phi_N(F)} = \frac{\sigma_c^2\ \eta_N(B)\eta_N(F)}{\sigma_a^2\theta(B)\theta(F)}\phi_s(B)\phi_s(F).$$

In practice, the S_t and N_t series are unobservable. Thus, without additional information, the component models (2.2) are unknown so that the weight functions $W_s(B)$ and $W_N(B)$ and, therefore, the estimates $\hat{S}_t$ and $\hat{N}_t$ cannot be computed. However, an accurate estimate of the model (2.3) can be obtained from the observable Z_t series. Consequently, it is of interest to investigate to what extent the component models can be determined from the model for Z_t, and to what extent prior knowledge about the component models is required to achieve a decomposition.

2.1 Prior Knowledge About S_t

Because X-11 has been widely used for many years, some of the characteristics it attributes to the seasonal component must have been regarded as desirable by users of seasonally adjusted data. In particular, additive X-11 produces estimates of S_t which aproximately repeat every year and approximately sum to zero over any twelve consecutive months. Also, Young (1968) shows that the X-11 procedure for estimating S_t may be approximated by a linear filter with weights that decrease from the center toward the end of the filter. This feature implicitly assumes that S_t evolves over time. Judging from these considerations, it seems reasonable that S_t should be a (nondeterministic) stochastic process, but that locally a regular seasonal pattern should be preserved.

If it is sensible to require that the sum of S_t over any twelve consecutive months should vary about zero, then the moving sum $U(B)S_t$, where $U(B) = 1 + B + \ldots + B^{11}$ should be a stationary time series with mean zero. In this case we can write $U(B)S_t = \eta_s(B)b_t$, where the b_t's are i.i.d. $N(0,\sigma_b^2)$, and $\eta_s(B)$ is as yet of unspecified degree. Notice that $E[U(B)S_t] = E[\eta_s(B)b_t] = 0$. If we also require S_t to locally follow a fixed seasonal pattern, the forecasting function at a given time origin of the model for S_t should follow a fixed pattern of period 12 and should sum to zero over twelve consecutive months. In other words, the model for S_t should not allow for predictable changes in the seasonal pattern; such changes should be part of the trend component. It is easy to show that these requirements are equivalent to restricting $\eta_s(B)$ to be of degree at most eleven. Thus, we are led to the following model for the stochastic seasonal component S_t:

$$U(B)S_t = \eta_s(B)b_t \qquad (2.6)$$

where $\eta_s(B)$ is a polynomial in B of degree at most eleven. Seasonal components following (2.6) will locally follow a fixed pattern, but as long as $\sigma_b^2 > 0$ the forecasting function of (2.6) will be continually updated as the time origin advances, thus the pattern in S_t will evolve over time.

2.2 Restrictions Imposed By The Data

We shall investigate how information available from the data embodied in a known model for Z_t, together with the additivity and independence assumptions (2.1), restrict the possible models for S_t and N_t. From the discussion in the previous section, Z_t will have a seasonal component if $\phi(B)$ contains the factor $\phi_S(B) = U(B)$. The autoregressive polynomial of N_t, $\phi_N(B)$, must have no zeros in common with $U(B)$, otherwise N_t would contain a seasonal component. Thus,

$$\phi(B) = U(B)\,\phi_N(B) \tag{2.7}$$

which determines $\phi_N(B)$ from $\phi(B)$, and the relationship (2.4) becomes:

$$\sigma_a^2 \frac{\theta(B)\theta(F)}{\phi(B)\phi(F)} = \sigma_b^2 \frac{\eta_S(B)\eta_S(F)}{U(B)U(F)} + \sigma_c^2 \frac{\eta_N(B)\eta_N(F)}{\phi_N(B)\phi_N(F)} . \tag{2.8}$$

It remains to determine the moving average polynomials $\eta_S(b)$ and $\eta_N(B)$, and the innovation variances σ_b^2 and σ_c^2. Any choice of moving average polynomials and variances satisfying (2.8), subject to the restriction that $\eta_S(B)$ is at most of degree eleven, will be an acceptable decomposition.

To determine an acceptable decomposition, a partial fractions expansion of the left hand side of (2.8) may be performed to yield

$$\sigma_a^2 \frac{\theta(B)\theta(F)}{\phi(B)\phi(F)} = \frac{Q_S(B)}{U(B)U(F)} + \frac{Q_N(B)}{\phi_N(B)\phi_N(F)} \tag{2.9}$$

where $Q_s(B) = q_{0s} + \sum_{j=1}^{10} q_{js}(B^j + F^j)$ and $Q_N(B)$ may be obtained by subtraction. The expansion (2.9) is unique because the order of $Q_S(B)$ is less than that of $U(B)U(F)$. In the case where an acceptable decomposition exists, if we set $B = e^{-i\omega}$ in (2.8) both terms on the right hand side are nonnegative for $0 \leq \omega \leq \pi$. Thus, to achieve an acceptable decomposition from the partial fractions expansion (2.9), we may need to modify the terms on the right hand side so they are nonnegative for $0 \leq \omega \leq \pi$, while their sum remains the same. Because the degree of $\eta_s(B)$ is restricted to be at most 11, the only possible modifications are the addition of a constant, γ, to the first term and the subsubtraction of γ from the second term. It follows from (2.9) by letting

$$\epsilon_1 = \min_{0 \leq \omega \leq \pi} \frac{Q_s(e^{-i\omega})}{|U(e^{-i\omega})|^2} \quad \text{and} \quad \epsilon_2 = \min_{0 \leq \omega \leq \pi} \frac{Q_N(e^{-i\omega})}{|\phi_N(e^{-i\omega})|^2}$$

that an acceptable decomposition exists if and only if $\epsilon_1 + \epsilon_2 \geq 0$. If $\epsilon_1 + \epsilon_2 > 0$, then the acceptable decomposition is not unique because there is an interval for the constant γ such that both

$$\frac{Q_s(e^{-i\omega})}{|U(e^{-i\omega})|^2} + \gamma \quad \text{and} \quad \frac{Q_N(e^{-i\omega})}{|\phi_N(e^{-i\omega})|^2} - \gamma$$

are nonnegative. Thus, when $\epsilon_1 + \epsilon_2 > 0$ the prior knowledge about S_t used to this point and the restrictions upon S_t imposed by the model for Z_t are not sufficient to determine a unique model for S_t. In this case we must further restrict the model for S_t based upon additional a priori assumptions about the seasonal component.

2.3 Canonical Decomposition and Justifications

Following the ideas originally given in Tiao and Hillmer (1978), Box, Hillmer, and Tiao (1978), Pierce (1978), and Burman (1980), we define a canonical decomposition of Z_t into S_t and N_t as follows. Within the range of choices of $\eta_s(B)$, $\eta_N(B)$, σ_b^2, and σ_c^2 satisfying (2.8), the <u>canonical decomposition</u> is that one which minimizes σ_b^2, the innovation variance for the seasonal component. This defining property is intuitively pleasing since the randomness of S_t arises from the sequence of b_t's. Thus, minimizing σ_b^2 selects the model for the seasonal component which is as deterministic as possible while remaining consistent with the information in the data.

Some additional properties of the canonical decomposition can be cited (see Hillmer and Tiao 1982). (i) Among the set of all acceptable decompositions, the canonical decomposition minimizes $Var[U(B)S_t]$. This is appealing since making $Var[U(B)S_t]$ small, combined with the fact that $E[U(B)S_t] = 0$, will ensure that the sum of S_t over any twelve consecutive months remains close to zero. (ii) If $\bar{S}_t$ denotes the canonical seasonal component and $\tilde{S}_t$ denotes any other acceptable seasonal component, then $S_t = \bar{S}_t + e_t$ where e_t is a white noise series with variance $\sigma_e^2 > 0$. In other words, every acceptable seasonal component may be viewed as the sum of the canonical seasonal and white noise. To define the seasonal component to be S_t seems unreasonable, since $\bar{S}_t$ is a highly predictable component which accounts for all of the seasonality in Z_t, while e_t is nonseasonal and completely unpredictable. Therefore, in the absence of other information about S_t, it is reasonable to define the seasonal component to be $\bar{S}_t$.

We believe that the canonical decomposition is an appropriate choice. If, however, there was a priori knowledge about S_t leading to a different acceptable decomposition, that choice could be justified. It is important to note that there is not enough information in the data to uniquely determine the model for S_t, so that some additional defining assumptions about the seasonal component must be made in order to carry out the seasonal adjustment. It is a strength of the model based approach that this fact is emphasized and that the assumptions being made are clearly specified.

2.4 Consideration of Some Special Problems

To be complete, we discuss some special problems that may arise. First, we have defined the canonical components only when $\phi(B)$ can be written as $U(B)\phi_N(B)$. While this covers many cases, it does not include all models that might be fit to seasonal data. One case that arises is where $\phi(B) = (1 - \phi_{12}B^{12})\phi^*(B)$ with $|\phi_{12}| < 1$ and $\phi^*(B)$ nonseasonal. In this case we would not decompose Z_t because, unless ϕ_{12} is very near 1, the annual pattern of an estimate of S_t that might be produced will probably change very rapidly. This behavior does not correspond well to the general idea of what a seasonal component or its estimate should look like. Generally speaking, we would recommend seasonally adjusting a series only when $\phi(B)$ contains $U(B)$

In practice, $U(B)$ enters into the model through seasonal differencing, i.e. $1 - B^{12} = (1 - B)\,U(B)$. After applying $1 - B^{12}$, we have found it more appropriate to account for remaining seasonality in the model through $\theta(B)$, e.g. $\theta(B) = (1 - \theta_{12}B^{12})\,\theta^*(B)$ where $\theta^*(B)$ is nonseasonal, rather than with additional seasonal autoregressive terms. If θ_{12} is small (say $\leq .4$) this choice might be made for convenience since we could well approximate $(1-.4B^{12})^{-1}$ with a seasonal autoregressive operator of low order. However, θ_{12} is typically much larger than .4. Thus, we have not found it necessary to deal with seasonal autoregressive operators apart from $U(B)$.

A final consideration involves series for which the seasonality is thought to be fixed, either from modeling the data or from a priori considerations. In this case it is easy to estimate and remove S_t (see Pierce 1978). One can fit monthly means to Z_t, or more frequently to $(1-B)Z_t$, constrain these to sum to zero, and subtract them out. We do not advocate subtracting monthly means from all series being adjusted, although this approach has been suggested by Pierce (1978) and Cleveland, Dempster, and Stith (1980).

2.6 Example

As an example to illustrate the ARIMA model based seasonal adjustment approach we consider the monthly time series of employed males aged 16 to 19 in nonagricultural industries, from January, 1965 to August, 1979. This series was obtained from the Bureau of Labor Statistics and is given in the Appendix. The data is plotted in Figure 2.1a. Judging from that plot it is evident that this series is seasonal, the level of the series is changing over time, and the variability over time is relatively stable. The sample autocorrelation functions of the series and selected differences are plotted in Figures 2.2 through 2.5. Examination of the sample ACF's suggests regular and seasonal differences of the series to achieve stationarity. The sample ACF of the differenced series $W_t = (1 - B)(1 - B^{12}) Z_t$ indicates that the model

$$(1 - B)(1 - B^{12}) Z_t = (1 - \theta_1 B)(1 - \theta_{12} B^{12}) a_t \qquad (2.10)$$

may be appropriate for this series. The parameters in model (2.10) were estimated (using the TSPACK program of Liu 1979) and the estimates are reported in Table 2.1. The residual autocorrelations are plotted in Figure 2.6 and

TABLE 2.1

Parameter Estimates For Model (2.10)

Parameter	Estimate	Standard Error
θ_1	.27	.073
θ_{12}	.82	.037

the standardized residuals are plotted in Figure 2.7. These plots reveal no model inadequacies. In addition, the Ljung-Box (1978) test statistic for overall model fit based upon 36 lags equals 39.7. This is less than 48.6, the level .05 chi-squared critical value with 34 degrees of freedom. Thus, (2.10) appears to be an adquate model for this data.

Assuming that the parameter estimates reported in Table 2.1 are the true values, the theory of Section 2.3 can be applied to estimate the canonical seasonal and nonseasonal components for this example. The estimated canonical nonseasonal and seasonal conponents are plotted in Figures 2.1a and 2.1b. The nonseasonal component captures the underlying movements of the series and the seasonality in this series is relatively stable.

Figures 2.1a and 2.1b through 2.7 here

3. TRADING DAY AND HOLIDAY VARIATION

It is not unusual for monthly economic time series to be affected by the composition of the calendar. The two primary calendar influences are trading day effects and holiday effects. In this section, we summarize the results in Bell and Hillmer (1981) on modeling series containing these effects.

3.1 Trading Day Variation

Suppose the level of retail sales in a type of business (e.g. grocery stores) is greater on Friday and Saturday than on other days of the week. Over the years the same month, say January, will contain a different number of Fridays and Saturdays so that the level of retail sales for January in a given year will be affected by the number of Fridays and Saturdays in that particular January. The variation in a monthly time series that is related to the day of the week composition of the calendar is called trading day variation. We might expect that economic time series on sales, production, shipments, monetary activity, and service activity may all be subject to trading day variation. Thus, it is important to be able to deal with this in modeling and seasonally adjusting these time series.

In modeling series with trading day variation we suppose that the trading day effect can be approximated by a deterministic model. Let TD_t denote the trading day factor for month t, then TD_t will be a function of the number of distinct types of days in month t. In particular, we suppose that

$$TD_t = \sum_{i=1}^{7} \gamma_i X_{it} \tag{3.1}$$

where X_{it}, i=1, ..., 7, are respectively the number of Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays, and Sundays in month t, and γ_i, i=1, ..., 7, are parameters. The model (3.1) is appropriate for flow series, such as sales, where the monthly values can be thought of as the accumulation of daily values. For stock series other approaches should be considered (see, for example, Cleveland and Grupe 1982).

The model (3.1) can be written as

$$TD_t = \sum_{i=1}^{6} (\gamma_i - \bar{\gamma})(X_{it} - X_{7t}) + \bar{\gamma} \sum_{i=1}^{7} X_{it} = \sum_{i=1}^{7} \beta_i T_{it} \tag{3.2}$$

where $\bar{\gamma} = 1/7 \sum_{i=1}^{7} \gamma_i$, $\beta_i = \gamma_i - \bar{\gamma}$ and $T_{it} = X_{it} - X_{7t}$ for i=1,...,6, $\beta_7 = \bar{\gamma}$, and

$T_{7t} = \sum_{i=1}^{7} X_{it}$ denotes the length of month t. The parameterization (3.2) is more convenient than (3.1) because estimates of the γ_i's tend to be highly correlated while estimates of the β_i's are less correlated. Also, when making inferences the differential effects $\beta_1,\ldots,\beta_6$ are of more interest than the γ_i's.

In addition to the trading day variation characterized by (3.2), we must also deal with autocorrelation, trends, and seasonality. To do this we assume that apart from trading day effects the series follows an ARIMA model. Thus, letting Z_t^* denote the value of an observed time series at month t including trading day effects, an additive model for Z_t^* is

$$Z_t^* = TD_t + Z_t \tag{3.3}$$

where TD_t is defined by (3.2) and Z_t follows (2.3).

Modeling Strategy

In building models of the form (3.3) the following approach has been found effective.

(i) Direct identification of the ARIMA model for Z_t is often difficult because of the confounding of the autocorrelation pattern of Z_t with the influences of TD_t. However, one can typically determine the appropriate degree of differencing (d and D of $(1-B)^d (1-B^{12})^D$ in the model for Z_t) by examining the sample autocorrelation function of Z_t^* and its differences in the usual way.

(ii) Given d and D, the trading day effects can be approximately removed by regressing $(1 - B)^d (1 - B^{12})^D Z_t^*$ on $(1 - B)^d (1 - B^{12})^D T_{it}$, $i=1, \ldots, 7$. The model for Z_t can then be identified by examining the sample autocorrelation function (SACF) and partial autocorrelation function (SPACF) of the residuals from this regression.

(iii) Once a model for Z_t has been tentatively identified, maximum likelihood estimates of the trading day and time series parameters can be computed, and the results of Pierce (1971) can be used to make inferenc[e]

(iv) Standard diagnostic checks described in Box and Jenkins (1970) can used to assess model inadequacy and suggest directions of improvement.

Removal of Estimated Trading Day Variation in Seasonal Adjustment

Given a model for Z_t^*, the estimated trading day effects are $\hat{TD}_t$

$\sum_{i=1}^{7} \hat{\beta}_i T_{it}$ where $\hat{\beta}_i$ $i = 1, \ldots, 7$ are the estimated trading day parameters.

For seasonal adjustment it is desirable that the long run average of the trading day adjustment factors be zero. Now

$$\frac{1}{n} \sum_{t=1}^{n} \sum_{i=1}^{6} \beta_i T_{it} = \sum_{i=1}^{7} (\gamma_i - \bar{\gamma})[\frac{1}{n} \sum_{t=1}^{n} (X_{it} - X_{7t})] \doteq 0$$

for large n since the $1/n \sum_{t=1}^{n} (X_{it} - X_{7t})$ are approximately constant while $\sum_{i=1}^{7} (\gamma_i - \bar{\gamma}) = 0$. Furthermore, we can write

$$T_{7t} = (T_{7t} - LF_t - 30.4375) + LF_t + 30.4375$$

where

$$LF_t = \begin{cases} .75 \text{ for a February in a leap year} \\ -.25 \text{ for a February in a non-leap year} \\ 0 \text{ otherwise} \end{cases}$$

and

$$30.4375 = \frac{365.25}{12} .$$

It can be easily seen that $(T_{7t} - LF_t - 30.4375)$ sums to zero over any twelve consecutive months and LF_t sums to zero over any 48 consecutive months. Thus, from

$$TD_t = \sum_{i=1}^{6} \beta_i T_{it} + \beta_7 LF_t + \beta_7 (T_{7t} - LF_t - 30.4375) + \beta_7 (30.4375) \quad (3.4)$$

we see the sum of the first two terms on the right hand side of (3.4) is the trading day adjustment factor, while the third term is part of the seasonal component, and the fourth part of the nonseasonal component.

To seasonally adjust a series with trading day effects we first form $\dot{Z}_t = Z_t^* - \hat{TD}_t$ where $\hat{TD}_t = \sum_{i=1}^{6} \hat{\beta}_i T_{it} + \hat{\beta}_7 LF_t$. Then we apply the results of

Section 2 by computing $\hat{N}_t = W_N(B)\dot{Z}_t$ and $\hat{S}_t = W_S(B)\dot{Z}_t$. When this is done the deterministic seasonal effect $\hat{\beta}_7$ $(T_{7t} - LF_t - 30.4375)$ is automatically assigned to $\hat{S}_t$ since $W_N(B) + W_S(B) = 1$ and $W_N(B)[T_{7t} - LF_t - 30.4375] = 0$ because $U(B)[T_{7t} - LF_t-30.4375] = 0$ (see(2.4) and (2.5) and recall $\phi_S(B) = U(B)$). Also, assuming $\phi(B)$ contains 1-B (in most applications it contains the factor $1-B^{12} = (1-B)U(B)$), $\hat{\beta}_7(30.4375)$ is automatically assigned to $\hat{N}_t$ since then $\phi_N(B)$ contains 1-B and $W_S(B)(30.4375) = 0$ because $\phi_N(B)(30.4375) = 0$. An estimate of the combined trading day and seasonal component is $\overline{TD}_t + \hat{S}_t$.

Trading Day Example

To illustrate the ideas in this section we analyze the monthly series of U.S. wholesale sales of hardware, plumbing, heating equipment, and supplies from January, 1967 through November, 1979. The data were obtained from the U.S. Census Bureau and are given in the Appendix. The data are plotted in Figure 3.11a. From the plot of the data it is apparent that the series is seasonal, the level increases over time, and the variability increases with the level. To stabilize the variability we took logarithms of the data; these are plotted in Figure 3.1.

The SACFs of the logged series and selected differences are plotted in Figures 3.2 to 3.5. We conclude that a first difference and possibly a twelfth difference are necessary to induce stationarity. The SACF of the first and twelfth differenced series exhibits a complex pattern. Analysts at the Census

Figures 3.1 - 3.5 here

Bureau indicated that this series is influenced by trading day variation; consequently, the pattern in Figure 3.5 may be due to confounding by the trading day effects. To approximately remove the trading day variation the model

$$(1 - B)(1 - B^{12})Z_t^* = \sum_{i=1}^{7} \beta_i (1 - B)(1 - B^{12}) T_{it} + N_t \qquad (3.5)$$

was fit by least squares to the logged data, Z_t^*. The SACF of the residuals from (3.5) is plotted in Figure 3.6. The most pronounced feature is the negative spike at lag 12 suggesting the model

$$Z_t^* = \sum_{i=1}^{7} \beta_i T_{it} + \frac{(1 - \theta_{12}B^{12})}{(1 - B)(1 - B^{12})} a_t \quad . \qquad (3.6)$$

Note that before allowing for trading day efects, the autocorrelation function (Figure 3.3) does not necessarily indicate that twelfth differencing is needed since the autocorrelations at multiples of lag 12 die out relatively quickly. Assuming we only need to first difference, the trading day effects may be approximately removed by consideration of the residuals from the least squares fit of the model

$$(1 - B) Z_t^* = \sum_{i=1}^{7} \beta_i (1 - B) T_{it} + N_t . \qquad (3.7)$$

The persistence of the autocorrelations of these residuals (plotted in Figure 3.7) at multiples of lag twelve now indicates that a seasonal difference is required to obtain a stationary noise model.

After estimation of the model (3.6), examination of the SACF of the residuals, plotted in Figure 3.8, reveals a significant negative spike at lag one.

Figures 3.6 - 3.8 here

Thus, we are led to consider the model

$$Z_t^* = \sum_{i=1}^{7} \beta_i T_{it} + \frac{(1 - \theta_1 B)(1 - \theta_{12} B^{12})}{(1 - B)(1 - B^{12})} a_t \qquad (3.8)$$

We see from the parameter estimates and t-ratios for model (3.8), reported in Table 3.1, that $\hat{\theta}_1$, $\hat{\theta}_{12}$, $\hat{\beta}_2$, $\hat{\beta}_4$, $\hat{\beta}_6$, and $\hat{\beta}_7$ are all statistically significant. The estimated trading day parameters $\hat{\beta}_1,\ldots,\hat{\beta}_6$ indicate that sales are higher on Tuesdays and Thursdays while lower on weekends (the estimated effect of Sunday, $\gamma_7 - \bar{\gamma}$, is -.015). However, note from Table 3.1 that the trading day parameter estimates are not independent, so that individual inferences about these parameters must be made with care. Since at least one of the β_i's is signifcantly different from zero, all of the trading day parameters will be retained for the purpose of seasonal adjustment.

Table 3.1 here

As diagnostic checks upon the adequacy of model (3.8), the SACF of the residuals is given in Figure 3.9 and the standardized residuals are plotted in Figure 3.10. The only possible concern is the series of negative residuals around the year 1975; for our purposes these were ignored. The Ljung-Box test statistic for overall model fit based upon 36 lags is 39.2 which is less than 48.6, the .05 chi-squared value with 34 degrees of freedom.

Figures 3.9 and 3.10 here

Assuming the model for the wholesale sales of hardware is (3.8) and the parameter estimates in Table 3.1 are the true values, it is possible to compute estimates of the trading day, canonical seasonal, canonical nonseasonal, and combined trading day-seasonal components. These are plotted in Figures 3.11a to 3.11d, where the estimated nonseasonal component has been transformed to the original metric of sales by exponentiation, and the estimates of the other components have been exponentiated and multiplied by 100.

From Figures 3.11c and 3.11d we notice that the estimated canonical seaadjustment for the seasonal is greater than that of the trading day. The trading day factors vary about 100 percent in an irregular fashion, which results in the combined trading day-seasonal component (Figure 3.11b) showing much more erratic behavior than the canonical seasonal. Finally, the estimated nonseasonal component seems to follow the underlying movements of the original data.

Figures 3.11a - 3.11d here

3.2 Easter Holiday Variation

Some economic time series are affected by holidays that recur each year at different times. The principle example of this for U.S. series is the increase in the level of some retail sales series in the days preceding Easter. Because Easter occurs at various dates during March and April, the monthly values for these months can be affected by the date of Easter each year. When the placement of holidays impacts the level of a series, it is important to develop models to account for these effects. Here we shall restrict attention to Easter holiday effects, although a similar approach could be used to model the effect of other holidays.

Series influenced by the placement of Easter may also exhibit trading day variation. Let Z_t^* denote an appropriately transformed time series containing Easter and trading day effects, and let E_t denote the Easter effect for month t. Then an additive model for Z_t^* is

$$Z_t^* = E_t + TD_t + Z_t \qquad (3.9)$$

where TD_t is given by (3.2) and Z_t by (2.3).

Specifying a functional form for E_t is not as easy as for TD_t because the placement of Easter affects daily sales for a period shorter than a month prior.

to Easter, while we typically have monthly data. As a first approximation, suppose there is a constant increase $\tilde{\alpha}$ in sales each day for τ days before Easter. Let $H(\tau,t)$ denote the proportion of the time period τ days before Easter that falls in month t. Then the Easter effect is

$$E_t = \tilde{\alpha}\,[\,\tau\, H(\tau,t)] = \alpha\, H(\tau\,,\,t) \qquad (3.10)$$

$$\text{where } \alpha = \tilde{\alpha}\tau\,.$$

The effect of the placement of Easter need not follow the pattern assumed in (3.10). However, a limited number of Easter dates will occur during the time frame of any series, and this limits the type of Easter effect that can be estimated. Bell and Hillmer (1981) show a way to test if the simple effect in (3.10) is reasonable for a given series and discuss more general effects.

Modeling Strategy

To model Easter variation we proceed in a manner similar to that discussed for trading day variation. Thus, we (i) identify the degree of differencing, $(1 - B)^d\,(1 - B^{12})^D$, from the SACF of the original series and its differences, (ii) remove the trading day and Easter effects in a preliminary fashion to allow a model for $Z_t = Z_t^* - TD_t - E_t$ to be identified, and (iii) efficiently estimate and check the entire model. To remove E_t and TD_t in (ii) and identify a model for Z_t we make a specific choice of τ in (3.10) (e.g. $\tau = 14$), regress $(1 - B)^d(1 - B^{12})^D Z_t^*$ on $(1 - B)^d(1 - B^{12})^D H(14,t)$ and $(1 - B)^d(1 - B^{12})^D T_{it}$, $i=1,\ldots,7$, and examine the SACF and SPACF of the residuals from this regression.

Parameter Estimation

The model is (3.9) with E_t given by (3.10) and TD_t by (3.2), so

$$Z_t^* = \alpha\, H(\tau\,,t) + \sum_{i=1}^{7} \beta_i\, T_{it} + Z_t. \qquad (3.11)$$

Assuming a suitable ARIMA model has been identified for Z_t, we wish to estimate the parameters of the model for Z_t^* including τ. Gaussian maximum likelihood estimates (MLE's) of the parameters can be obtained as follows.

Let $L(\alpha, \tau, \underset{\sim}{\beta}, \underset{\sim}{\phi}, \underset{\sim}{\theta}, \sigma_a^2)$ denote the log -likelihood function, where $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_7)'$, $\underset{\sim}{\phi} = (\phi_1, \ldots, \phi_p)'$, and $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_q)'$. Maximizing this over α, $\underset{\sim}{\beta}$, $\underset{\sim}{\phi}$, $\underset{\sim}{\theta}$, and σ_a^2 for fixed τ gives (asymptotically)

$$L_{max}(\tau) = \max_{\alpha, \underset{\sim}{\beta}, \underset{\sim}{\phi}, \underset{\sim}{\theta}, \sigma_a^2} L(\alpha, \tau, \underset{\sim}{\beta}, \underset{\sim}{\phi}, \underset{\sim}{\theta}, \sigma_a^2) = -\frac{n}{2} \ln \hat{\sigma}_a^2(\tau)$$

where $\hat{\sigma}_a^2(\tau)$ is the MLE of σ_a^2. Fitting (3.11) for a suitable set of τ's and picking the τ that minimizes $\hat{\sigma}_a^2(\tau)$ gives $\hat{\tau}$, the MLE of τ. The estimates $\hat{\alpha}$, $\hat{\underset{\sim}{\beta}}$, $\hat{\underset{\sim}{\phi}}$, $\hat{\underset{\sim}{\theta}}$, $\hat{\sigma}_a^2(\hat{\tau})$ from the fit with $\hat{\tau}$ are the MLE's of α, $\underset{\sim}{\beta}$, $\underset{\sim}{\phi}$, $\underset{\sim}{\theta}$, and σ_a^2.

For fixed τ (3.11) is linear in α and $\underset{\sim}{\beta}$, so the results of Pierce (1971) may be used to make inferences conditional upon τ. Unconditional inferences are difficult to make because $H(\tau,t)$ is a nonlinear and nondifferentiable function of τ. In practice, due to the limited number of observed Easter dates, it is unlikely that τ can be estimated with great accuracy, so that a range of values for τ will yield broadly similar estimates of the other parameters. In this case inferences made conditional upon $\tau = \hat{\tau}$ should not be misleading. In any event this can be checked by examining the estimates of the other parameters and their standard errors for various values of τ.

Seasonally Adjusting Series With Easter Effects

The Census Bureau's adjustment for Easter is done in the context of a multiplicative model and uses the reciprocal of the March adjustment factor for April. We use an additive decomposition on a suitably transformed series so a similar requirement imposed on our Easter factors would be that they sum to zero. So far we have accounted for the effect of Easter by including $E_t = H(\tau,t)$ in our model without worrying about this restriction.

Notice that $H(\tau,t)$ sums to 1 over any calendar year. We assume $\tau \leq 21$; then $H(\tau,t)$ is nonzero only for March and April. We let MA_t equal 1 in March and April and 0 otherwise and write

$$E_t = \alpha[H(\tau,t) - \frac{1}{2}MA_t] + \alpha[\frac{1}{2}MA_t - \frac{1}{12}] + \frac{\alpha}{12} \tag{3.12}$$

Now in (3.12) $[H(\tau,t) - \frac{1}{2}MA_t]$ is zero outside of March and April, and it sums to zero over March and April, thus satisfying the restrictions we desire for an Easter effect. Since $U(B)MA_t = 2$, $[\frac{1}{2}MA_t - \frac{1}{12}]$ in (3.12) sums to zero over any twelve consecutive months, a condition we desire of a seasonal effect. Thus, in (3.12) we let $\alpha[H(\tau,t) - \frac{1}{2}MA_t]$ be the Easter effect, $\alpha[\frac{1}{2}MA_t - \frac{1}{12}]$ be a seasonal effect, and $\frac{\alpha}{12}$ be part of the level of the series.

To seasonally adjust a series with Easter and trading day effects we first remove the estimated Easter and trading day effects and then apply the results of Section 2. Thus, we form $\dot{Z}_t = Z_t^* - \hat{\alpha}[H(\hat{\tau},t) - \frac{1}{2}MA_t] - \sum_{i=1}^{6}\hat{\beta}_i T_{it} - \hat{\beta}_7 LF_t$ and compute $\hat{N}_t = W_N(B)\dot{Z}_t$ and $\hat{S}_t = W_S(B)\dot{Z}_t$. When this is done the deterministic seasonal effect, $\hat{\alpha}[\frac{1}{2}MA_t - \frac{1}{12}]$, is automatically assigned to $\hat{S}_t$ since $W_N(B) + W_S(B) = 1$ and $W_N(B)[1/2\ MA_t - 1/12] = 0$ (see 2.5)). Similarly, $\hat{\alpha}/12$ is automatically assigned to N_t.

Holiday Example

We consider the time series of monthly retail sales of U.S. men's and boys' clothing stores from January, 1967 through September, 1979. The data is given in the Appendix. The plot of the original data in Figure 3.21a shows that it increases in level over time and that the seasonal amplitude varies with the level. Taking logarithms (Figure 3.12) seems to stabilize the seasonal amplitude. SACFs of the log series Z_t^*, and of $(1 - B)Z_t^*$, $(1 - B^{12})Z_t^*$, and $(1 - B)(1\ B^{12})Z_t^*$, are given in Figures 3.13-3.16. We need to take $(1 - B)(1 - B^{12})Z_t^*$ to get the autocorrelations to die out.

The behavior of the SACF of $(1 - B)(1 - B^{12}) Z_t^*$ at and near lags 36 and 48 is indicative of the Easter effect in this series as pointed out by Bell and Hillmer (1981). Figures 3.17 and 3.18 present the SACF and SPACF of the residuals from a regression of $(1 - B)(1 - B^{12})Z_t^*$ on $(1 - B)(1 - B^{12})H(14,t)$ and $(1 - B)(1 - B^{12})T_{it}$, $i=1, \ldots, 7$. Notice there are no longer spikes at lags 36 and 48 in the SACF. The SACF and SPACF suggest the tentative model:

$$Z_t^* = \alpha H(\tau, t) + \sum_{i=1}^{7} \beta_i T_{it} + Z_t \tag{3.13}$$

where $(1 - B)(1 - B^{12})Z_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta_{12} B^{12}) a_t$.

This model was fitted to the data with $\tau = 14$ and diagnostic checks revealed no model inadequacies.

Figures 3.12 - 3.20 here

We now estimate τ jointly with the other parameters in (3.13). This was done by estimating (3.13) for integer values of τ from 1 to 25. Some of the results are given in Table 3.2. We see $100\,\hat{\sigma}_a^2(\tau)$ is minimized around $\hat{\tau} = 9$, although a wide range of τ values work about as well. The residual ACF for the model with $\tau = 9$ is shown in Figure 3.19. It reveals no model inadequacies, and the Ljung-Box test statistic using 36 lags is 27.4, which is less than the χ^2_{33} five percent critical value of 47.4. The standardized residuals, plotted in Figure 3.20, do not indicate any problems with the model.

Table 3.2 here

We can now estimate the canonical seasonal and nonseasonal components in the manner discussed earlier. The components were first estimated in the log-metric and then exponentiated. Figure 3.21a shows the original data and estimated nonseasonal component and Figure 3.21b the estimate of the combined holiday-trading day-seasonal component. The pattern in Figure 3.21b varies from year to year, especially for the months of March, April, and December. Figures 3.21c-3.21e give the estimated seasonal, trading day, and holiday components respectively. We notice from Figure 3.21c that December has far and away the largest seasonal influence (due to Christmas) and that the pattern is fairly stable from year to year. The less regular pattern in Figure 3.21b is due to the erratic patterns of the trading day and holiday components.

Figures 3.21a - 3.21e here

4. DETECTION AND REMOVAL OF THE EFFECTS OF OUTLIERS

Economic and business time series observations are often subject to the influence of nonrepetitive exogenous interventions, e.g. strikes, outbreaks of wars, sudden changes in the market structure of a commodity, and unexpected heat or cold waves. When the timings of such interventions are known their effects can often be accounted for in a model using intervention analysis techniques proposed by Box and Tiao (1975). As an illustration, again let Z_t^* be the observable time series, and suppose that an intervention occurs at time t_0. The effect can often be modeled as

$$Z_t^* = \frac{\omega(B)}{\delta(B)} \xi_t^{(t_0)} + Z_t \qquad (4.1)$$

where $\xi_t^{(t_0)} = \begin{cases} 1 & \text{for } t = t_0 \text{ and} \\ 0 & \text{otherwise} \end{cases}$

signifies the time of occurrence of the intervention, $\omega(B) = \omega_0 - \omega_1 B - \ldots - \omega_s B^s$, $\delta(B) = 1 - \delta_1 B - \ldots - \delta_r B^r$, and the ratio $\omega(B)/\delta(B)$ describes the dynamic behavior of the intervention.

In practice, the timings of exogenous interventions are often unknown to the statistical analyst. Since the effects of the interventions can lead to bias in parameter estimates, and hence in forecasts and seasonal adjustments, it is important to develop procedures which can help detect and remove such effects.

This has come to be known as the problem of "outliers" or "spurious observations". In the Census X-11 procedure, outliers are handled in the filtering process. However, we believe that it is best to treat this problem as another element in the modeling process. If the time t_0 is known, this problem is not different in principle from that of modeling trading day or holiday variation. For the situation when t_0 is unknown, we summarize the results on outliers in time series of Chang and Tiao (1982), following earlier work by Fox (1972).

4.1 Additive and Innovational Outliers

We shall concentrate on two types of outliers, additive and innovational. An additive outlier (AO) is defined as

$$Z_t^* = Z_t + \omega \xi_t^{(t_0)} \tag{4.2}$$

while an innovational outlier (IO) is defined as

$$Z_t^* = Z_t + \frac{\theta(B)}{\phi(B)} \omega \xi_t^{(t_0)} \tag{4.3}$$

where Z_t follows model (2.3). In terms of the a_t's in (2.3), we have that

$$\text{(AO)} \quad Z_t^* = \frac{\theta(B)}{\phi(B)} a_t + \omega \xi_t^{(t_0)} \tag{4.4}$$

and

$$\text{(IO)} \quad Z_t^* = \frac{\theta(B)}{\phi(B)} \left\{ a_t + \omega \xi_t^{(t_0)} \right\}. \tag{4.5}$$

Thus, the AO case may be called a "gross error" model, since only the level of the t_0^{th} observation is affected. On the other hand, an IO represents an extraordinary shock at t_0 influencing Z_{t_0}, $Z_{t_0+1}, \ldots$ through the memory of the system described by $\theta(B)/\phi(B)$.

4.2 Estimation of ω When t_0 Is Known

To motivate the procedures for the detection of AO and IO, we discuss the situation when t_o and all time series parameters in the model (2.3) are known. Defining the residuals $e_t = \Pi(B)Z_t^*$, where $\Pi(B) = \phi(B)/\theta(B) = (1+\Pi_1 B + \Pi_2 B^2 + \ldots)$ we have that

$$\text{(AO)} \quad e_t = \omega \Pi(B)\xi_t^{(t_o)} + a_t \qquad (4.6)$$
$$\text{(IO)} \quad e_t = \omega \xi_t^{(t_o)} + a_t.$$

From least squares theory, estimators of the impact, ω, of the intervention and the variances of these estimators are

$$\text{(AO)} \quad \hat{\omega}_A = \rho^2 \Pi(F) e_{t_o}, \quad \mathrm{Var}(\hat{\omega}_A) = \rho^2 \sigma_a^2 \qquad (4.7)$$
$$\text{(IO)} \quad \hat{\omega}_I = e_{t_o}, \quad \mathrm{Var}(\hat{\omega}_I) = \sigma_a^2$$

where $\rho^2 = (1 + \Pi_1^2 + \Pi_2^2 + \ldots)^{-1}$. Thus, the best estimate of the effect of an IO at time t_o is the residual e_{t_o}, while the best estimate of the effect for an AO is a linear combination of e_{t_o}, e_{t_o+1}, ... with weights depending on the structure of the time series model. Note that the variance of $\hat{\omega}_A$ can be much smaller than σ_a^2.

If desired, one may perform various tests among the hypotheses,

H_o : $Z_{t_o}^*$ is neither an IO nor an AO
H_1 : $Z_{t_o}^*$ is an IO
H_2 : $Z_{t_o}^*$ is an AO

The likelihood ratio test statistics for IO and AO are

H_1 vs H_o $\quad \lambda_1 = \hat{\omega}_I / \sigma_a$
H_2 vs H_o $\quad \lambda_2 = \hat{\omega}_A / (\rho \sigma_a)$

On the null hypothesis H_o, λ_1 and λ_2 are both distributed as $N(0,1)$.

4.3 Detection of Outliers

In practice, t_o as well as the time series parameters are all unknown

If only t_0 is unknown, one may proceed by calculating λ_1 and λ_2 for each t, denoted by λ_{1t} and λ_{2t}, and then make decisions based on the sampling properties given above. The time series parameters (ϕ's, θ's, and σ_a) are also unknown, and it can be shown that the estimates of these parameters can be seriously biased by the existence of outliers. In particular, σ_a will tend to be overestimated. These considerations have led to the following iterative procedure to handle a situation in which there may exist an unknown number of AO or IO outliers.

(i) Model the series Z_t^* by supposing that there are no outliers (i.e. $Z_t^* = Z_t$) and from the estimated model compute the residuals

$$\hat{e}_t = \hat{\Pi}(B) Z_t^*$$

Let $\hat{\sigma}_a^2 = \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2$ be the initial estimate of σ_a^2.

(ii) Compute $\hat{\lambda}_{it}$, i=1,2 and t=1,...n, these being λ_{1t} and λ_{2t} with the estimated model. Let $|\hat{\lambda}_{t_0}| = \max_t \max_i [|\hat{\lambda}_{it}|]$. If $|\hat{\lambda}_{t_0}| = |\hat{\lambda}_{1t_0}| > c$, where c is a predetermined positive constant usually taken to be 3, then there is the possibility of an IO at t_0, and the best estimate of ω is $\hat{\omega}_{It_0}$. Eliminate the effect of this possible IO by defining a new residual $\tilde{e}_{t_0} = \hat{e}_{t_0} - \hat{\omega}_{It_0} = 0$. If, on the other hand $|\hat{\lambda}_{t_0}| = |\hat{\lambda}_{2t_0}| > c$, then there is the possibility of an AO at t_0, and the best estimate of its effect is $\hat{\omega}_{At_0}$. The effect of this AO can be removed by defining the new residuals $\tilde{e}_t = \hat{e}_t - \hat{\omega}_{At_0} \hat{\Pi}(B) \xi_t^{(t_0)}$, $t \geq t_0$. A new estimate $\tilde{\sigma}_a^2$ is computed from the modified residuals.

(iii) Recompute $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$ based on the same initial parameter estimates of the ϕ's and θ's but using the modified residuals and $\tilde{\sigma}_a^2$, and repeat the process (ii).

(iv) When no more outliers are found in (iii), suppose that k outliers (either IO or AO) have been tentatively identified at times $t_1, \ldots, t_k$. Treat these times as if they are known, and estimate the outlier parameters $\omega_1, \ldots, \omega_k$ and the time series parameters simultaneously using models of the form

$$Z_t^* = \sum_{j=1}^{k} \omega_j L_j(B) \xi_t^{(t_j)} + \frac{\theta(B)}{\phi(B)} a_t \qquad (4.9)$$

where $L_j(B) = 1$ for an AO and $L_j(B) = \frac{\theta(B)}{\phi(B)}$ for an IO at $t=t_j$. The new residuals are

$$\hat{e}_t^{(1)} = \pi^{(1)}(B) \left[Z_t^* - \sum_{j=1}^{k} \hat{\omega}_j L_j(B) \xi_t^{(t_j)}\right]. \qquad (4.10)$$

The entire process is repeated until all outliers are identified and their effects simultaneously estimated.

The above procedure is easy to implement since very few modifications to existing software capable of dealing with ARIMA and transfer function models are needed to carry out the required computations. Based on simulation studies, the performance of this procedure for estimating the autoregressive coefficient of a simple AR(1) model compares favorably with the robust estimation procedures proposed by Denby and Martin (1979) and Martin (1980). While the latter procedures cover only the AR case, our iterative procedure can be used for any ARIMA model.

4.4 Seasonal Adjustment

For the purposes of seasonal adjustment, trading day variation, holiday variation, and outliers can all be incorporated into the model based procedure by writing the model as

$$Z_t^* = TD_t + E_t + O_t + Z_t \qquad (4.11)$$

where $O_t = \sum^{k} \omega_j L_j(B) \xi_t^{(t_j)}$. The influences of all these effects can then be removed by setting $\dot{Z}_t = \hat{Z}_t^* - \widehat{TD}_t - \hat{E}_t - \hat{O}_t$ after which the techniques of Section 2.5 can then be used to decompose $\dot{Z}_t$ into the canonical seasonal, $\hat{S}_t$, and the canonical nonseasonal, $\hat{N}_t$ (allocating appropriate parts of $\widehat{TD}_t$ and $\hat{E}_t$ to $\hat{S}_t$ and $\hat{N}_t$ as was discussed in Sections 3.1 and 3.2). Finally, in keeping with current practice, the series adjusted for seasonal, trading day, and holiday variation is $\hat{N}_t + \hat{O}_t$ (in the transformed metric).

4.5 An Example

To illustrate the iterative procedure described above, we consider the logarithms of the monthly retail sales of variety stores from January, 1967 to September, 1979 obtained from the U.S. Bureau of the Census. This series contains trading day and Easter variation; as a preliminary step these effects were modeled as described in Section 3 and removed from the data.[2] The preprocessed data are reported in the Appendix and plotted in Figure 4.1.

The ARIMA model obtained from the preliminary analysis is

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - B^{12})Z_t^* = (1 - \theta_{12} B^{12})a_t \qquad (4.12)$$

where $\phi_1 = -.40$, $\phi_2 = -.27$, $\theta_{12} = .81$, and Z_t^* denotes the logarithms of the trading day and Easter adjusted sales. Examination of the statistics $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$ as described in (ii) of Section 4.3 indicates the possibility of an IO at t=112. After modification of the residuals, updating of the estimate of σ_a^2 and construction of new sequences for $\hat{\lambda}_{1t}$ and $\hat{\lambda}_{2t}$, an AO is identified at t=96. Further inner iterations revealed two additional candidate IO outliers, at t=113 and t=45.

For the second major iteration, the four tentatively identified outliers and the ARIMA parameters are simultaneously estimated. The process is repeated and an AO at t=121 together with an IO at t=114 are identified. The process is continued through five major iterations; Table 4.1 summarizes the results. The elimination of the eight outliers leads to changes in the estimates of the time series parameters and a large reduction in the estimated variance, $\hat{\sigma}_a^2$.

Figure 4.1 then Table 4.1 here

Discussion

For this example it is interesting to note that three consecutive IO's are identified at t=112, 113, and 114, suggesting an intervention effect different from that of an individual AO or IO. Discussions with analysts at the Census Bureau revealed that at around t=112 (which is April, 1976) a major variety store chain, W.T. Grant, went out of business, and that as a result a significant proportion of retail sales previously made at variety stores were shifted to department stores. Based upon this information it is reasonable to expect a level drop in the series starting at t=112. A model for this is

$$Z_t^* = Z_t + \frac{\omega}{1-B} \xi_t^{(112)} \quad \text{with } \omega < 0. \tag{4.13}$$

The important point is that the outlier analysis summarized in Table 4.1 is consistent with the above explanation, and a closer examination of consecutive outliers may reveal the nature of the intervention. To illustrate, the standardized residuals from the preliminary model for t=111, ..., 115 are reported below.

t	$\hat{e}_t/\hat{\sigma}_a$
111	.0
112	-5.3
113	-3.1
114	-2.1
115	-.2

If the model for Z_t^* is described by (4.13) with Z_t following the preliminary model (4.12), then the residuals are

$$e_t = a_t + \frac{(1 - \phi_1 B - \phi_2 B^2)(1 - B^{12})}{(1 - \theta_{12} B^{12})} \omega \xi_t^{(112)} . \qquad (4.14)$$

From the initial estimate of θ_{12} in Table 4.1 it follows that $\frac{1 - B^{12}}{1 - \theta_{12} B^{12}}$ is approximately one. From 4.14 and the initial estimates of ϕ_1 and ϕ_2 we would expect to observe large negative residual values at t=112, 113, and 114 that would behave somewhat like ω, $.4\omega$, and $.3\omega$. Also, there would be no effect on the residual at t=115. The three consecutive IO outliers noted above are thus reasonably consistent with the model (4.13).

If we consider the first differences of the data, $W_t^* = Z_t^* - Z_{t-1}$, then the model (4.13) reduces to

$$W_t^* = W_t + \omega \xi_t^{(112)}$$

where $W_t = Z_t - Z_{t-1}$. In other words, the intervention in (4.13) can be viewed as an AO at t=112 for the differenced data W_t^*. This is confirmed by applying the outlier detection procedure to W_t^* resulting in the identification of a single AO at t=112 instead of three consecutive outliers starting at that point

5. REVISIONS

Most seasonal adjustment methods recently proposed or currently in use use two sided filters. When current data are being adjusted the future values of the series required for the use of the two sided filters are not available. In practice the estimates of N_t are computed and subsequently changed as more data become available--these changes are known as <u>revisions</u>. Revisions present a practical problem since it can be difficult to explain to users of seasonally adjusted data why the current adjusted data get changed in subsequent years -- especially if the changes are large.

Retail Sales of Men's and Boys' Clothing Stores 1/67 - 9/79

237	187	241	245	259	296	252	260	271	267	320	54
266	216	252	297	302	310	270	288	280	316	372	59
319	249	287	320	342	329	291	321	315	361	400	68
338	268	304	313	348	350	321	317	333	364	396	71
336	267	303	375	382	401	341	351	357	382	447	77
364	310	379	408	439	451	390	413	424	469	534	88
452	361	426	470	477	502	424	442	442	479	562	96
437	368	427	495	514	492	443	500	458	492	542	88
459	403	490	467	556	542	474	510	483	527	591	104
495	404	463	540	518	552	505	502	496	558	629	113
511	440	496	578	542	550	492	518	507	569	708	114
480	421	532	536	542	563	508	554	552	609	763	129
561	462	564	582	586	615	553	612	570			

Retail Sales of Variety Stores 1/67 - 9/79
(Modified for Trading Day and Holiday Effects)

296	303	365	363	417	421	404	436	421	429	499	91
331	361	402	426	460	457	451	476	436	464	525	93
345	364	427	445	478	492	469	501	459	494	548	102
370	378	453	470	534	510	485	527	536	553	621	112
394	411	482	484	550	525	494	537	513	521	596	106
473	425	503	529	581	558	547	588	549	593	649	119
463	459	554	576	615	619	589	637	601	642	737	127
490	490	598	615	681	654	637	694	645	684	749	124
499	511	612	623	725	692	623	734	662	684	781	130
503	537	636	560	607	585	559	608	556	596	665	122
427	450	573	579	615	601	608	617	550	616	673	119
430	458	548	584	639	616	614	647	588	648	713	126
483	483	593	620	672	650	643	702	654			

References

Bell, W.R. and Hillmer, S.C. (1981), "Modeling Time Series with Calendar Variation," University of Kansas, School of Business Working Paper.

Box, G.E.P., Hillmer, S.C., and Tiao, G.C. (1978), "Analysis and Modeling of Seasonal Time Series," in Seasonal Analysis of Economic Time Series, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, 309-334.

Box, G.E.P. and Jenkins, G.M. (1970), Time Series Analysis: Forecasting and Control, San Francisco: Holden Day.

Box, G.E.P. and Tiao, G.C. (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," Journal of the American Statistical Association, 70, 70-79.

Burman, J.P. (1980), "Seasonal Adjustment by Signal Extraction," Journal of the Royal Statistical Society, Ser. A, 143, 321-337.

Chang, I. and Tiao, G.C. (1982), unpublished paper.

Cleveland, W.P. (1972), "Analysis and Forecasting of Seasonal Time Series," unpublished Ph.D. thesis, University of Wisconsin, Dept. of Statistics.

Cleveland, W.P., Dempster, A.P., and Stith, J. (1980), "Advances in Model-Based Seasonal Adjustment," Proceedings of the American Statistical Association, Business and Economics Statistics Section, 30-36.

Cleveland, W.P. and Grupe, M.R. (1982), "Modeling Time Series When Calendar Effects are Present," included in this volume.

Cleveland, W.P. and Tiao, G.C. (1976), "Decomposition of Seasonal Time Series: A Model for the Census X-11 Program," Journal of the American Statistical Association, 71, 581-587.

Cleveland, W.S, Dunn, D.M., and Terpenning, I.J. (1978), "SABL: A Resistant Seasonal Adjustment Procedure with Graphical Methods for Interpretation and Diagnosis," in Seasonal Analysis of Economic Time Series, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, 309-334.

Dagum, E.B. (1975), "Seasonal Factor Forecasts from ARIMA Models," Proceedings of the 40th Session of the International Statistical Insitute, Warsaw, 206-219.

Denby, L. and Martin, R.D. (1979), "Robust Estimation of the First-Order Autoregressive Parameter," Journal of the American Statistical Association, 74, 140-146.

Fox, A.J. (1972), "Outliers in Time Series," Journal of the Royal Statistical Society, Ser. B, 34, 350-363.

Geweke, J. (1978), "Revision of Seasonally Adjusted Time Series," SSRI Report Number 7822, Department of Economics, University of Wisconsin.

Hannan, E.J. (1967), "Measurement of a Wandering Signal Amid Noise," Journal of Applied Probability, 4, 90-102.

Hillmer, S.C. (1981), "An Empirical Comparison of Some Seasonal Adjustment Methods," Proceedings of the American Statistical Association, Business and Economic Statistics Section.

Hillmer, S.C. and Tiao, G.C. (1982), "An ARIMA Model Based Approach to Seasonal Adjustment," to appear in Journal of the American Statistical Association.

Ljung, G.M. and Box, G.E.P. (1978), "On a Measure of Lack of Fit in Time Series Models," Biometrika, 65, 297-303.

Macauley, F.R. (1931), The Smoothing of Time Series, New York: National Bureau of Economic Research.

Martin, R.D. (1980), "Robust Estimation of Autoregressive Models," in Directions in Time Series, eds. D.R. Brillinger and G.C. Tiao, Institute of Mathematical Statistics, 228-254.

Persons, W.M. (1919), "Indices of Business Conditions," Review of Economics and Statistics, 1, 5-107.

Pierce, D.A. (1971), "Least Squares Estimation in the Regression Model with Autoregressive-Moving Average Errors," Biometrika, 58, 299-312.

________ (1978), "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality are Present," in Seasonal Analysis of Economic Time Series, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, 242-269.

________ (1980), "Data Revisions With Moving Average Seasonal Adjustment Procedures," Journal of Econometrics, 14, 95-114.

Shiskin, J., Young, A.H., and Musgrave, J.C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Technical Paper No. 15, U.S. Bureau of the Census.

Sobel, E.L. (1967), "Prediction of a Noise-distorted, Multivariate, Non-stationary Signal," Journal of Applied Probability, 4, 330-342.

Tiao, G.C. and Hillmer, S.C. (1978), "Some Consideration of Decomposition of a Time Series," Biometrika, 65, 497-502.

Wecker, W. (1979), "A New Approach to Seasonal Adjustment," Proceedings of the American Statistical Association, Business and Economic Statistics Section, 322-323.

Whittle, P. (1963), Prediction and Regulation, New York: D. Van Nostrand.

Young, A.H. (1968), "Linear Approximation to the Census and BLS Seasonal Adjustment Methods," *Journal of the American Statistical Association*, 63, 445-471.

FOOTNOTES:

[1] Using (3.12) effectively assumes that the long run average Easter effect is $\omega/2$ in both March and April, whereas it really depends on τ. A more refined approach would be to replace MA_t with the long run average of $H(\tau,t)$ for each month.

[2] The effects of possible outliers were approximately removed so that the trading day and Easter parameter estimates would not be badly biased.

[3] This is not how X-11 is actually used, and will give smaller mean squared revisions than the way X-11 actually operates (by results of Pierce (1980) and Geweke(1978)).

[4] For X-11 ARIMA no account was taken of trading day variation in forecasting the observed series (it was handled in the usual X-11 way in adjusting the extended series) because the X-11 ARIMA computer program did not allow for it. Smaller revisions for X-11 ARIMA may have resulted if some allowance for trading day effects had been made in the forecasting.

[5] The ratios were computed in order to make comparisons across the 76 series. The revisions for the individual series for each of the three methods are available and will be provided on request.

TABLE 3.1

Parameter Estimates and Correlation Matrix for (3.8)

Parameter	Estimate	t-Ratio
θ_1	.22	2.6
θ_{12}	.75	12.1
β_1	.001	.3
β_2	.013	3.5
β_3	.004	1.1
β_4	.011	3.0
β_5	.001	.2
β_6	-.015	-4.0
β_7	.026	2.0

Correlation Matrix

θ_1	1.00								
θ_{12}	.07	1.00							
β_1	.04	.03	1.00						
β_2	.02	.04	-.57	1.00					
β_3	-.03	-.04	-.05	-.53	1.00				
β_4	.03	.01	.11	-.02	-.55	1.00			
β_5	-.08	-.02	.02	.09	-.03	-.56	1.00		
β_6	.04	-.02	.02	.02	.09	.01	-.56	1.00	
β_7	.05	.02	.14	-.13	.13	-.13	.05	.09	1.00

TABLE 3.2

Estimation of τ

τ	1	2	3	4	5	6	7	8	9	10	11	12
$100\tilde{\sigma}_a^2(\tau)$	.126	.122	.122	.123	.123	.122	.121	.1203	.1201	.1205	.122	.123

τ	14	15	16	17	18	19	20	21	22	23	24	25
$100\tilde{\sigma}_a^2(\tau)$	.127	.127	.128	.129	.129	.130	.131	.131	.131	.131	.131	.131

Estimation for $\tau = \hat{\tau} = 9$

Parameter	Estimate	t-Ratio
θ_1	.26	3.2
θ_2	.37	4.5
θ_{12}	.78	15.5
β_1	-.010	-1.8
β_2	-.002	-.4
β_3	.005	1.1
β_4	-.002	-.3
β_5	.011	2.2
β_6	.013	2.5
β_7	.014	.8
α	.070	7.7

TABLE 4.1

Variety Stores Outlier Detection

Major Iteration	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_{12}$	$\hat{\sigma}_a^2 \times 10^3$	Outlier Time	Outlier Type
(1)	-.40	-.27	.81	1.02	112	IO
					96	AO
					113	IO
					45	IO
(2)	-.53	-.27	.83	.65	121	AO
					114	IO
(3)	-.61	-.38	.85	.58	129	IO
(4)	-.64	-.36	.87	.54	103	IO
(5)	-.66	-.33	.89	.50		

TABLE 5.1

Ratio of Mean Squared First Year Revisions
(Model based/X-11).

One Month Ahead Forecasts

θ_{12} \ θ_1	.1	.3	.5	.7	.9
.1	1.56	1.38	1.19	1.19	1.28
.3	1.35	1.22	1.08	1.08	1.14
.5	1.06	.98	.89	.89	.93
.7	.66	.62	.58	.58	.60
.9	.15	.15	.15	.14	.15

Six Month Ahead Forecasts

θ_{12} \ θ_1	.1	.3	.5	.7	.9
.1	1.48	1.48	1.48	1.48	1.48
.3	1.27	1.27	1.27	1.27	1.27
.5	1.00	1.00	1.00	1.00	1.00
.7	.62	.62	.62	.62	.62
.9	.15	.15	.15	.15	.15

Twelve Month Ahead Forecasts

θ_{12} \ θ_1	.1	.3	.5	.7	.9
.1	1.28	1.31	1.36	1.41	1.45
.3	1.14	1.16	1.19	1.23	1.25
.5	.94	.95	.96	.98	.99
.7	.63	.63	.63	.63	.62
.9	.18	.17	.16	.16	.15

TABLE 5.2

Geometric Means of the Ratio of Model Based to X-11

Revisions in the Level of the Series

First Year	Second Year	Third Year
.59	.56	.56

Revisions in Month to Month Percentage Change

First Year	Second Year	Third Year
.56	.54	.54

TABLE 5.3

Geometric Means of the Ratio of Model Based to X-11 ARIMA

Revisions in the Level of the Series

First Year	Second Year	Third Year
.53	.55	.59

Revisions in Month to Month Percentage Change

First Year	Second Year	Third Year
.58	.56	.58

EMPLOYED NONAGRI MALES 16-19

Figure 2.1a

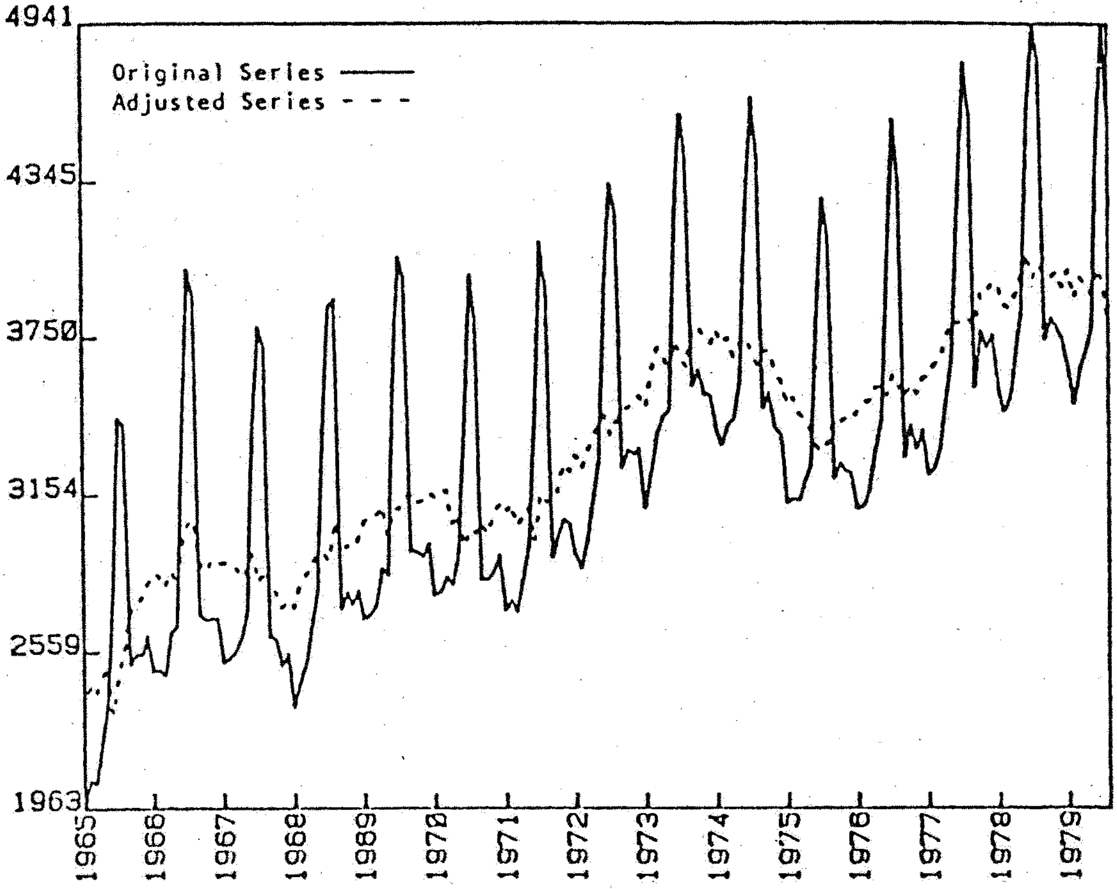


Figure 2.1b Seasonal Factors

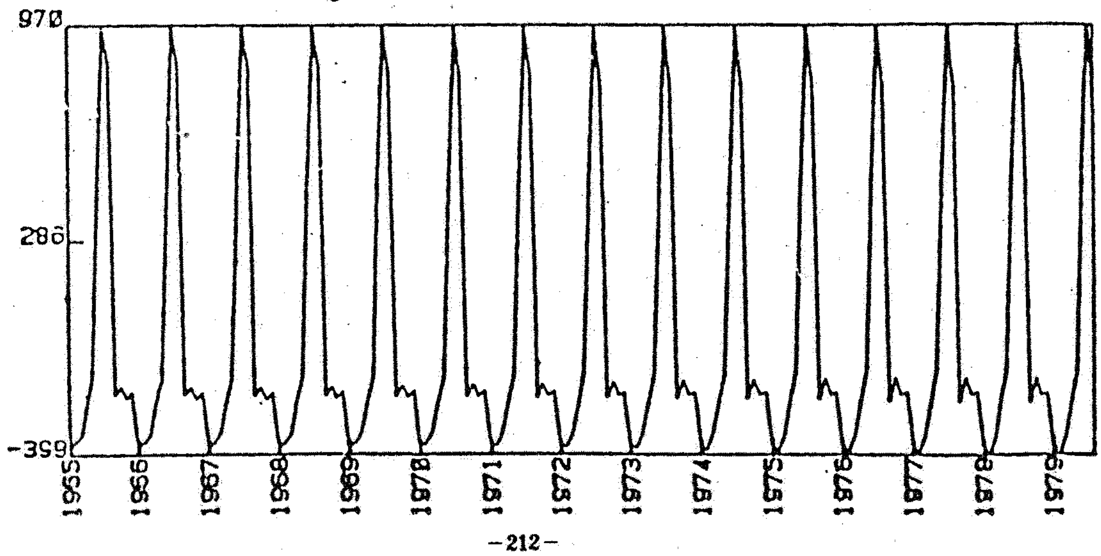
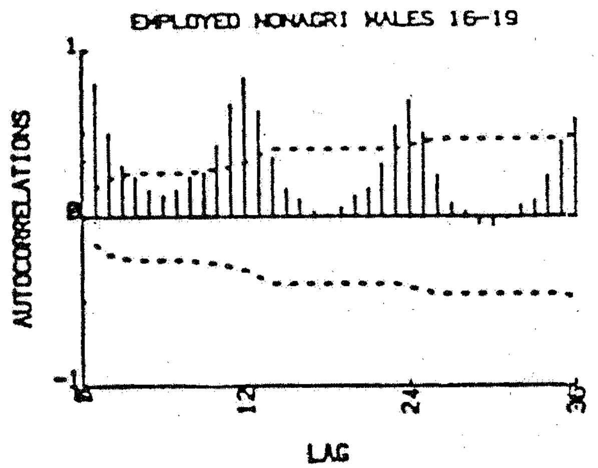


Figure 2.3



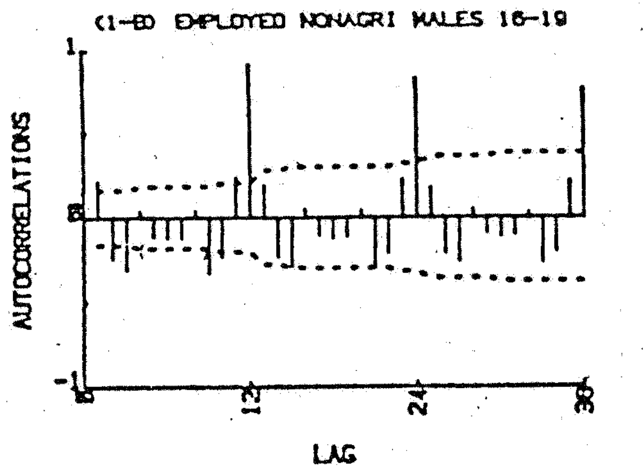


Figure 2.4

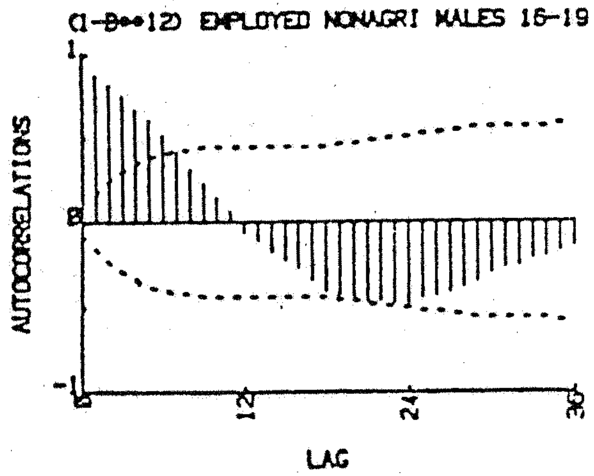


Figure 2.5

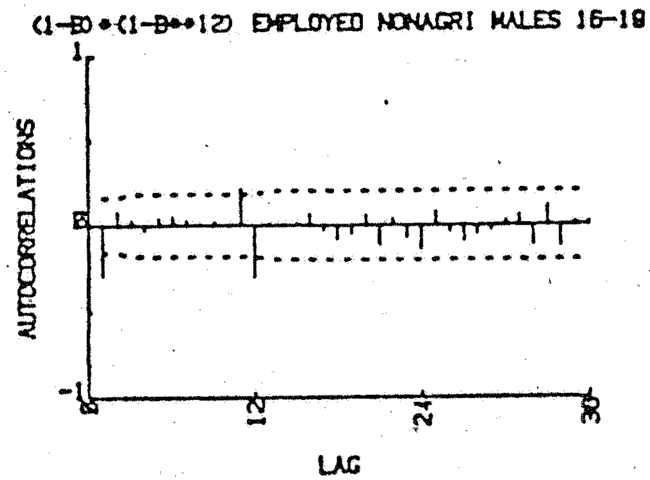


Figure 2.6

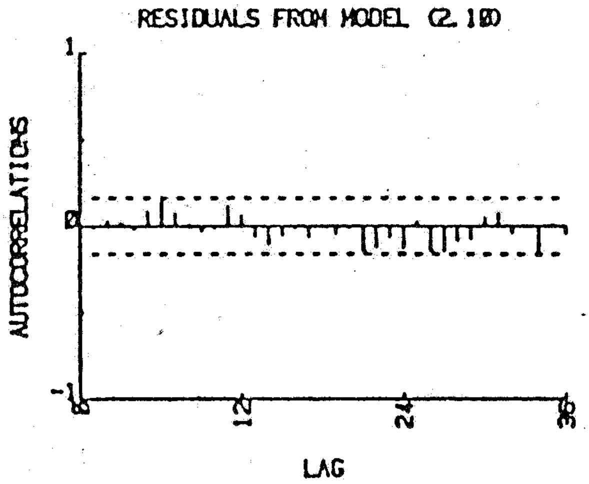
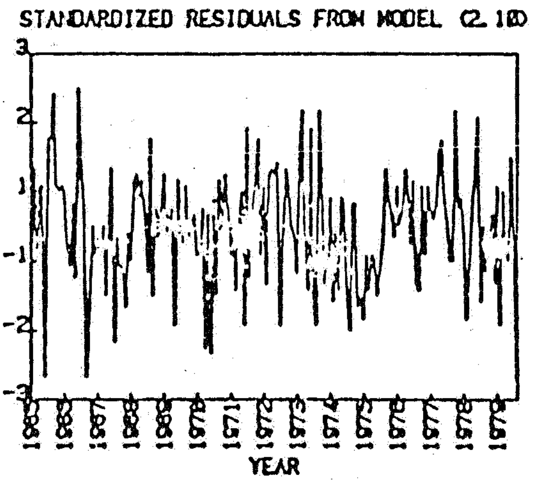


Figure 2.7



LOG WHOLESALE SALES HARDWARE

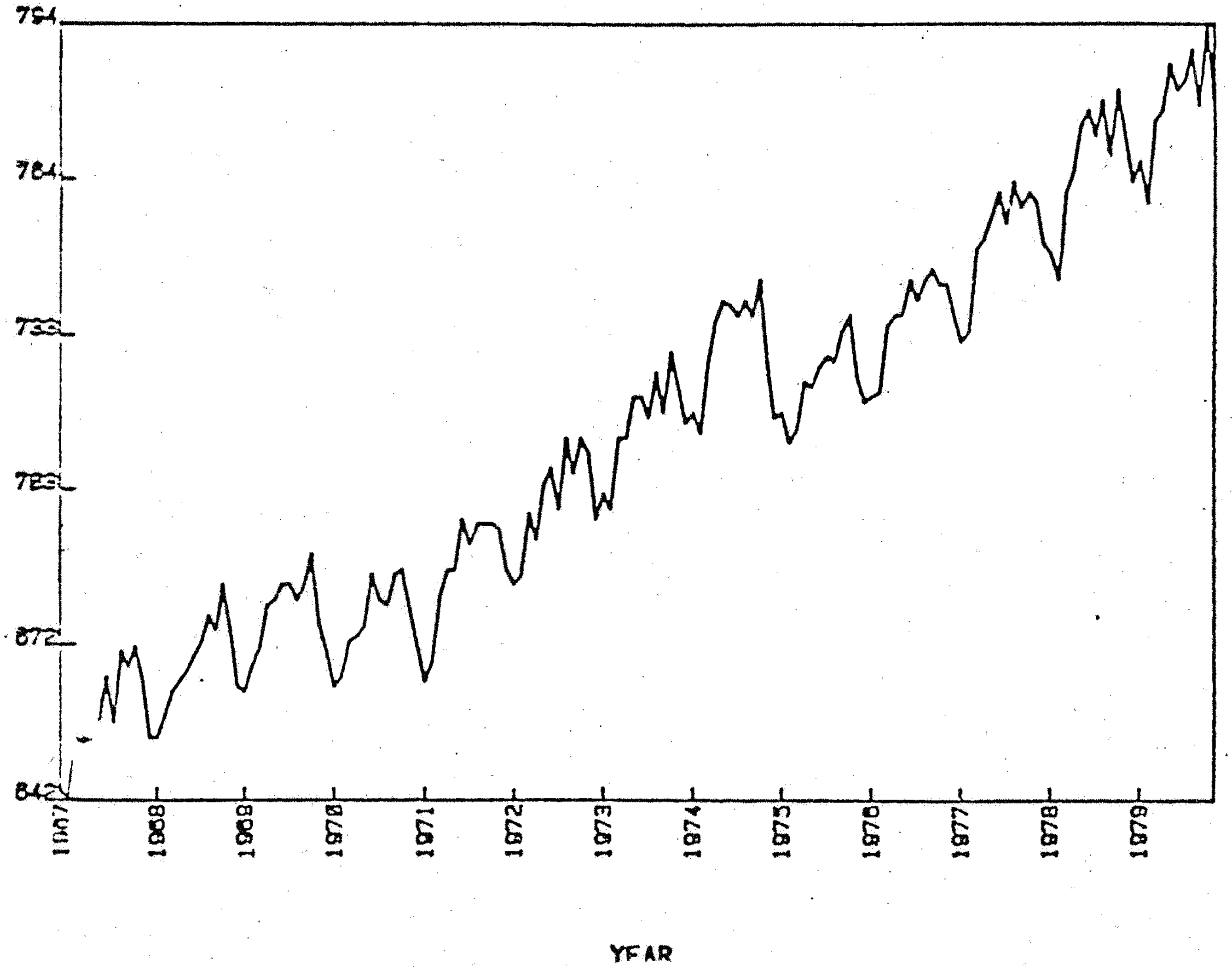


Figure 3.2

LOG WHOLESALE SALES HARDWARE

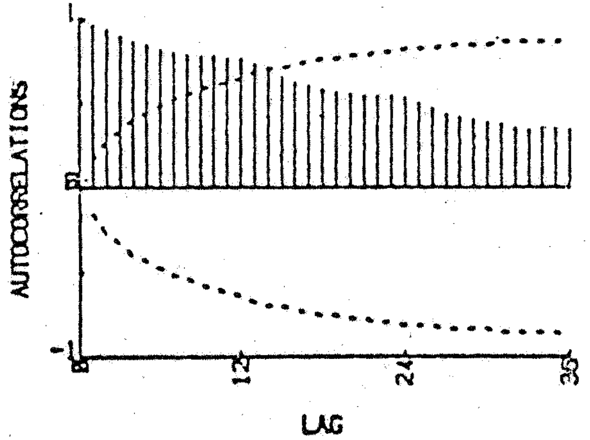


Figure 3.3

(1-B) LOG WHOLESALE SALES HARDWARE

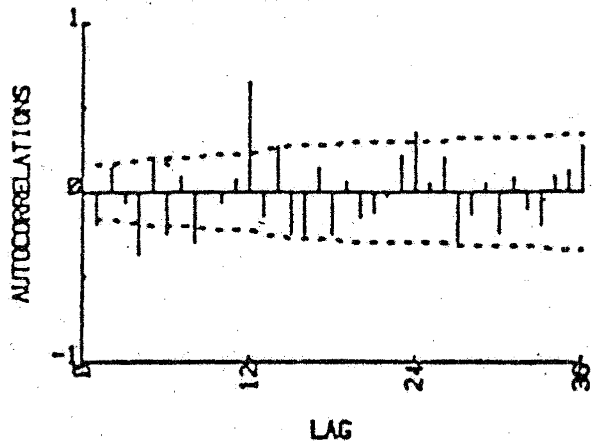


Figure 3.4

(1-B**12) LOG WHOLESALE SALES HARDWARE

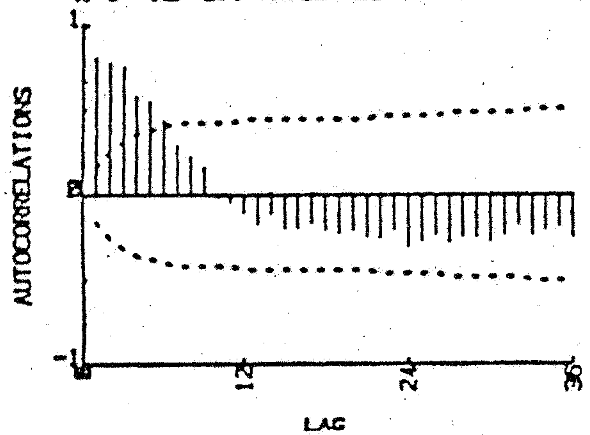
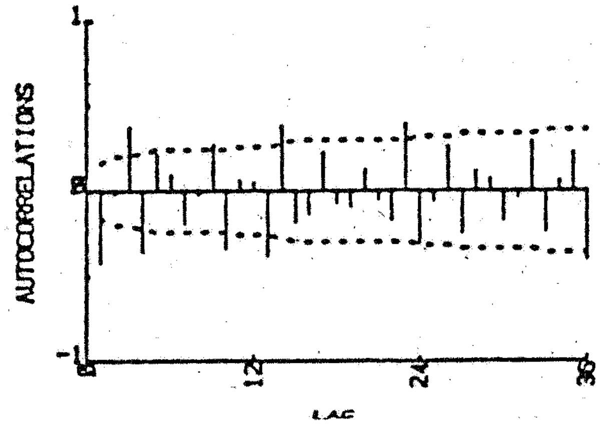
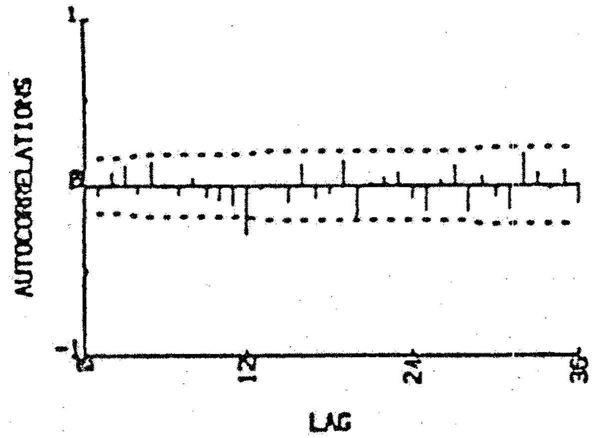


Figure 3.5

(1-B)*(1-B**12) LOG WHOLESALE SALES HARDWARE



RESIDUALS FROM MODEL (3.5)



RESIDUALS FROM MODEL (3.7)

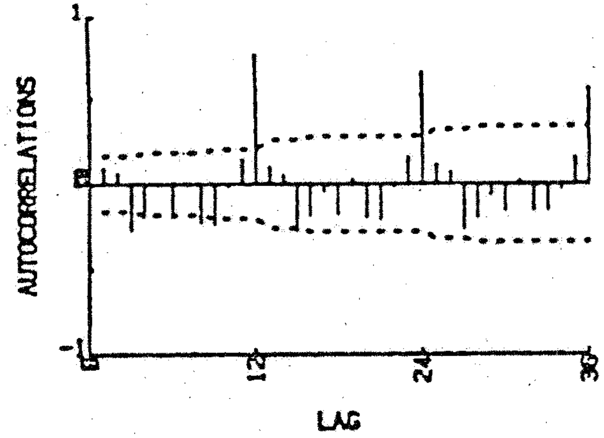


Figure 3.8

RESIDUALS FROM MODEL (3.8)

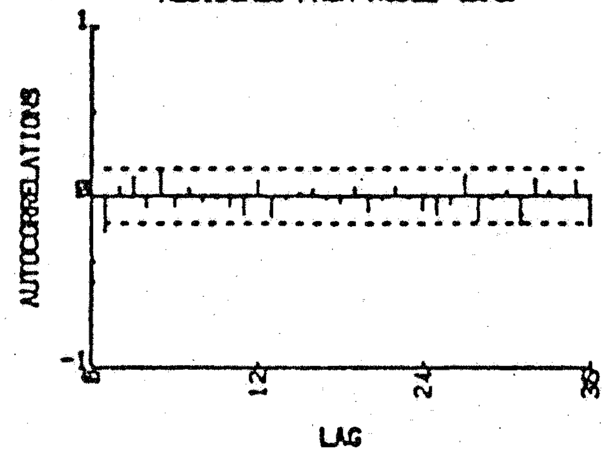


Figure 3.9

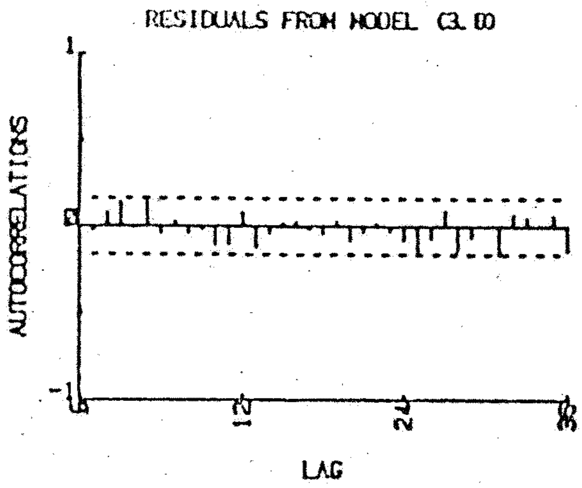
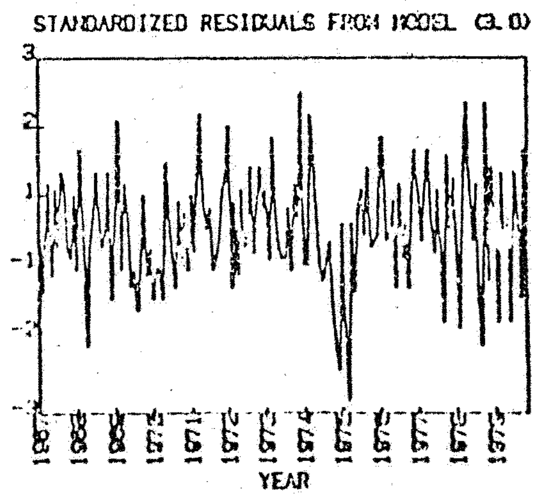


Figure 3.10



WHOLESALE SALES: HARDWARE

Figure 3.11a

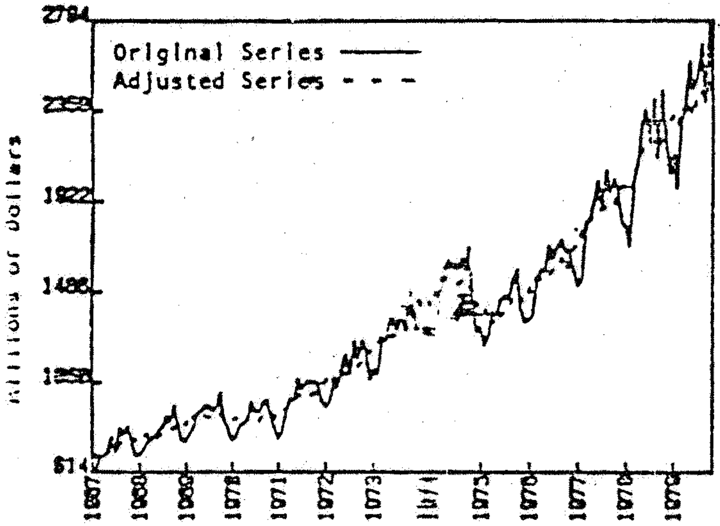


Figure 3.11a Combined Seasonal and TD Factors

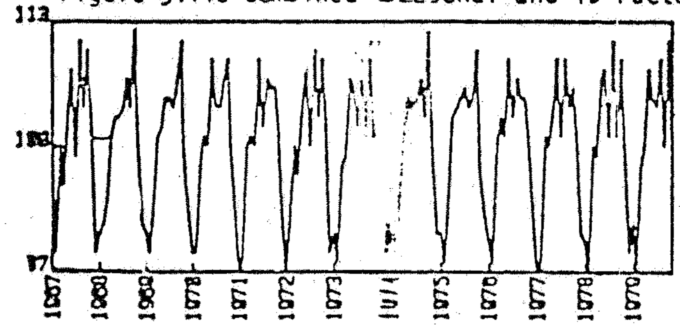


Figure 3.11c Seasonal Factors

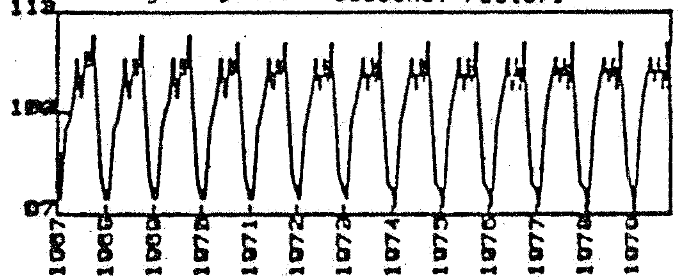


Figure 3.11d Trading Day Factors

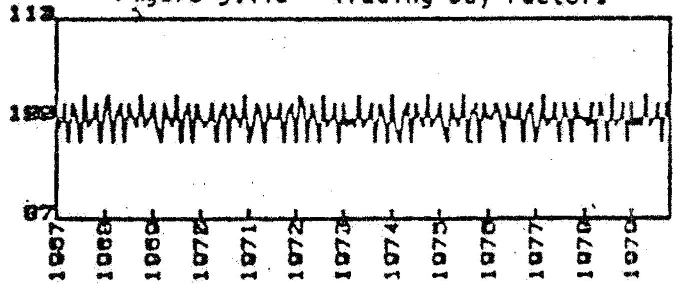


Figure 3.12

LOG RETAIL SALES MEN AND BOYS CLOTHING

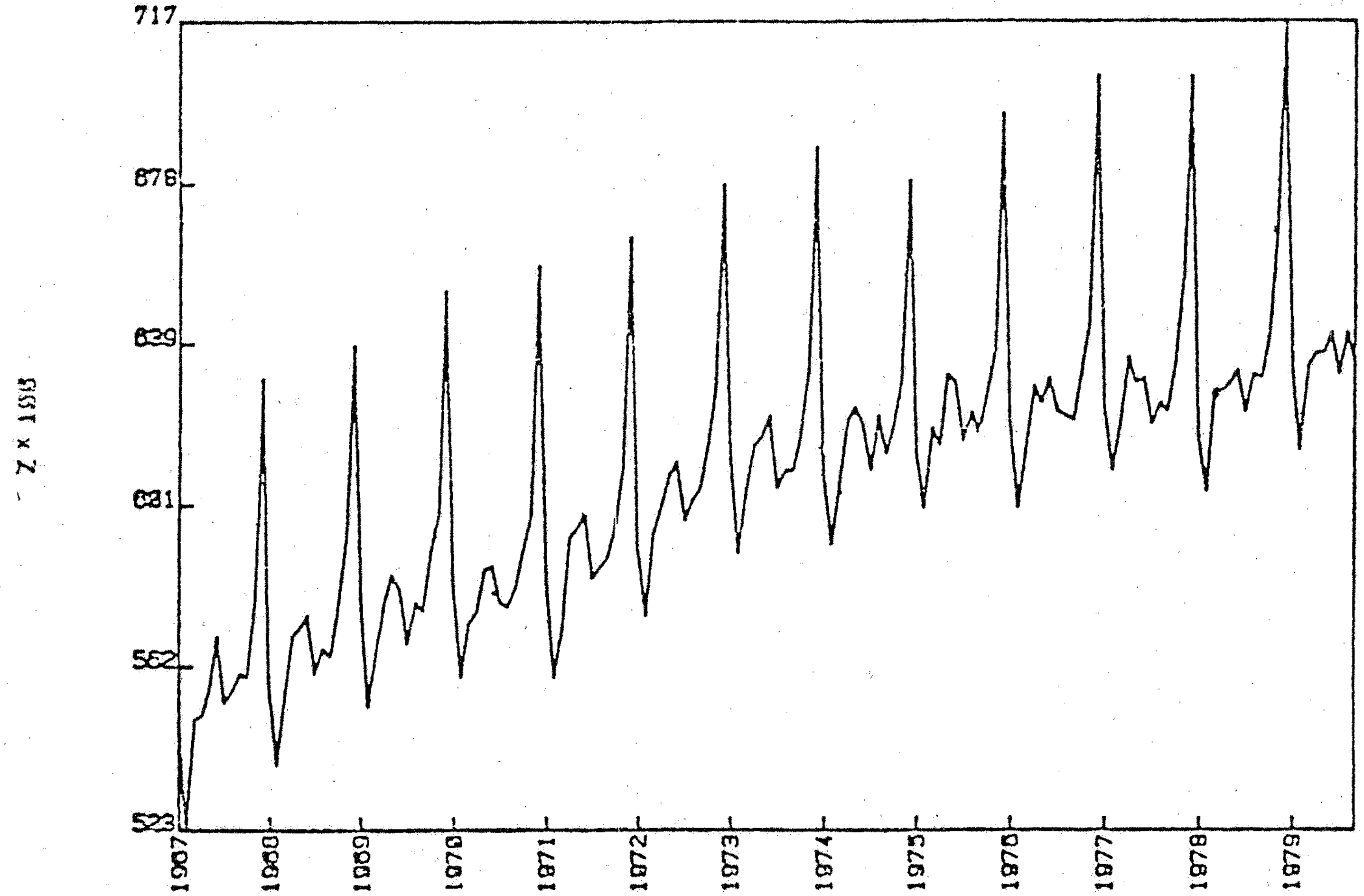


Figure 3.13

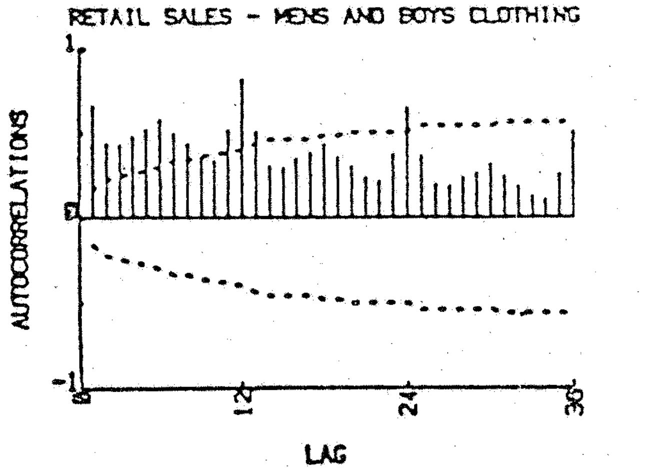


Figure 3.14

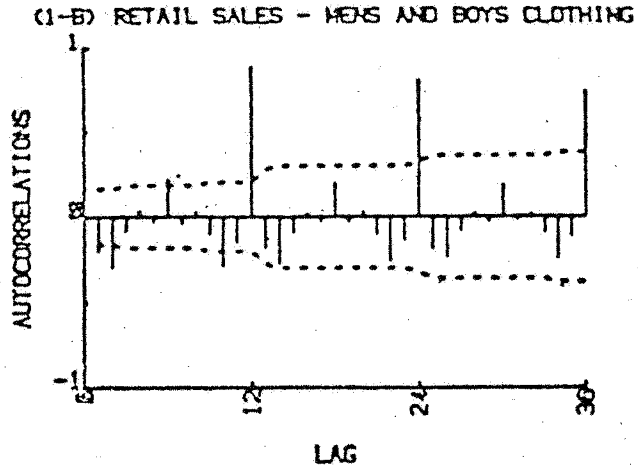


Figure 3.15

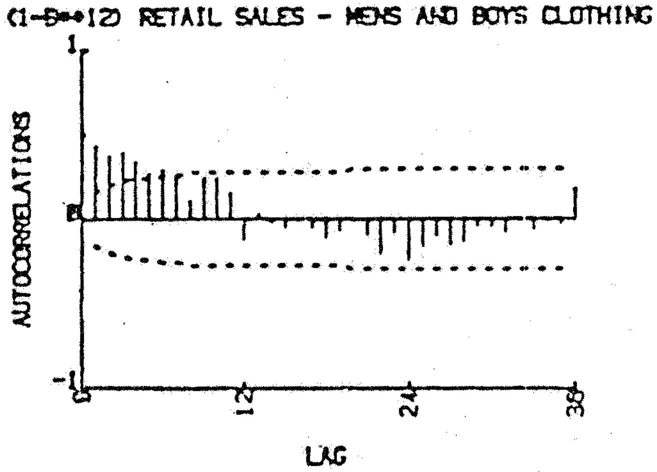


Figure 3.16

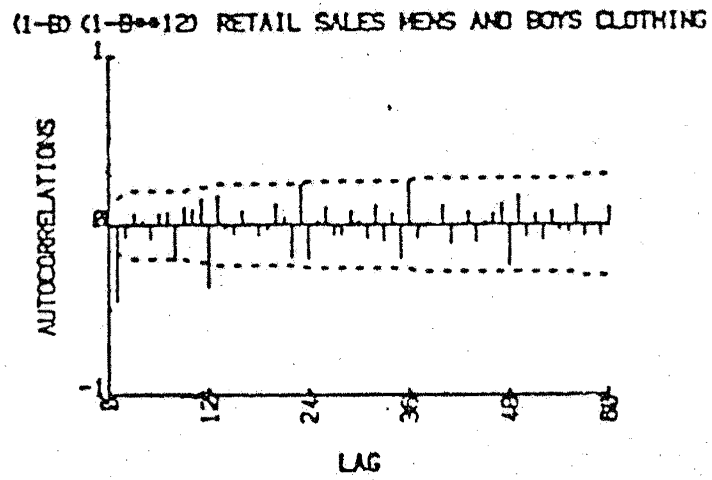


Figure 3.17

RESIDUALS FROM EASTER AND TD REGRESSION

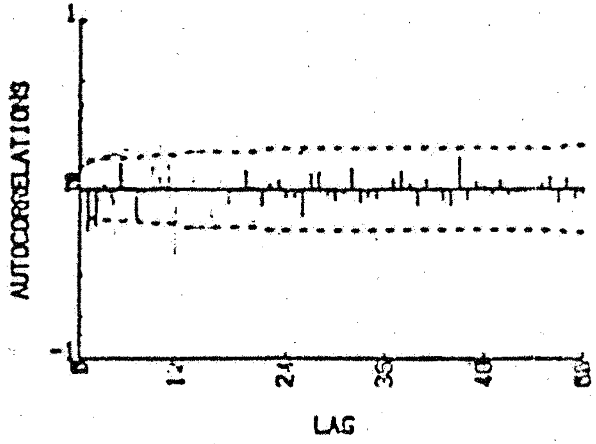


Figure 3.18

RESIDUALS FROM EASTER AND TD REGRESSION

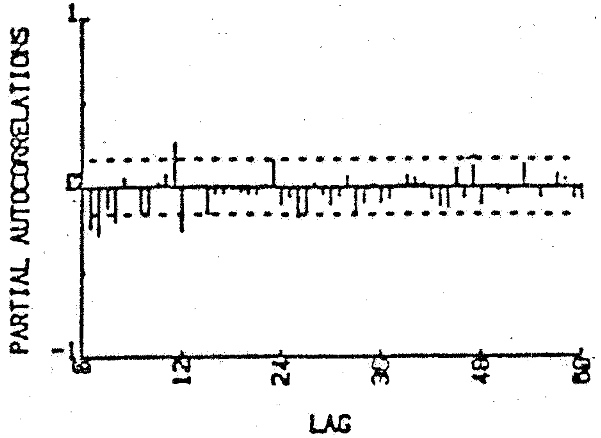


Figure 3.19

RESIDUALS FROM MODEL GL 13D WITH TAU = 0

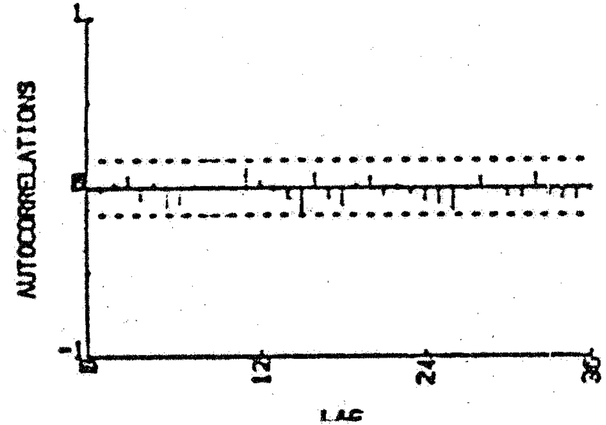
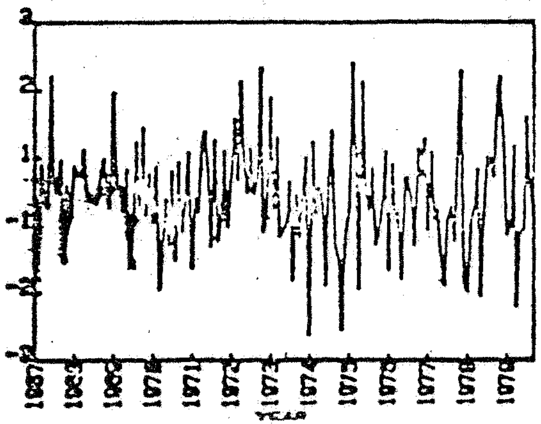


Figure 3.20

STANDARDIZED RESIDUALS MODEL GL 13D WITH TAU = 0



RETAIL SALES — MEN AND BOYS CLOTHING

Figure 3.21a

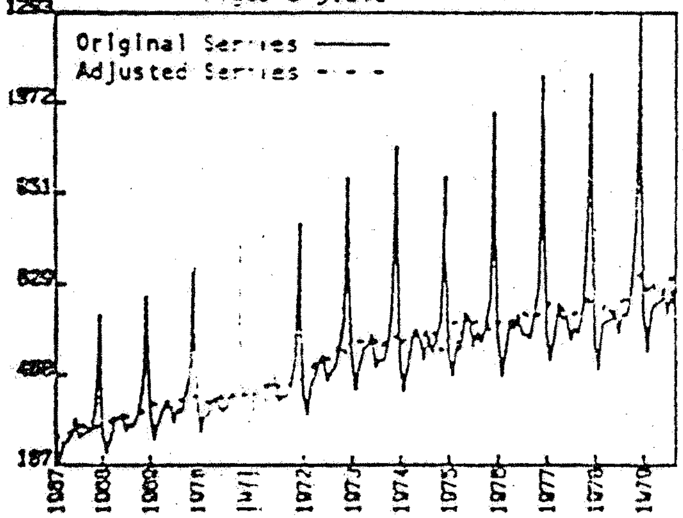


Figure 3.21b Combined Seasonal, TD, and Easter Factors

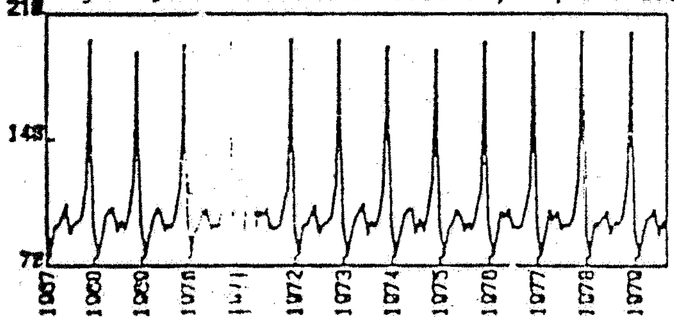


Figure 3.21c Seasonal Factors

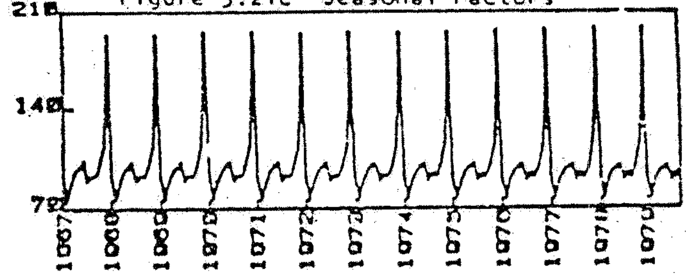


Figure 3.21d Trading Day Factors

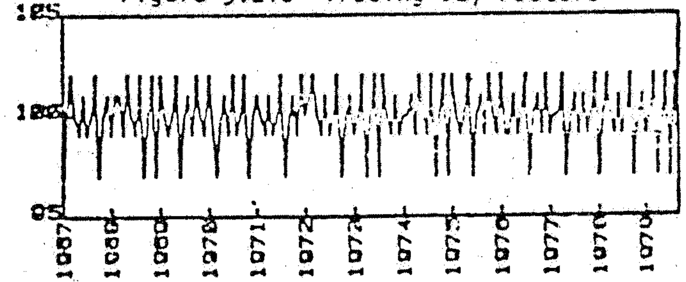


Figure 3.21e Easter Factors

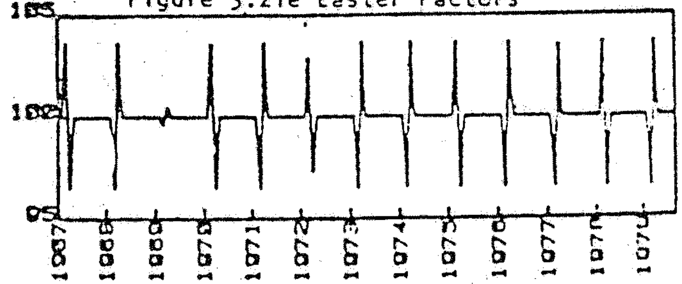


Figure 4.1

LOG RETAIL SALES VARIETY STORES

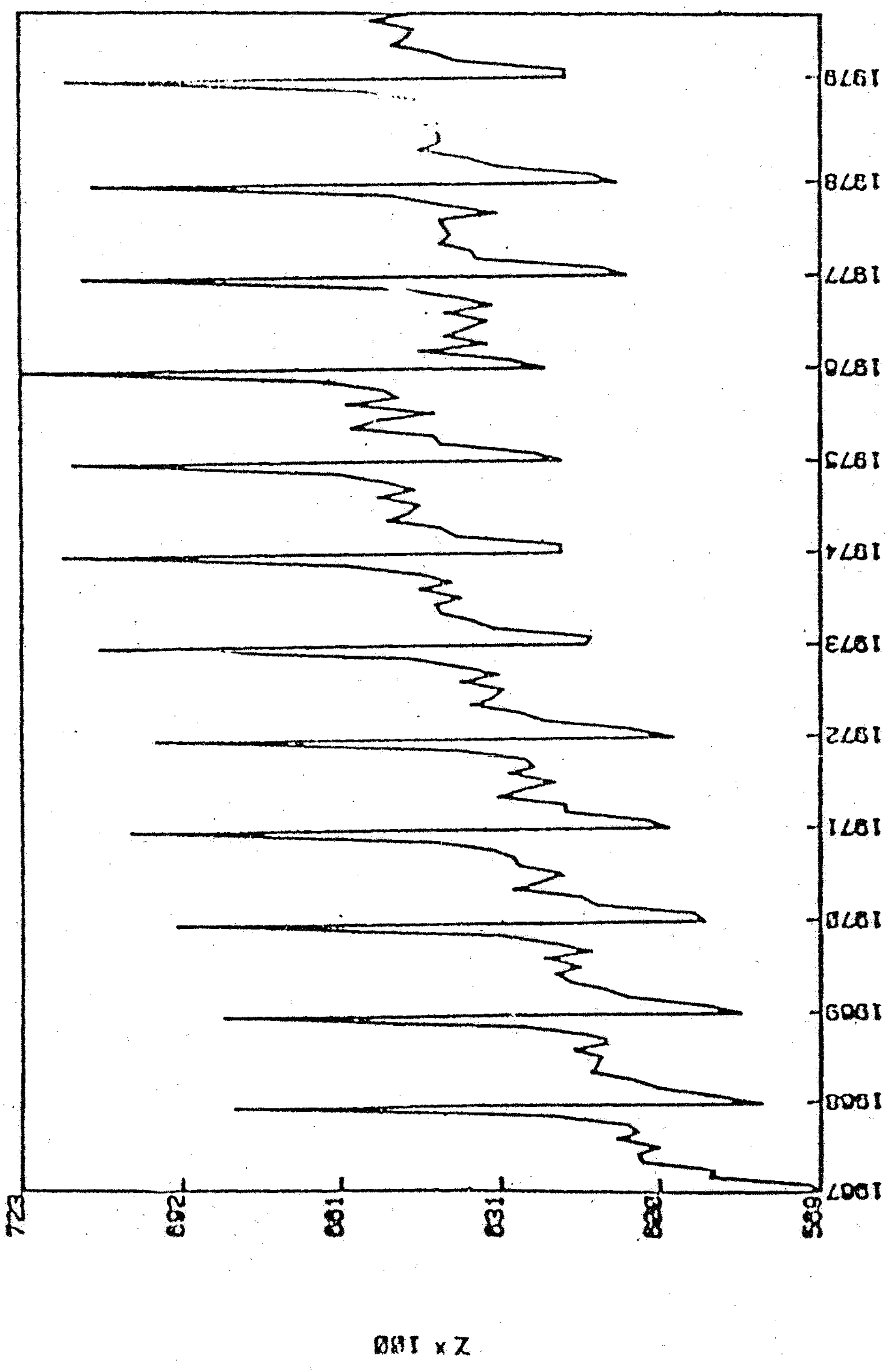
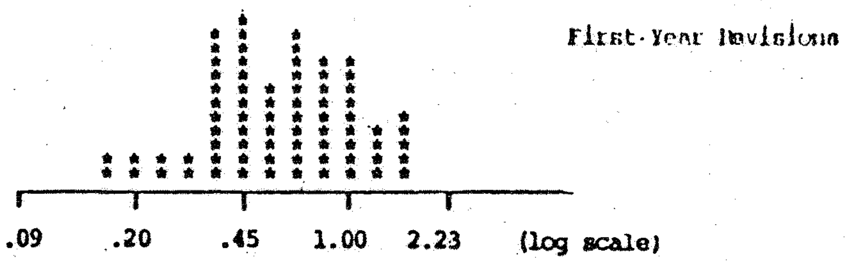
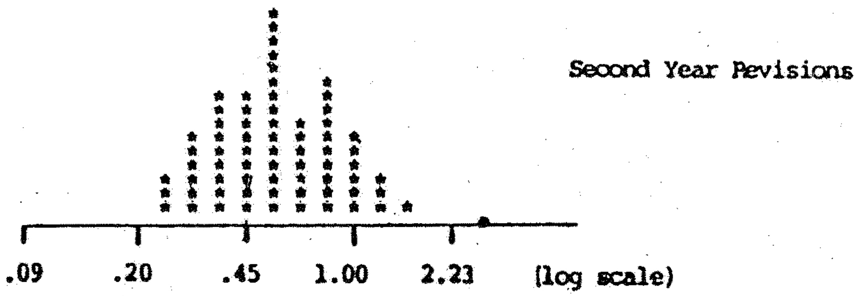


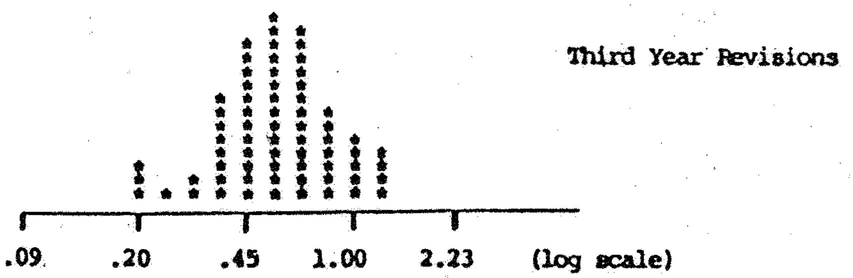
Figure 5.1

Histogram of Ratios (Model Based/X-11)

for Revisions in Level







© Journal of the American Statistical Association
December 1970, Volume 65, Number 332
Theory and Methods Section

Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models

G. E. P. BOX and DAVID A. PIERCE*

Many statistical models, and in particular autoregressive—moving average time series models, can be regarded as means of transforming the data to white noise, that is, to an uncorrelated sequence of errors. If the parameters are known exactly, this random sequence can be computed directly from the observations; when this calculation is made with estimates substituted for the true parameter values, the resulting sequence is referred to as the "residuals," which can be regarded as estimates of the errors.

If the appropriate model has been chosen, there will be zero autocorrelation in the errors. In checking adequacy of fit it is therefore logical to study the sample autocorrelation function of the residuals. For large samples the residuals from a correctly fitted model resemble very closely the true errors of the process; however, care is needed in interpreting the serial correlations of the residuals. It is shown here that the residual autocorrelations are to a close approximation representable as a *singular* linear transformation of the autocorrelations of the errors so that they possess a singular normal distribution. Failing to allow for this results in a tendency to overlook evidence of lack of fit. Tests of fit and diagnostic checks are devised which take these facts into account.

1. INTRODUCTION

An approach to the modeling of stationary and non-stationary time series such as commonly occur in economic situations and control problems is discussed by Box and Jenkins [4, 5], building on the earlier work of several authors beginning with Yule [19] and Wold [17], and involves iterative use of the three-stage process of identification, estimation, and diagnostic checking. Given a discrete time series $z_t, z_{t-1}, z_{t-2}, \cdots$ and using B for the backward shift operator such that $Bz_t = z_{t-1}$, the general autoregressive—integrated moving average (ARIMA) model of order (p, d, q) discussed in [4, 5] may be written

$$\phi(B)\nabla^d z_t = \theta(B)a_t \tag{1.1}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots \theta_q B^q$, $\{a_t\}$ is a sequence of independent normal deviates with common variance σ_a^2, to be referred to as "white noise," and where the roots of $\phi(B) = 0$ and $\theta(B) = 0$ lie outside the unit circle. In other words, if $w_t = \nabla^d z_t = (1-B)^d z_t$ is the dth difference of the series z_t, then w_t is the stationary, invertible, mixed autoregressive (AR)—moving average (MA) process given by

$$w_t = \sum_{i=1}^{p} \phi_i w_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j} + a_t,$$

and permitting $d > 0$ allows the original series to be (homogeneously) nonsta-

* G. E. P. Box is professor of statistics, University of Wisconsin. David A. Pierce is on leave from the Department of Statistics, University of Missouri, Columbia, as statistician, Research Department, Federal Reserve Bank of Cleveland. This work was supported jointly by the Air Force Office of Scientific Research under Grant AFOSR-69-1803 and by the U. S. Army Research Office under Grant DA-ARO-D-31-124-G917.

tionary. In some instances the model (1.1) will be appropriate after a suitable transformation is made on z; in others z may represent the noise structure after allowing for some systematic model.

This general class of models is too rich to allow immediate fitting to a particular sample series $\{z_t\} = z_1, z_2, \cdots, z_n$, and the following strategy is therefore employed:

1. A process of identification is used to find a smaller subclass of models worth considering to represent the stochastic process.
2. A model in this subclass is fitted by efficient statistical methods.
3. An examination of the adequacy of the fit is made.

The object of the third or diagnostic checking stage is not merely to determine whether there is evidence of lack of fit but also to suggest ways in which the model may be modified when this is necessary. Two basic methods for doing this are suggested:

Overfitting. The model may be deliberately overparameterized in a way it is feared may be needed and in a manner such that the entertained model is obtained by setting certain parameters in the more general model at fixed values, usually zero. One can then check the adequacy of the original model by fitting the more general model and considering whether or not the additional parameters could reasonably take on the specified values appropriate to the simpler model.

Diagnostic checks applied to the residuals. The method of overfitting is most useful where the nature of the alternative feared model is known. Unfortunately, this information may not always be available, and less powerful but more general techniques are needed to indicate the way in which a particular model might be wrong. It is natural to consider the stochastic properties of the residuals $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_n)'$ calculated from the sample series using the model (1.1) with estimates $\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p; \hat{\theta}_1, \hat{\theta}_2, \cdots \hat{\theta}_q$ substituted for the parameters. In particular their autocorrelation function

$$\hat{r}_k = \sum \hat{a}_t\hat{a}_{t-k} / \sum \hat{a}_t^2 \tag{1.2}$$

may be studied.

Now if the model were appropriate and the a's for the particular sample series were calculated using the *true* parameter values, then these a's would be uncorrelated random deviates, and their first m sample autocorrelations $r = (r_1, r_2, \cdots, r_m)'$, where m is small relative to n and

$$r_k = \frac{\sum a_t a_{t-k}}{\sum a_t^2}, \tag{1.3}$$

would for moderate or large n possess a multivariate normal distribution [1]. Also it can readily be shown that the $\{r_k\}$ are uncorrelated with variances

$$V(r_k) = \frac{n-k}{n(n+2)} \approx 1/n, \tag{1.4}$$

from which it follows in particular that the statistic $n(n+2)\sum_{k=1}^{m}(n-k)^{-1}r_k^2$ would for large n be distributed as χ^2 with m degrees of freedom; or as a further approximation,

$$n \sum_{k=1}^{m} r_k^2 \sim \chi_m^2. \tag{1.5}$$

It is tempting to suppose that these same properties might to a sufficient approximation be enjoyed by the $\hat{r}$'s from the *fitted* model; and diagnostic checks based on this supposition were suggested by Box and Jenkins [4] and Box, Jenkins, and Bacon [6]. If this assumption were warranted, approximate standard errors of $1/\sqrt{n}$ [or more accurate standard errors of $\sqrt{n-k}/n(n+2)$] could be attached to the $\hat{r}$'s and a quality-control-chart type of approach used, with particular attention being paid to the $\hat{r}$'s of low order for the indication of possible model inadequacies. Also it might be supposed that Equation (1.5) with $\hat{r}$'s replacing r's would still be approximately valid, so that large values of this statistic would place the model under suspicion.

It was pointed out by Durbin [10], however, that this approximation is invalid when applied to the residual autocorrelations from a fitted autoregressive model. For example, he showed that $\hat{r}_1$ calculated from the residuals of a first order autoregressive process could have a much smaller variance than r_1 for white noise.

The present paper therefore considers in some detail the properties of the $\hat{r}$'s and in particular their covariance matrix, both for AR processes (Sections 2 and 3) and for MA and ARIMA processes (Section 5). This is done with the intention of obtaining a suitable modification to the above diagnostic checking procedures (Sections 4 and 5.3)

The problem of testing fit in time series models has been considered previously by several authors. Quenouille [14][1] developed a large-sample procedure for AR processes based on their sample partial autocorrelations, which possesses the same degree of accuracy as the present one.[2] Quenouille's test was subsequently extended [3, 15, 18] to cover MA and mixed models. Whittle [16] proposed tests based on the likelihood ratio and resembling the overfitting method above. The present procedure (a) is a unified method equally applicable to AR, MA, and general ARIMA models, (b) is motivated by the intuitive idea that the residuals from a correct fit should resemble the true errors of the process, and (c) can be used to suggest particular modifications in the model when lack of fit is found [5].

2. DISTRIBUTION OF RESIDUAL AUTOCORRELATIONS FOR THE AUTOREGRESSIVE PROCESS

In this section we obtain the joint large-sample distribution of the residual autocorrelations $\hat{\mathbf{r}}=(\hat{r}_1, \cdots, \hat{r}_m)'$ where $\hat{r}_k$ is given by (1.2), for an autoregressive process. This is done by first setting forth some general properties of AR processes, using these to obtain a set of linear constraints (2.9) satisfied by the $\{\hat{r}_k\}$, and then approximating $\hat{r}_k$ by a first order Taylor expansion (2.22) about the white noise autocorrelation r_k. Finally, these results are combined in matrix form to establish a linear relationship (2.27) between $\hat{\mathbf{r}}$ and $\mathbf{r}$ analogous to that between the residuals and true errors in a standard regression model, from which the distribution (2.29) of $\hat{\mathbf{r}}$ readily follows. Subsections 2.5–2.7 then discuss examples and applications of this distribution.

[1] See also [11].

[2] The authors are grateful to a referee for this observation.

2.1 The Autoregressive Process

The general AR process of order p,

$$\phi(B)y_t = a_t, \tag{2.1}$$

where B, $\phi(B)$, and $\{a_t\}$ are as in (1.1), can also be expressed as a moving average of infinite order by writing $\psi(B)=\phi^{-1}(B)=(1+\psi_1B+\psi_2B^2+\cdots)$ to obtain

$$y_t = \psi(B)a_t = \sum_{j=0}^{\infty}\psi_j a_{t-j}, \tag{2.2}$$

where $\psi_0=1$. By equating coefficients in the relation $\psi(B)\cdot\phi(B)=1$, it is seen that the ψ's and ϕ's satisfy the relation

$$\psi_\nu = \begin{cases}\phi_1\psi_{\nu-1}+\cdots+\phi_{\nu-1}\psi_1+\phi_\nu, & \nu\le p\\ \phi_1\psi_{\nu-1}+\cdots+\phi_p\psi_{\nu-p}, & \nu\ge p.\end{cases} \tag{2.3}$$

Therefore by setting $\psi_\nu=0$ for $\nu<0$, we have

$$\psi_0 = 1;\quad \phi(B)\psi_\nu = 0, \qquad \nu\ne 0. \tag{2.4}$$

Suppose then we have a series $\{y_t\}$ generated by the model (2.1) or (2.2), where in general $y_t=\nabla^d z_t$ can be the dth difference ($d=0, 1, 2, \cdots$) of the actual observations. Then for given values $\dot{\phi}=(\dot\phi_1, \cdots, \dot\phi_p)'$ of the parameters we can define

$$\dot a_t = a_t(\dot\phi) = y_t - \dot\phi_1 y_{t-1} - \cdots - \dot\phi_p y_{t-p} = \dot\phi(B)y_t \tag{2.5}$$

and the corresponding autocorrelation

$$\dot r_k = r_k(\dot\phi) = \frac{\sum \dot a_t \dot a_{t-k}}{\sum \dot a_t^2}. \tag{2.6}$$

Thus, in particular,

1. $a_t(\phi)=a_t$ as in (2.1), (2.2);
2. $a_t(\hat\phi)=\hat a_t$ are the residuals when (2.1) is fitted and least squares estimated $\hat\phi$ obtained; and
3. $r_k(\hat\phi)$ and $r_k(\phi)$ are respectively the residual and white noise autocorrelations (1.2) and (1.3).

2.2 Linear Constraints on the $\hat r$'s

It is known that the residuals $\{\hat a_t\}$ above satisfy the orthogonality conditions

$$\sum_{t=p+1}^{n}\hat a_t y_{t-j} = 0, \qquad 1\le j\le p. \tag{2.7}$$

Therefore if we let

$$\hat\psi(B) = \hat\phi^{-1}(B) = (1-\hat\phi_1 B - \cdots - \hat\phi_p B^p)^{-1}, \tag{2.8}$$

then $y_t=\hat\psi(B)\hat a_t$, and from (2.7) we have

$$\begin{aligned}0 &= \textstyle\sum_t\sum_k \hat\psi_k \hat a_t \hat a_{t-k-j}\\ &= \textstyle\sum_k \hat\psi_k \hat r_{k+j}\\ &= \textstyle\sum \psi_k \hat r_{k+j} + O_p(1/n)\end{aligned} \tag{2.9}$$

where the symbol introduced in (2.9) denotes "order in probability" as defined in [13].

In leading up to (2.9) we have presumably summed an infinite number of autocorrelations from a finite series. However since $\{y_t\}$ is stationary we have $\psi_k \to 0$ as k becomes large; and unless ϕ is extremely close to the boundary of the stationarity region, this dying off of ψ_k is fast so that the summation can generally be stopped at a value of k much less than n. More precisely, we are assuming that n is larger than a fixed number N and for such n there exists a sequence of numbers m_n such that

(a) all ψ_j where $j \geq m_n - p$ are of order $1/\sqrt{n}$ or smaller, and
(b) the ratio m_n/n is itself of order $1/\sqrt{n}$.

Then in (2.9) and in all following discussion the error in stopping the summations at $k=m$ (we write m for m_n in the sequel) can to the present degree of approximation be ignored; and (b) also ensures that "end effects" (such as there being only $n-k$ terms summed in the numerator of $\hat{r}_k$ compared with n terms in the denominator) can also be neglected.

2.3 Linear Expansion of $\hat{r}_k$ about r_k

The root mean square error of $\hat{\phi}_j$, $1 \leq j \leq p$, defined by $\sqrt{E(\phi_j - \hat{\phi}_j)^2}$, is of order $1/\sqrt{n}$, and we can therefore approximate $\hat{r}_k$ by a first order Taylor expansion about $\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}$ (evaluating the derivatives, however, at $\hat{\boldsymbol{\phi}}$ rather than $\boldsymbol{\phi}$ in order to obtain the simplification (2.12) below). Thus

$$\hat{r}_k = r_k + \textstyle\sum_{j=1}^{p} (\phi_j - \hat{\phi}_j)\hat{\delta}_{jk} + O_p(1/n), \tag{2.10}$$

where

$$\hat{\delta}_{jk} = -\left.\frac{\partial \dot{r}_k}{\partial \dot{\phi}_j}\right|_{\dot{\phi} = \hat{\phi}}. \tag{2.11}$$

Now

$$\frac{\partial}{\partial \dot{\phi}_j} [\textstyle\sum \dot{a}_t^2] = 0 \quad \text{at } \dot{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}}, \tag{2.12}$$

so that

$$\hat{\delta}_{jk} = -[\textstyle\sum \hat{a}_t^2]^{-1} \left.\frac{\partial \dot{c}_k}{\partial \dot{\phi}_j}\right|_{\dot{\phi} = \hat{\phi}} \tag{2.13}$$

where

$$\begin{aligned} \dot{c}_k &= \textstyle\sum \dot{a}_t \dot{a}_{t-k} = \sum [\dot{\phi}(B) y_t][\dot{\phi}(B) y_{t-k}] \\ &= \textstyle\sum_t \sum_{i=0}^{p} \sum_{j=0}^{p} \dot{\phi}_i \dot{\phi}_j y_{t-i} y_{t-k-j}, \end{aligned} \tag{2.14}$$

where in (2.14) and below, $\phi_0 = \dot{\phi}_0 = -1$. From (2.13) and (2.14) it follows that

$$\begin{aligned} \hat{\delta}_{jk} &= -\frac{\sum y_t^2}{\sum \hat{a}_t^2} \textstyle\sum_{i=0}^{p} \hat{\phi}_i [r^{(y)}_{k-i+j} + r^{(y)}_{k+i-j}] \\ &= -\frac{\sum_{i=0}^{p} \hat{\phi}_i [r^{(y)}_{k-i+j} + r^{(y)}_{k+i-j}]}{\sum_{i=0}^{p} \sum_{j=0}^{p} \hat{\phi}_i \hat{\phi}_j r^{(y)}_{i-j}}, \end{aligned} \tag{2.15}$$

where

$$r_\nu^{(y)} = \frac{\sum y_t y_{t-\nu}}{\sum y_t^2}.$$

Let us approximate $\hat{\delta}_{jk}$ by replacing $\hat{\phi}$'s and $r^{(y)}$'s in (2.15) by ϕ's and ρ's (the theoretical parameters and autocorrelations of the autoregressive process $\{y_t\}$) and denote the result by δ_{jk}. That is,

$$\delta_{jk} = \frac{\sum_{i=0}^{p} \phi_i[\rho_{k-i+j} + \rho_{k+i-j}]}{-\sum_{i=0}^{p}\sum_{j=0}^{p} \phi_i\phi_j\rho_{i-j}}. \tag{2.16}$$

Now from Bartlett's formula [2, Equation (7)] we have

$$r_k^{(y)} = \rho_k + O_p(1/\sqrt{n}), \tag{2.17}$$

and as in the discussion preceding (2.10), $\hat{\phi}_j = \phi_j + O_p(1/\sqrt{n})$;
thus

$$\hat{\delta}_{jk} = \delta_{jk} + O_p(1/\sqrt{n}), \tag{2.18}$$

so that equation (2.10) holds when $\hat{\delta}_{jk}$ is replaced by δ_{jk}.

By making use of the recursive relation which is satisfied by the autocorrelations of an autoregressive process, namely

$$\rho_\nu - \phi_1\rho_{\nu-1} - \cdots - \phi_p\rho_{\nu-p} = \phi(B)\rho_\nu = 0, \qquad \nu \geq 1, \tag{2.19}$$

expression (2.16) can be simplified to yield

$$\delta_{jk} = \frac{\sum_{i=0}^{p} \phi_i\rho_{k-j+i}}{\sum_{i=0}^{p} \phi_i\rho_i}. \tag{2.20}$$

Thus δ_{jk} depends only on $(k-j)$, and we therefore write $\delta_{k-j} = \delta_{jk}$. Then it is straightforward to show that

(a) $\delta_0 = 1$

(b) $\delta_\nu = 0, \quad \nu < 0,$ and thus

(c) $\phi(B)\delta_\nu = \dfrac{\sum_{i=0}^{p} \phi_i[\phi(B)\rho_{\nu+i}]}{\sum_{i=0}^{p} \phi_i\rho_i} = 0, \qquad \nu \geq 1.$

Comparing (a), (b), and (c) with the corresponding results (2.4) for ψ_ν, we therefore have $\delta_\nu = \psi_\nu$, that is

$$\delta_{jk} = \psi_{k-j}, \tag{2.21}$$

whence, for $k = 1, 2, \cdots, m$,

$$\hat{r}_k = r_k + \sum_{j=1}^{p} (\phi_j - \hat{\phi}_j)\psi_{k-j} + O_p(1/n). \tag{2.22}$$

2.4 Representation of $\hat{r}$ as a Linear Transformation of r

We can now establish a relationship between the residual autocorrelations $\hat{r}$ and the white noise autocorrelations r. Let

$$X = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \psi_1 & 1 & & \vdots \\ \psi_2 & \psi_1 & \ddots & 0 \\ \vdots & \vdots & & 1 \\ \vdots & \vdots & & \vdots \\ \psi_{m-1} & \psi_{m-2} & \cdots & \psi_{m-p} \end{bmatrix} \tag{2.23}$$
$$= [x_1 \mid x_2 \mid \cdots \mid x_p].$$

Then to $O_p(1/n)$ we can write (2.22) in matrix form as

$$\hat{r} = r + X(\phi - \hat{\phi}), \tag{2.24}$$

where from (2.9)

$$\hat{r}'X = 0. \tag{2.25}$$

If we now multiply (2.24) on both sides by

$$Q = X(X'X)^{-1}X', \tag{2.26}$$

then using (2.25) we obtain

$$\hat{r} = (I - Q)r. \tag{2.27}$$

It is known [1] that r is very nearly normal for n moderately large. The vector of residual autocorrelations is thus approximately a linear transformation of a multi-normal variable and is therefore itself normally distributed. Specifically,

$$r \sim N(0, (1/n)I), \tag{2.28}$$

and hence

$$\hat{r} \sim N(0, (1/n)[I - Q]). \tag{2.29}$$

Note that the matrix $I-Q$ is idempotent of rank $m-p$, so that the distribution of $\hat{r}$ has a p-dimensional singularity.

2.5 Further Consideration of the Covariance Structure of the $\hat{r}$'s

It is illuminating to examine in greater detail the covariance matrix of $\hat{r}$, or equivalently the matrix Q. The latter matrix is idempotent of rank p, and its non-null latent vectors are the columns of X. Also,

$$X'X = \begin{bmatrix} \sum \psi_j^2 & \sum \psi_j\psi_{j-1} & \cdots & \sum \psi_j\psi_{j-p+1} \\ \sum \psi_j\psi_{j-1} & \sum \psi_j^2 & \cdots & \sum \psi_j\psi_{j-p+2} \\ \vdots & \vdots & & \vdots \\ \sum \psi_j\psi_{j-p+1} & \sum \psi_j\psi_{j-p+2} & \cdots & \sum \psi_j^2 \end{bmatrix}$$
$$\doteq \frac{\sigma_y^2}{\sigma_a^2}\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix} \tag{2.30}$$

which when multiplied by σ_a^2 is the autocovariance matrix of the process itself. Let c^{ij} be the (ij)th element of $(X'X)^{-1}$ (given explicitly in [9]), and similarly q_{ij} for Q. If $\xi_j' = (\psi_{j-1}, \cdots, \psi_{j-p})$ denotes the jth row of X, then

$$\begin{aligned} q_{ij} &= \xi_i'(X'X)^{-1}\xi_j \\ &= \sum_{k=1}^{p} \sum_{l=1}^{p} \psi_{i-k} c^{kl} \psi_{j-l} \qquad (2.31) \\ &= (-n) \operatorname{cov}[\hat{r}_i, \hat{r}_j] \quad \text{if } i \neq j. \end{aligned}$$

Since the elements of each column of X satisfy the recursive relation (2.4), we have $\phi(B)\xi_j = 0$, and hence

$$\phi(B) q_{ij} = 0, \qquad (2.32)$$

where in (2.32) B can operate either on i or on j. This establishes an interesting recursive structure in the residual autocorrelation covariance matrix $(1/n) \cdot (I-Q)$ and provides an important clue as to how rapidly the covariances die out and the variances approach 1. Also, because of this property the entire covariance matrix is determined by specifying the elements

$$\begin{matrix} q_{11} & q_{12} & \cdots & q_{1p} \\ & q_{22} & \cdots & q_{2p} \\ & & \ddots & \vdots \\ & & & q_{pp} \end{matrix} \qquad (2.33)$$

of Q, which are readily obtained by inverting the $X'X$ matrix (2.30).

2.6 Covariance Matrix of $\hat{r}$ for first and second order processes

Consider, for example, the first order autoregressive process $y_t = \phi y_{t-1} + a_t$, which in accordance with (2.2) we can write as

$$y_t = (1 - \phi L)^{-1} a_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}. \qquad (2.34)$$

For this process, $\psi_j = \phi^j$ and $(X'X)^{-1} = 1 - \phi^2$. From (2.31) the (ij)th element of Q is therefore $\phi^{i+j-2}(1-\phi^2)$, so that approximately the covariance matrix of the sample residual autocorrelations is

$$\sum \hat{r} = (1/n)(I - Q) = 1/n \begin{bmatrix} \phi^2 & -\phi + \phi^3 & -\phi^2 + \phi^4 & \cdots \\ -\phi + \phi^3 & 1 - \phi^2 + \phi^4 & -\phi^3 + \phi^5 & \cdots \\ -\phi^2 + \phi^4 & -\phi^3 + \phi^5 & 1 - \phi^4 + \phi^6 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \qquad (2.35)$$

For the second order process

$$y_t = (1 - \phi_1 B - \phi_2 B^2)^{-1} a_t = \psi(B) a_t, \qquad (2.36)$$

we have

$$X = \begin{bmatrix} 1 & 0 \\ \psi_1 & 1 \\ \psi_2 & \psi_1 \\ \vdots & \vdots \end{bmatrix}, \qquad X'X = \frac{\sigma_y^2}{\sigma_a^2} \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix},$$

$$(X'X)^{-1} = \frac{\sigma_a^2}{\sigma_y^2(1-\rho_1^2)}\begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix}, \qquad \sigma_y^2 = \frac{(1-\phi_2)\sigma_a^2}{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]}.$$

Thus

$$q_{11} = 1-\phi_2^2, \qquad q_{12} = -\phi_1\phi_2(1+\phi_2), \qquad q_{22} = 1-\phi_2^2-\phi_1^2(1+\phi_2)^2,$$

from which Q and $\sum_{\hat{r}} = 1/n(I-Q)$ may be determined using (2.32). In particular,

$$\left.\begin{aligned} V(\hat{r}_1) &= 1/n\cdot\phi_2^2, \\ V(\hat{r}_2) &= 1/n[\phi_2^2+\phi_1^2(1+\phi_2)^2], \quad \text{and} \\ V(\hat{r}_k) &= 1/n[1-\phi_1 q_{k,k-1}-\phi_2 q_{k,k-2}],\ k \geq 3. \end{aligned}\right\} \tag{2.37}$$

From these examples we can see a general pattern emerging. As in (2.33) the first p variances and corresponding covariances will be heavily dependent on the parameters $\phi_1, \cdots, \phi_p$ and in general can depart sharply from the corresponding values for white noise autocorrelations, whereas for $k \geq p+1$ a "1" is introduced into the expression for variances (as in (2.35) and (2.37)), and the recursion (2.32) ensures that as k increases the $\{\hat{r}_k\}$ behave increasingly like the corresponding $\{r_k\}$ with respect to both their variances and covariances.

2.7 The distribution of $n\sum_1^m \hat{r}_k^2$

We have remarked earlier that if the fitted model is appropriate and the parameters $\boldsymbol{\phi}$ are exactly known, then the calculated a_t's would be uncorrelated normal deviates, their serial correlations $\boldsymbol{r}$ would be approximately $N(\mathbf{0}, (1/n)I)$, and thus $n\sum_1^m r_k^2$ would possess a χ^2 distribution with m degrees of freedom. We now see that if m is taken sufficiently large so that the elements after the mth in the latent vectors of Q are essentially zero, then we should expect that to the order of approximation we are here employing, the statistic

$$n\sum_1^m \hat{r}_k^2, \tag{2.38}$$

obtained when estimates $\hat{\boldsymbol{\phi}}$ are substituted for the true parameters $\boldsymbol{\phi}$ in the model, will still be distributed as χ^2, only now with $m-p$ rather than m degrees of freedom. This result is of considerable practical interest because it suggests that an overall test of the type discussed in [4] can in fact be justified when suitable modifications coming from a more careful analysis are applied. Later we consider in more detail the use of this test, along with procedures on individual $\hat{r}$'s, in diagnostic checking.

3. MONTE CARLO EXPERIMENT

We have made certain approximations in deriving the distribution of the residual autocorrelations, and it is therefore of interest to investigate this distribution empirically through repeated sampling and to compare the results with (2.29). This was done for the first order AR process for $\phi=0$, $\pm.1$, $\pm.3$, $\pm.5$, $\pm.7$, $\pm.9$. For given ϕ, $s=50$ sets of $n=200$ random normal deviates were generated on the computer using a method described in [7], with separate aggregates of deviates obtained for each parameter value. For the jth set a series $\{y_t^{(j)}\}$ was generated using formula (2.34), $\hat{\phi}^{(j)}$ was estimated, $\{\hat{a}_t^{(j)}\}$ determined, and the quantities

$$\hat{r}_k^{(j)} = \frac{\sum \hat{a}_t^{(j)} \hat{a}_{t-k}^{(j)}}{\sum [\hat{a}_t^{(j)}]^2} \tag{3.1}$$

computed for $1 \leq k \leq m = 20$, $1 \leq j \leq s = 50$. This yielded sample variances and covariances

$$C_{kt} = \frac{1}{50} \sum_{j=1}^{50} \hat{r}_k^{(j)} \hat{r}_t^{(j)} \tag{3.2}$$

and sample correlations

$$R_{kt} = C_{kt}/\sqrt{C_{kk}C_{tt}}\,. \tag{3.3}$$

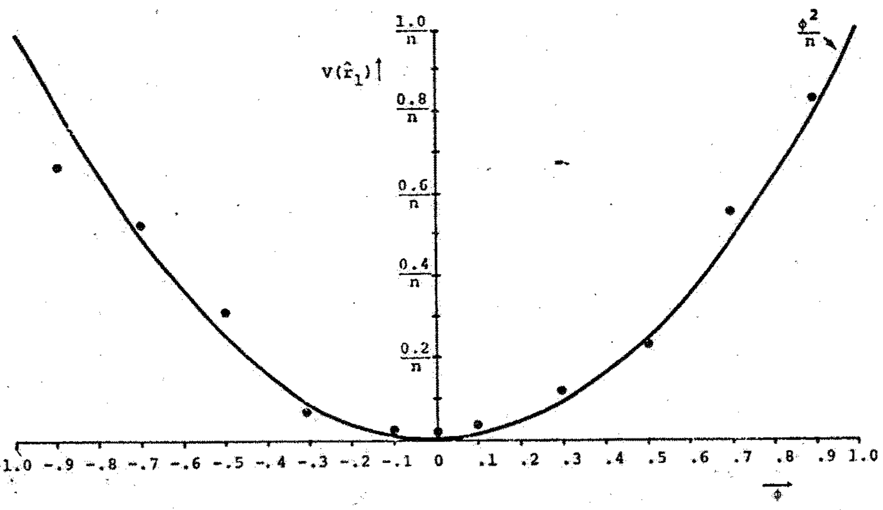
The results of this Monte Carlo sampling are set out in detail in [8] and in general confirm the adequacy of the approximations used. As an example of these calculations, Table 1 compares the empirical variances (3.2) of $\hat{r}_k$ and correlations (3.3) of $(\hat{r}_1, \hat{r}_k)$ with their theoretical counterparts obtained from (2.35). Allowing for the sampling error of the Monte Carlo estimates themselves, there is good agreement between the two sets of quantities, a phenomenon which occurred also for the other values of ϕ considered.

Since the large-sample variance ϕ^2/n of $\hat{r}_1$ departs the most from the common variance of $1/n$ for white noise autocorrelations, an examination of the empirical behavior of this quantity is of particular interest. Thus Figure 1 shows the sample variance of $\hat{r}_1$ for $\phi = 0, \pm.1, \pm.3, \pm.5, \pm.7, \pm.9$ in relation to the parabola $V(\hat{r}_1) = \phi^2/n$, with reasonable agreement between the two. (The coefficient of variation of the sample variance of $\hat{r}_k$ for $\phi \neq 0$ is approximately $\sqrt{2/s} = 1/5$, independent of k and n; at $\phi = 0$, $V(\hat{r}_1) = O(1/n^2)$.

Table 1. *THEORETICAL (AS IN (2.35)) AND EMPIRICAL (FROM MONTE-CARLO SAMPLING) VARIANCES AND CORRELATIONS OF SAMPLE RESIDUAL AUTOCORRELATIONS FROM FIRST-ORDER AR PROCESS WITH $\phi = .5$*

k	*Variance of $\hat{r}_k$ (multiplied by n)*		*Correlation between $\hat{r}_1$ and $\hat{r}_k$*	
	Theoretical	*Empirical*	*Theoretical*	*Empirical*
1	.250	.244	1.000	1.000
2	.813	.676	−.832	−.812
3	.953	.741	−.384	−.301
4	.988	.864	−.189	−.186
5	.997	1.240	−.094	−.366
6	.999	.967	−.047	−.221
7	1.000	.870	−.023	.083
8	1.000	1.203	−.012	−.148
9	1.000	.982	−.006	−.009
10	1.000	.881	−.003	−.080

Figure 1. THEORETICAL (LINE) AND EMPIRICAL (DOTS) VARIANCES OF $\hat{r}_1$



There are several additional comparisons which can be made based on certain functions of the $\hat{r}$'s. Thus we have seen that

$$\hat{l} = \sum \hat{\phi}^{k-1}\hat{r}_k = 0, \tag{3.4}$$

and in the course of our derivations we have had to make the approximation

$$l = \sum \phi^{k-1}\hat{r}_k \approx 0. \tag{3.5}$$

Some indication of the validity of this approximation is gained by examining the actual values of l from the sampling experiment, which were found to be distributed about zero with a variance of about one-hundredth that which would have been expected from the same linear form in white noise autocorrelations.

Of considerable importance because of its role in diagnostic checking is an examination of the quantity

$$n \sum_{k=1}^{m} \hat{r}_k^2 = 200 \sum_{k=1}^{20} \hat{r}_k^2, \tag{3.6}$$

which as in (2.38) should possess a χ^2—distribution with $\nu = m-1 = 19$ degrees of freedom. Such a distribution has a mean and variance of 19 and 38, respectively, with which the Monte Carlo values can be compared. When this was done, the overall or pooled empirical mean was found to be 18.1 and significantly different from 19. This difference is plausible, however, when it is realized that the statistic $n\sum_1^m \hat{r}_k^2$ possesses a χ^2_{m-p} distribution only insofar as the white noise autocorrelations $r = (r_1, \cdots, r_m)'$ have a common variance of $1/n$; and from (1.4) it is seen that this approximation overestimates the true variance of a given r_k by a factor of $(n+2)/(n-k)$. In particular, for $n = 200$, $m = 20$, and a typical value of $k = 10$, the actual variance $V(r_k)$ is $190/202 \approx 94$ percent of the $1/n$ approximation. Since the residual autocorrelations $\hat{r}$ are by (2.27) a linear transformation of r, it is reasonable to expect that a comparable depression of the variances of $\{\hat{r}_k\}$ would occur, and this would account for the discrepancy between the theoretical and empirical means of the statistic $200\sum_1^{20}\hat{r}_k^2$ encountered above. (This phenomenon would also explain the tendency for the empirical variances themselves, such as those in Table 1, to take on values averaging about 5 percent lower than those based on the matrix $(1/n)(I-Q)$ of (2.20).)

4. USE OF RESIDUAL AUTOCORRELATIONS IN DIAGNOSTIC CHECKING

We have obtained the large sample distribution of the residual autocorrelations $\hat{r}$ from fitting the correct model to a time series, and we have discussed the ways in which this distribution departs significantly from that of the white noise autocorrelations r. It is desirable now to consider the practical implications of these results in examining the adequacy of fit of a model.

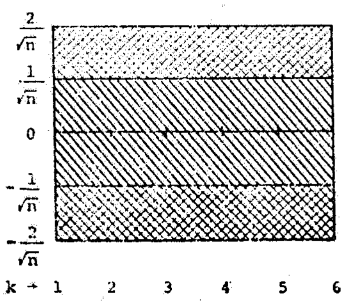
First of all it appears that even though the $\hat{r}$'s have a variance/covariance matrix which can differ very considerably from that of the r's, the statistic $n\sum_{k=1}^{m}\hat{r}_k^2$ will (since the matrix $I-Q$ is idempotent) still possess a χ^2-distribution, only now with $m-p$ rather than m degrees of freedom. Thus the overall χ^2-test discussed in Section 1 may be justified to the same degree of approximation as before when the number of degrees of freedom is appropriately modified.

However, regarding the "quality-control-chart" procedure, that is the comparison of the $\{\hat{r}_k\}$ with their standard errors, some modification is clearly needed.

Figure 2 shows the straight-line standard error bands of width $1/\sqrt{n}$ associated with any set of white noise autocorrelations $\{r_k\}$. These stand in marked contrast to the corresponding bands for the residual autocorrelations $\{\hat{r}_k\}$, derived from their covariance matrix $(1/n)(I-Q)$ and shown in Figure 3 for selected first and second order AR processes. Since it is primarily the $\hat{r}$'s of small lags that are most useful in revealing model inadequacies, we see that the consequence of treating $\hat{r}$'s as r's in the diagnostic checking procedure can be a serious underestimation of significance, that is, a failure to detect lack of fit in the model when it exists. Of course, if the model would have been judged inadequate anyway, our conviction in this regard is now strengthened.

Suppose, for example, that we identify a series of length 200 as first order

Figure 2. STANDARD ERROR LIMITS FOR WHITE NOISE AUTOCORRELATIONS r_k



autoregressive and after fitting $\hat{\phi}=.5$. Suppose also that $\hat{r}_1=.10$. Now the standard error of r_1 for white noise is $1/\sqrt{n}=.07$, so that $\hat{r}_1$ is well within the limits in Figure 2. Therefore if we erroneously regarded these as limits on $\hat{r}_1$ we would probably not conclude that this model was inadequate. However, if the true process actually were first order autoregressive (say with $\phi=.5$), the standard error of $\hat{r}_1$ would be $|\phi|/\sqrt{n}=.035$; since the observed $\hat{r}_1=.10$ is almost three times this value, we should be very suspicious of the adequacy of this fit.

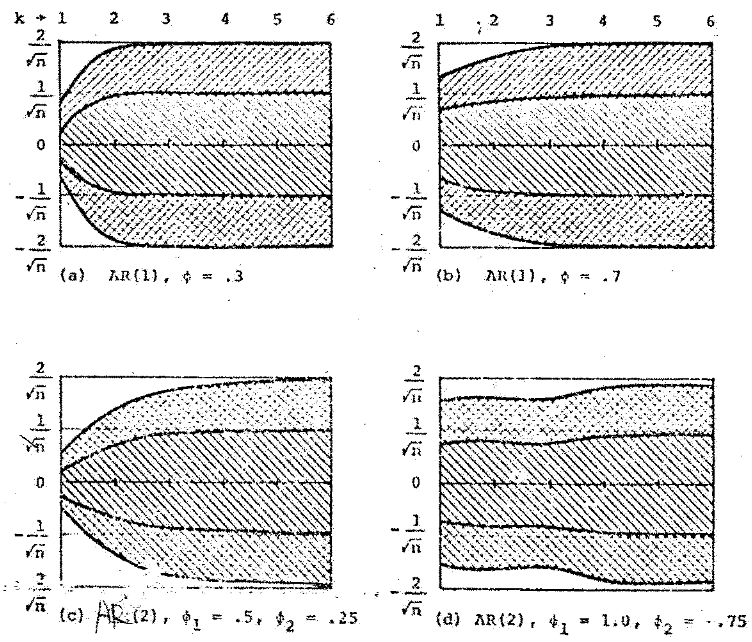
The situation is further complicated by the existence of rather high correlations between the $\hat{r}$'s, especially between those of small lags. For the first order process, the most serious correlation is

$$\rho[\hat{r}_1, \hat{r}_2] = -\frac{\phi}{|\phi|}\,\frac{1-\phi^2}{\sqrt{1-\phi^2+\phi^4}}$$

which, for example, approaches -1 as $\phi\to 0^+$ and is still as large as $-.6$ for $\phi=.7$. Correlation among the $\hat{r}$'s is even more prevalent in second and higher-order processes, where (as for variances) those involving lags up to $k=p$ can be particularly serious. From then on their magnitude is controlled by the recursive relationship (2.32); in particular, the closer ϕ is to the boundary of the stationarity region, the slower will be the dying out of $\mathrm{cov}(\hat{r}_k, \hat{r}_l)$ or $\rho(\hat{r}_k, \hat{r}_l)$ although often in these situations the less serious will the initial correlations $\rho(\hat{r}_1, \hat{r}_2)$, $\rho(\hat{r}_2, \hat{r}_3)$, $\rho(\hat{r}_1, \hat{r}_3)$, etc., tend to be.

We have thus seen that the departure of the distribution of the residual autocorrelations $\hat{r}$ from that of white noise autocorrelations r is serious enough to

Figure 3. STANDARD ERROR LIMITS FOR RESIDUAL AUTOCORRELATIONS $\hat{r}_k$



warrant some modifications in their use in diagnostic checking. The residual autocorrelation function, however, remains a powerful device for this purpose.

5. DISTRIBUTION OF RESIDUAL AUTOCORRELATIONS FOR THE MOVING AVERAGE AND GENERAL ARIMA PROCESSES

In obtaining the distribution of $\hat{r} = (\hat{r}_1, \cdots, \hat{r}_m)'$ for the pure autoregressive process in Section 2, considerable use was made of the recursive relation $\phi(B)\rho_k = 0$, which is not satisfied by moving average models $y_t = \theta(B)a_t$, or more generally by mixed models of the form (1.1) with $w_t = \nabla^d z_t$ denoting the stationary dth difference.

It is fortunate, therefore, that these models have in common with the pure AR models (2.1) an important property (derived in Section 5.1) because of which the distribution of their residual autocorrelations can be found as an immediate consequence of the autoregressive solution (2.29). This property is that if two time series, (a) the mixed autoregressive—moving average series (1.1), and (b) an autoregressive series

$$\pi(B)x_t = (1 - \pi_1 B - \cdots - \pi_{p+q}B^{p+q})x_t = a_t \tag{5.1}$$

are both generated from the *same set* of deviates $\{a_t\}$, and moreover if

$$\pi(B) = \phi(B)\theta(B), \tag{5.2}$$

then when these models are each fitted by least squares, their residuals, and hence also their residual autocorrelations, will be very nearly the same. Therefore if a mixed model of order (p, d, q) is correctly identified and fitted, its residual autocorrelations for n sufficiently large will be distributed as though the model had been of order $(p+q, d, 0)$ with the relations between the two sets of parameters given by (5.2). In particular the ψ's comprising the X-matrix (2.23) for the model (1.1) are the coefficients in $\psi(B) = [\phi(B)\theta(B)]^{-1}$.

5.1 Equality of Residuals in AR and ARIMA Models

Let w_t and x_t be as in (1.1) and (5.1); (5.2) then implies

$$w_t = \theta^2(B)x_t. \tag{5.3}$$

As in (2.5), define

$$\dot{a}_t^{AR} = a_t^{AR}(\dot{\pi}) = \dot{\pi}(B)x_t = -\sum_{j=0}^{p+q} \dot{\pi}_j x_{t-j} \tag{5.4}$$

where $\pi_0 = -1$, and now also

$$\dot{a}_t^* = a_t^*(\dot{\phi}, \dot{\theta}) = \dot{\phi}(B)\dot{\theta}^{-1}(B)w_t = [\sum_{i=0}^{p} \dot{\phi}_i B^i][\sum_{j=0}^{q} \dot{\theta}_j B^j]^{-1} w_t, \tag{5.5}$$

where $\dot{\phi}_0 = \dot{\theta}_0 = -1$. We will expand these quantities about the true parameter values and go through a least squares estimation in each case which is analogous to writing the linear regression model $y = X\beta + \varepsilon$ as

$$\dot{e} = y - \dot{y} = X(\beta - \dot{\beta}) + \varepsilon = X\delta + \varepsilon, \tag{5.6}$$

for fixed $\dot{\beta}$, and then performing the regression directly on $\dot{e}$ rather than on y. The equality of the residuals in the two cases depends heavily on the fact that the derivatives in each expansion involve the same autoregressive variable x_t.

Thus

$$\frac{\partial \dot{a}_t^{AR}}{\partial \dot{\pi}_j} = - x_{t-j}, \qquad 1 \leq j \leq p+q, \text{ irrespective of } \dot{\pi};$$

$$\frac{\partial \dot{a}_t^*}{\partial \dot{\phi}_j} = - \dot{\theta}^{-1}(B) w_{t-j}, \qquad 1 \leq j \leq p$$

$$= - \theta(B) x_{t-j} \qquad \text{at } (\dot{\phi}, \dot{\theta}) = (\phi, \theta); \quad \text{and}$$

$$\frac{\partial \dot{a}_t^*}{\partial \dot{\theta}_j} = \dot{\phi}(B) \dot{\theta}^{-2}(B) w_{t-j}, \qquad 1 \leq j \leq q$$

$$= \phi(B) x_{t-j} \qquad \text{at } (\dot{\phi}, \dot{\theta}) = (\phi, \theta).$$

Then

$$\dot{a}_t^{AR} = a_t^{AR} + \sum_{j=1}^{p+q} (\pi_j - \dot{\pi}_j) x_{t-j}, \tag{5.7}$$

and approximately

$$\dot{a}_t^* = a_t^* + \sum_{i=1}^{p} (\phi_i - \dot{\phi}_i)\theta(B) x_{t-i} - \sum_{j=1}^{q} (\theta_j - \dot{\theta}_j)\phi(B) x_{t-j} \tag{5.8}$$

$$= a_t^* + \sum_{i=1}^{p} (\phi_i - \dot{\phi}_i) x_{t-i} - \sum_{j=1}^{q} (\theta_j - \dot{\theta}_j) x_{t-j}$$
$$+ \sum_{i=1}^{p} \sum_{j=1}^{q} [\phi_i(\theta_j - \dot{\theta}_j) - \theta_j(\phi_i - \dot{\phi}_i)] x_{t-i-j}$$
$$= a_t^* + \sum_{i=1}^{p} (\phi_i - \dot{\phi}_i) x_{t-i} - \sum_{j=1}^{q} (\theta_j - \dot{\theta}_j) x_{t-j}$$
$$+ \sum_{i=1}^{p} \sum_{j=1}^{q} [\dot{\phi}_i(\theta_j - \dot{\theta}_j) - \dot{\theta}_j(\phi_i - \dot{\phi}_i)] x_{t-i-j} \tag{5.9}$$
$$= a_t^* + \sum_{j=1}^{p+q} (\beta_j - \dot{\beta}_j) x_{t-j}.$$

Thus letting $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_{p+q})'$ and $\boldsymbol{\lambda} = \begin{bmatrix}\boldsymbol{\phi}\\ \boldsymbol{\theta}\end{bmatrix}$, we see that

$$\boldsymbol{\beta} = A\boldsymbol{\lambda}, \tag{5.10}$$

where A is a $(p+q)$−square matrix whose elements involve $\dot{\boldsymbol{\lambda}}$ but not the true parameter values $\boldsymbol{\lambda}$. For example, if $p=q=1$, we would have

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix}1 & -1\\ -\dot{\theta} & \dot{\phi}\end{bmatrix}\begin{bmatrix}\phi\\ \theta\end{bmatrix} \tag{5.11}$$

Now equations (5.7) and (5.9) can be written as

$$\dot{\mathbf{a}}^{AR} = \mathbf{a} + X(\boldsymbol{\pi} - \dot{\boldsymbol{\pi}}) \tag{5.12}$$

$$\dot{\mathbf{a}}^* = \mathbf{a} + X(\boldsymbol{\beta} - \dot{\boldsymbol{\beta}}) \tag{5.13}$$

where the error in (5.13) is $O(|\boldsymbol{\beta} - \dot{\boldsymbol{\beta}}|^2)$, and where we have made use of the fact that, at $\dot{\boldsymbol{\pi}} = \boldsymbol{\pi}$, $\dot{\boldsymbol{\theta}} = \boldsymbol{\theta}$, and $\dot{\boldsymbol{\phi}} = \boldsymbol{\phi}$,

$$a_t^{AR} = a_t^* = a_t. \tag{5.14}$$

Thus in (5.12) the sum of squares

$$\mathbf{a}'\mathbf{a} = \sum a_t^2 = \sum [a_t^{AR}(\boldsymbol{\pi})]^2$$

is minimized as a function of $\boldsymbol{\pi}$ when

$$\boldsymbol{\pi} - \dot{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}} - \dot{\boldsymbol{\pi}} = (X'X)^{-1}X'\dot{\mathbf{a}}^{AR}, \tag{5.15}$$

while in (5.13) if we write

$$a^* = a + X[A(\lambda - \dot{\lambda})] = a + Z(\lambda - \dot{\lambda}),$$

then the sum of squares

$$a'a = \sum a_t^2 = \sum [a_t^*(\lambda)]^2$$

is minimized as a function of λ when

$$\lambda - \dot{\lambda} = \hat{\lambda} - \dot{\lambda} = (Z'Z)^{-1}Z'\dot{a}^* = A^{-1}(\hat{\beta} - \dot{\beta});$$

that is,

$$\hat{\beta} - \dot{\beta} = (X'X)^{-1}X'\dot{a}^*. \tag{5.16}$$

Then by setting $\dot{a} = a$ in (5.15) and (5.16), we have from (5.14) the important equality

$$\hat{\pi} - \pi = (X'X)^{-1}X'a = \hat{\beta} - \beta; \tag{5.17}$$

and finally by setting "." = "^" in (5.12) and (5.13), it follows from (5.17) that to $O_p(1/n)$

$$\hat{a}^{AR} = a + X(\pi - \hat{\pi}) = a + X(\beta - \hat{\beta}) = \hat{a}^*, \tag{5.18}$$

and thus (to the same order) $\hat{r}^{AR} = \hat{r}^*$, as we set out to show.

5.2 Monte Carlo Experiment

The equality (5.18) between the residuals from the autoregressive and mixed models depends on the accuracy of the expansion (5.8), that is, on the extent of linearity in the moving average model, between the true and estimated values θ and $\hat{\theta}$. It is therefore worthwhile to confirm this model-duality by generating and fitting pairs of series of the form (1.1) and (5.1) and comparing their residuals, or more to our purpose, their residual autocorrelations. This was done for $p+q=1$ and $p+q=2$ for series of length 200. Some indication of the close-

Table 2. RESIDUAL CORRELATIONS FROM FIRST ORDER AR AND MA TIME SERIES GENERATED FROM SAME WHITE NOISE ($n=200$)

k	$\phi=\theta=.1$		$\phi=\theta=.5$		$\phi=\theta=.9$	
	$\hat{r}_k^{AR}$	$\hat{r}_k^{MA}$	$\hat{r}_k^{AR}$	$\hat{r}_k^{MA}$	$\hat{r}_k^{AR}$	$\hat{r}_k^{MA}$
1	−.029	−.010	.003	−.005	−.048	−.057
2	.164	.169	.044	.045	.157	.151
3	.096	.099	−.098	−.096	.008	.009
4	−.050	−.049	.014	.021	−.126	−.127
5	−.003	−.006	.057	.058	.034	.035
6	−.143	−.144	.010	.012	−.091	−.090
7	−.023	−.026	−.004	.001	−.001	−.000
8	−.040	−.041	−.054	−.046	−.038	−.035
9	.010	.009	.052	.052	−.004	.000
10	−.049	−.049	−.065	−.067	.113	.116
$\hat{\phi}$ or $\hat{\theta}\rightarrow$	.159	.057	.543	.451	.922	.870

ness of the agreement is obtained from the few results for first order AR and MA processes shown in Table 2, where it is seen that the residual autocorrelation $\hat{r}_k^{AR}$ and $\hat{r}_k^{MA}$ are equal or nearly equal to the second decimal place.

A sampling experiment of the type described in Section 3 was also performed for the first order MA process. The results were very similar, which is to be expected in view of (5.18).

5.3 Conclusions

We have shown above that to a close approximation the residuals from any moving average or mixed autoregressive-moving average process will be the same as those from a suitably chosen autoregressive process. We have further confirmed the adequacy of this approximation by empirical calculation. It follows from this that we need not consider separately these two classes of processes; more precisely,

1. We can immediately use the AR result to write down the variance/covariance matrix of $\hat{r}$ for any autoregressive-integrated moving average process (1.1) by considering the corresponding variance/covariance matrix of $\hat{r}$ from the pure AR process

$$\pi(B)x_t = \theta(B)\phi(B)x_t = a_t. \qquad (5.19)$$

2. All considerations regarding the use of residual autocorrelations in tests of fit and diagnostic checking discussed in Section 4 for the autoregressive model therefore apply equally to moving average and mixed models.
3. In particular it follows from the above that a "portmanteau" test for the adequacy of any ARIMA process is obtained by referring $n\sum_{k=1}^{m} \hat{r}_k^2$ to a χ^2 distribution with ν degrees of freedom, where $\nu = m - p - q$.

REFERENCES

[1] Anderson, R. L., "Distribution of the Serial Correlation Coefficient," *The Annals of Mathematical Statistics*, 13 (March 1942), 1–13.

[2] Bartlett, M. S., "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series," *Journal of the Royal Statistical Society*, Series B, 8 (April 1946), 27–41.

[3] ——— and Diananda, P. H., "Extensions of Quenouille's Tests for Autoregressive Schemes," *Journal of the Royal Statistical Society*, Series B, 12 (April 1950), 108–15.

[4] Box, G. E. P. and Jenkins, G. M., *Statistical Models for Prediction and Control*, Technical Reports #72, 77, 79, 94, 95, 99, 103, 104, 116, 121, and 122, Department of Statistics, University of Wisconsin, Madison, Wisconsin, 1967.

[5] ———, *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day, Inc., 1970.

[6] ——— and Bacon, D. W., "Models for Forecasting Seasonal and Non-Seasonal Time Series," in B. Harris, ed., *Spectral Analysis of Time Series*, New York: John Wiley & Sons, Inc., 1967.

[7] Box, G. E. P. and Muller, M. E., "Note on the Generation of Random Normal Deviates," *The Annals of Mathematical Statistics*, 29 (June 1958), 610–11.

[8] Box, G. E. P. and Pierce, D. A., "Distribution of Residual Autocorrelations in Integrated Autoregressive-Moving Average Time Series Models," Technical Report #154, Department of Statistics, University of Wisconsin, Madison, April, 1968.

[9] Durbin, J., "Efficient Estimation of Parameters in Moving Average Models," *Biometrika*, 46 (December 1959), 306–16.

[10] ———, "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica*, 38 (May 1970), 410–21.

[11] Grenander, U. and Rosenblatt, M., *Statistical Analysis of Stationary Time Series*, New York: John Wiley & Sons, Inc., 1957.

[12] Mann, H. B. and Wald, A., "On the Statistical Treatment of Linear Stochastic Difference Equations," *Econometrica*, 11 (July 1943), 173–220.

[13] ——— "On Stochastic Limit and Order Relationships," *The Annals of Mathematical Statistics*, 14 (September 1943), 217–26.

[14] Quenouille, M. H., "A Large-Sample Test for the Goodness of Fit of Autoregressive Schemes," *Journal of the Royal Statistical Society*, Series A, 110 (June 1947), 123–9.

[15] Walker, A. M., "Note on a Generalization of the Large-Sample Goodness of Fit Test for Linear Autoregressive Schemes," *Journal of the Royal Statistical Society*, Series B, 12 (April 1950), 102-7.

[16] Whittle, P., "Tests of Fit in Time Series," *Biometrika*, 39 (December 1952), 309–18.

[17] Wold, H., *A Study in the Analysis of Stationary Time Series*, Stockholm: Almquist and Wiksell, 1938.

[18] ———, "A Large-Sample Test for Moving Averages," *Journal of the Royal Statistical Society*, Series B, 11 (April 1949), 297–305.

[19] Yule, G. U., "On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers," *Philosophical Transactions*, (A) 226 (July 1927), 267–98.

O. D. Anderson, ed., FORECASTING

FIRST ORDER AUTO-REGRESSIVE REGRESSION ANALYSIS

E. M. L. BEALE and R. G. SEELEY

Scicon Computer Services Ltd, Milton Keynes

This paper shows how a general-purpose linear multiple regression program can easily be modified to allow for serial correlation in the residuals when analysing time-series data. The methods are illustrated by re-examining the analyses of the *Financial Times* Ordinary Share Index made by Coen, Gomme and Kendall (1969) and by Box and Newbold (1971). We find that we are in substantial agreement with both sets of authors, although they come to opposite conclusions.

1. Introduction

We often want to fit values of some dependent variable y_i as a linear function of n independent (or predetermined or regressor) variables x_{ij}. In econometric work we may in particular wish to predict the future value of y_i from known future values of x_{ij}. The standard way to do this is by Ordinary Least Squares, which means finding parameters b_j to minimize

$$S = \sum_{i=1}^{N}\left(y_i - b_0 - \sum_{j=1}^{n} b_j x_{ij}\right)^2 \tag{1.1}$$

where summation extends over all observations i for which y_i and the x_{ij} are known. The justification for this procedure is the assumption that

$$y_i = \beta_0 + \sum_j \beta_j x_{ij} + \epsilon_i \tag{1.2}$$

where the ϵ_i are random errors, assumed to have zero means and equal variances, and to be statistically independent of each other.

Now in time series data it is common to find that successive ϵ_i are positively correlated. It is therefore usual to compute the Durbin–Watson statistic (the sum of the squares of the differences between successive residuals divided by the sum of the squares of individual residuals), since a value significantly less than 2.0 is evidence of positive serial correlation. But even if this statistic is significant one may be reluctant to give up the convenience of using the standard facilities of a good general-purpose multiple regression program.

On the other hand it is not very satisfactory to ignore this serial correlation, since when it is present Ordinary Least Squares does not give the best possible estimates b_j of the coefficients β_j, and more importantly the standard theory tends to exaggerate the precision to which the b_j are determined. In particular it may suggest that a coefficient is statistically significantly different from zero when it is not.

So we have taken the simplest possible generalization of the assumption that the ϵ_i are independent, and developed a way of finding maximum likelihood estimates of the parameters in the generalized model while preserving the format – and hence the standard facilities – of the general-purpose multiple regression program. Specifically we assume that the errors form a first order autoregressive process, so that

$$\epsilon_i = \rho\epsilon_{i-1} + \omega_i \tag{1.3}$$

where ρ is another parameter (to be determined) and the ω_i are assumed to be statistically independent. We also assume that ω_i is independent of ϵ_{i-1}. In Box–Jenkins notation, we assume that the residuals form a (1, 0, 0) ARIMA process.

In Section 2 we derive the formulae for the maximum likelihood estimates, and in Section 3 we discuss the results of applying the procedure to the models for the Financial Times Ordinary Share Index discussed by Coen, Gomme and Kendall (1969) and Box and Newbold (1971).

2. Theory

The theory of maximum likelihood estimation for autoregressive series is fairly elementary, but it seems worth presenting since this can be done compactly. The first part of the theory, which converts most of the problem into a weighted least squares problem, is well known. Other people also seem to know about the small modification that converts the formula for weighted least squares into one for true maximum likelihood estimation; but we have been unable to find a published reference.

From (1.2) and (1.3) we deduce that

$$y'_i = \beta'_0 + \sum_{j=1}^{n} \beta_j x'_{ij} + \omega_i \quad \text{for } i = 2, \ldots, N \tag{2.1}$$

where

$$\begin{aligned} y'_i &= y_i - \rho y_{i-1}, \\ \beta'_0 &= (1-\rho)\beta_0, \\ x'_{ij} &= x_{ij} - \rho x_{i-1,j}. \end{aligned} \tag{2.2}$$

We can now seek values of the parameters β'_0, β_j and ρ to minimize S_1 defined by

$$S_1 = \sum_{i=2}^{N} \left(y'_i - \beta'_0 - \sum_{j=1}^{n} \beta_j x'_{ij} \right)^2. \tag{2.3}$$

This is an example of a nonlinear regression model, but it is a particularly simple one since it is linear in the β_j for any given value of ρ. It can therefore most easily be solved by taking trial values for ρ and then using linear regression to find the β_j that minimize S_1 for each value of ρ. The optimum value of ρ is then found by a one-dimensional search for the value that minimizes S_1.

This approach omits some information contained in the first observation. This omission is not serious if we have a fair number of observations, but it can easily be rectified. A complete analysis requires a proper maximum likelihood approach, but the extra computing involved is negligible.

To incorporate the first observation we multiply (1.2) by $(1-\rho)$. It is then in the form (2.1) with $i = 1$ if

$$y'_1 = (1-\rho)y_1 \quad \text{and} \quad x'_{1j} = (1-\rho)x_{1j}. \tag{2.4}$$

But now ω_1 represents $(1-\rho)\epsilon_1$. So if ϵ_i have variance σ^2, then ω_1 has variance $(1-\rho)^2\sigma^2$ while the remaining ω_i all have variance $\tau^2 = (1-\rho^2)\sigma^2$. The optimum estimates of the β_j for any given ρ are therefore found by minimizing

$$S(\rho) = \sum_{i=1}^{N} w_i (y'_i - \beta'_0 - \sum_{j=1}^{n} \beta_j x'_{ij})^2 \tag{2.5}$$

where

$$w_1 = (1-\rho^2)/(1-\rho)^2 = (1+\rho)/(1-\rho) \tag{2.6}$$

and

$$w_i = 1 \quad \text{for } i > 1.$$

Maximum likelihood estimates are not obtained by choosing ρ to minimize $S(\rho)$, because the variances of the observations depend on ρ, and least squares is then not equivalent to maximum likelihood. We must also take note of the fact that the Jacobian of the transformation from the variables y_i' to y_i also depends on ρ, being $(1 - \rho)$. The likelihood function can therefore be written as

$$\sqrt{\frac{1+\rho}{1-\rho}}\left(\prod_{i=1}^{N}\frac{1}{\tau\sqrt{2\pi}}\right)\exp(-\tfrac{1}{2}S(\rho)/\tau^2)(1-\rho)\prod_{i=1}^{N}dy_i$$

and, omitting the irrelevant term $\Pi\, dy_i$, the log likelihood is

$$\tfrac{1}{2}\ln(1-\rho^2) - \tfrac{1}{2}N\ln 2\pi - N\ln\tau - \tfrac{1}{2}S(\rho)/\tau^2.$$

This is maximized with respect to τ by putting

$$\tau^2 = S(\rho)/N$$

in which case the log-likelihood becomes

$$\tfrac{1}{2}\ln(1-\rho^2) - \tfrac{1}{2}N\ln 2\pi - \tfrac{1}{2}N\ln S(\rho) + \tfrac{1}{2}N\ln N - \tfrac{1}{2}N.$$

The maximum of this expression with respect to ρ is therefore found by maximizing

$$\ln(1-\rho^2) - N\ln S(\rho)$$

or equivalently by minimizing

$$(1-\rho^2)^{-1/N}S(\rho). \tag{2.7}$$

For any given value of ρ it is almost as easy to minimize $S(\rho)$ defined by (2.5) as it is to minimize S_1 defined by (2.3). And the one-dimensional search for an optimum value of ρ is not significantly affected by the multiplication by $(1-\rho^2)^{-1/N}$.

If $\hat{\rho}$ denotes the value of ρ that minimizes (2.7), we can estimate the significance of the difference between this quantity and zero by an approximate F test in which $S(0) - S(\hat{\rho})$ is assigned one degree of freedom and $S(\hat{\rho})$ is assigned $N - n - 2$ degrees of freedom. We suggest that this is a simpler and more satisfactory test than any of those discussed by Durbin and Watson (1971), although we have not investigated its power.

Often in regression work one wants to choose between equations using different subsets of the independent variables. One could then consider re–

estimating ρ for each subset, but this seems overelaborate. Our program therefore estimates ρ once using all available independent variables. We can then use the optimum regression procedure of Beale, Kendall and Mann (1967) to select alternative subsets that minimize $S(\rho)$ defined by (2.5) for different numbers of nonzero β_j.

Our optimum regression procedure analyses the data as if ρ were known to equal $\hat{\rho}$, although we reduce the number of degrees of freedom by one. So, if $b_j(\rho)$ denotes the value of β_j that minimizes $S(\rho)$, we approximate the variance of β_j by var $b_j(\hat{\rho})$. But, once we have abandoned the hypothesis that $\rho = 0$, an exact analysis should allow for the fact that ρ is itself an unknown parameter. In fact we have the formula

$$\text{var } \hat{\beta}_j = \text{var } b_j(\hat{\rho}) + \left(\frac{\mathrm{d}b_j}{\mathrm{d}\rho}\right)^2 . \text{var } \hat{\rho} \tag{2.8}$$

provided that $\mathrm{d}b_j/\mathrm{d}\rho$ is approximately constant.

We can estimate $\mathrm{d}b_j/\mathrm{d}\rho$ by comparing $b_j(\hat{\rho})$ with $b_j(0)$, i.e. by comparing the estimates obtained by First Order Autoregressive Analysis with those obtained by Ordinary Least Squares. In practice we have found that the true standard error of $\hat{\beta}_j$ rarely exceeds that of $b_j(\hat{\rho})$ by more than about 20%, so the approximate analysis is adequate for most practical purposes. But Formula (2.8) reminds us it is useful to compare a First Order Autoregressive Analysis with Ordinary Least Squares, and to be somewhat suspicious if the results are very different.

3. Application

We have applied our methods to some of the models analysed by Coen, Gomme and Kendall (1969). This paper was given a rather hostile reception when it was read to the Royal Statistical Society, since it seemed unreasonable that the Financial Times index could usefully be forecast by simple regression on such independent variables as UK Car Production 6 quarters earlier. Box and Newbold (1971) criticized the paper for analysing the data as if the random errors were statistically independent, and suggested that the data could be explained less controversially, but less usefully, by a random walk process.

We did not obtain identical results to those of either Coen, Gomme and Kendall or Box and Newbold. This was disappointing because we tried to use the same data. We found one error in the published data. The FT Index for 1963/3 should read 320.3 not 295.7. We reran our analyses with this error corrected, and the small discrepancies with the original analyses were reduced but not eliminated. In any case we believe that our results represent a valid comparison between Ordinary Least Squares and First Order Autoregressive Regression Analysis on econometric data.

In broad terms our conclusions are:

(1) In several models, fitting the FT Index over various time periods, ρ was always significantly positive, ranging from 0.71 to 0.93.

(2) The t-values on the regression coefficients b_j were smaller, and the magnitudes of some of the b_j were smaller; but on the whole the same coefficients were statistically significant in the same sense.

(3) The standard deviations of the apparently random errors were reduced from about 20 points to about 16 points.

(4) Forecasts given by the new method are more plausible, since they take more note of the latest observation. But we do not attach great importance to this fact, since the interest of the forecast is in the predicted turning points and not in the absolute values.

To appreciate these results in more detail, let us look at one particular model that was studied in both the cited papers. This relates the FT Index from the 2nd Quarter of 1954 to the 4th Quarter of 1966 to UK Car Production 6 quarters earlier (x_{i1}), and to a Commodity Price Index 7 quarters earlier. Data are quoted for both the FT Commodity Index (x_{i2}), and Reuters' Commodity Index (x_{i3}), and we do not know why Coen, Gomme and Kendall used the FT Index since our analysis leads us to prefer the Reuters Index. But the point is not important. Optimum regression analysis using Ordinary Least Squares produces the following 2-variable equation:

$$\begin{aligned} y_i = 563 &+ 0.000460\, x_{i1} - 0.933\, x_{i3} \pm \epsilon_i \\ &\quad (0.000033) \qquad (0.088) \quad (21.41) \end{aligned}$$

while the first-order autoregressive analysis gives $\rho = 0.89$, and the following 2-variable equation:

$$\begin{aligned} y_i = 467 &+ 0.000253\, x_{i1} - 0.604\, x_{i3} + 0.89\, \epsilon_{i-1} \pm \omega_i. \\ &\quad (0.000067) \qquad (0.181) \qquad (0.17) \qquad (15.41) \end{aligned}$$

If we insist on using the FT Commodities Index rather than the Reuters Index the Ordinary Least Squares equation becomes

$$\begin{aligned} y_i = 664 &+ 0.000474\, x_{i1} - 6.24\, x_{i2} \pm \epsilon_i \\ &\quad (0.000033) \qquad (0.61) \quad (21.92) \end{aligned}$$

while the first-order autoregressive analysis gives $\rho = 0.93$, and the following 2-variable equation

$$y_i = 350 + \underset{(0.000072)}{0.000247}\, x_{i1} - \underset{(1.07)}{1.93}\, x_{i2} + \underset{(0.16)}{0.93}\, \epsilon_{i-1} \pm \underset{(16.22)}{\omega_i}.$$

A repeat of the last analysis using the data as quoted by Coen, Gomme and Kendall gives $\rho = 0.92$

$$y_i = 350 + \underset{(0.000078)}{0.000224}\, x_{i1} - \underset{(1.16)}{1.84}\, x_{i2} + \underset{(0.19)}{0.92}\, \epsilon_{i-1} \pm \underset{(17.70)}{\omega_i}$$

In all cases, the standard errors are calculated by the approximate method, omitting the correction term in (2.8). Box and Newbold's analysis is similar, but the implication that the regression coefficients are typically not significantly different from zero seems unwarranted. When the error in the data is corrected, both regression coefficients in the best fitting model are significant at the 3σ level, and remain so if the correction term in (2.8) is included. Even when we use the less appropriate Commodities Index one coefficient is significant at the 3σ level while the other is nearly significant at the 2σ level.

It is perhaps of some interest to note that if we use the same independent variables to fit the FT Index over a more recent period of time, from 1st Quarter 1965 to 3rd Quarter 1973, the regression coefficients are all insignificantly different from zero. This certainly implies that the empirical relationships developed over the period leading up to the 1967 devaluation of the Pound do not have enduring validity; but we do not ourselves accept the further deduction that they are useless. Since this is an essentially methodological paper, we did not think it appropriate to ask permission to discuss the models that were actually used to predict the FT Index during this later period.

We are encouraged by the results of Box and Newbold's analysis of alternative error structures. They found that the First Order Autoregressive Model fits marginally better than any of the others that they tried, but that the differences in the residual mean squares are small. We believe that this will often be the case when there are several independent variables with unknown regression coefficients. So the fact that our computational system only allows one specific but convenient assumed error structure may not be a practical disadvantage.

We therefore conclude that first-order autoregressive regression analysis is a convenient and practical tool for the analysis of econometric data. We have not yet used it on other time-series data, but we think it might still be useful. The main practical advantage over Ordinary Least Squares is not so much that the answers will be very different but that they inspire more confidence.

Acknowledgements

We are happy to acknowledge constructive criticism of the first draft of the paper by Dr. D. J. Payne and Dr. P. Newbold.

Discussion

Mr. G. J. A. Stern.

This paper appears to me to be mistaken on three grounds:

(1) If there really is a clear relationship between the FT Share Index and variables which are known in advance then this could be used to make large sums of money. Indeed it would destroy the stock exchange. Sir M. G. Kendall stated that he and Gomme had done quite well out of it and Martin Beale said that he had not the resources to gamble in this way. This to me is like someone saying that he has discovered how to make gold and has made a little, but won't spend £500 on the equipment to make much more.

(2) In the paper by Coen et al. several series (or variables) were used and several lags were tested, and 51 data points were available. As a lagged series is, from the point of view of fitting a regression equation, pretty well a new variable, Coen et al. were trying to fit maybe 30 or more variables on 51 data points. Some of these must fit well by chance. Martin Beale answered this point by saying that in his paper he had only taken one relationship: maybe, but this was the relationship selected by Coen et al. as the best, so my criticism applies to Martin's paper also.

(3) Box and Newbold tried a more general error structure

$$n_t = (1 - \theta) \sum_{j=1}^{\infty} a_{t-j} + a_t$$

which is a "noisy random walk" (a_t independent). For $\theta = 1$, this is the error structure of Coen et al. For $\theta = 0$ it is a pure random walk. They also showed that if one generated 5 random walks and detrended as Coen et al. did, then large correlations were observed just as Coen et al. witnessed. It seems to me that Martin has not at all answered these points. No doubt his error structure results in significant regression coefficients, just as that of Coen et al. did, but so what? It could be explained by this paragraph or 2 and while such a natural solution as that of Box and Newbold has been found, so much in accord with experience and common sense, it seems odd to advocate any other solution. It is like devising an elaborate explanation for a man's winning at roulette when the most likely cause is good luck.

Professor E. M. L. Beale

This paper is primarily a description of a convenient way of doing regression on time-series data without assuming that the errors are statistically independent. The application to the FT data was just an illustration of this tool and I have no strong views about whether any good results from this application should be attributed to skill or luck.

I certainly do not believe that there is a rule of nature saying that the FT Index varies in some predetermined way based on the past performance of the world's economies. On the other hand, I do not find it intrinsically incredible that an analysis of these economies should give some guidance as to when the turning points in the FT Index might occur. I am surprised but not completely incredulous at the idea that the relevant aspects of these economies can be summarized in a few simple statistics.

There are various reasons why I personally did not use this investment aid to make money for myself, but laziness and timidity are probably the most important ones.

Concerning the two more statistical criticisms, it is true that Coen et al. looked at several series and several lags, so there is an element of selection. On the other hand I am fairly sure that we could have found better variables, or better lags, or both, if our objective had been to fit this particular range of values of the FT Index as well as possible. So it is hard to know how to assess this point. It is also true that the residual mean square under the random walk hypothesis is not much larger than the residual mean square under the regression hypothesis; but a regression equation can be significant in every sense without accounting for more than a small fraction of the total variance.

References

Beale, E. M. L., Kendall, M. G. and Mann, D. W. (1967). The discarding of variables in multivariate analysis. *Biometrika*, 54, 357–366.

Box, G. E. P. and Newbold, P. (1971). Some comments on a paper by Coen, Gomme and Kendall. *J. Roy. Statist. Soc*, (A) 134, 229–240.

Coen, P. J., Gomme, E. D. and Kendall, M. G. (1969). Lagged relationships in economic forecasting. *J. Roy. Statist. Soc.*, (A) 132, 133–163.

Durbin, J. and Watson, G. S. (1971). Testing for serial correlation in least squares regression III. *Biometrika*, 58, 1–19.

THE PROBLEM OF AUTOCORRELATION IN REGRESSION ANALYSIS*

R. L. ANDERSON
North Carolina State College

1. INTRODUCTION

IN LEAST squares analysis, the usual regression model is

$$Y_t = \beta_0 + \sum_{i=1}^{r} \beta_i X_{it} + \epsilon_t, \qquad t = 1, 2, \cdots, n,$$

where the predictors, the X's, are assumed fixed in repeated sampling and the ϵ's independently distributed with the same variance, σ^2. The X's may be merely dummy variates (0 or 1), as in classification data (often called analysis of variance data). When tests of significance or confidence limits for the parameters are used, one usually assumes normality of the ϵ's. Even if the X's and Y follow a multivariate normal distribution, the least squares point and interval estimates of the β's can be used, and the usual null tests applied.

If the nY's are successive observations in time, the experimenter frequently wishes to investigate the nature of the response curve over time. In this case he might set $X_{it} = t^i$, or he might use the method of *harmonic analysis* to search for periodicities in Y. In other cases, the assumed model might involve lagged values of Y as predictors. For example,

$$Y_t = \beta_0 + \sum_{i=1}^{r} \beta_i Y_{t-i} + \epsilon_t. \tag{1}$$

This is an *autoregressive model.* Finally one could use a combined regression model with lagged Y's, present X's, lagged X's, and time as predictors. The method of least squares is applicable for autoregressive models, provided n is large [see Mann and Wald [6]].

One of the major difficulties with the use of least squares methods with time series is the strong possibility that the ϵ's are not independent. Aitken [1] pointed out that it is correlation of the ϵ's and not of the Y's which is to be avoided. It is possible that if the X's and Y's are both correlated in time, the errors will be relatively uncorrelated. A considerable amount of research has been devoted to the problem

* Paper presented at Annual Meeting of American Statistical Association, Chicago, December 27, 1952.

of testing for the existence of correlation in the errors, but all too little on the more important problem of the best estimation procedure when correlations do exist. Summaries of current methods of analyzing time series are given by Kendall [5] and Tintner [7].

The correlation of successive items in a time series was called a *lagged serial correlation* by Yule [9]. At the present time, it is more popular to use the term serial correlation to apply to the correlation between two series and the word *autocorrelation* for this correlation between successive items in a given series [see, for example, Tintner [7]]. I shall use this distinction. Many of the earlier papers on this subject, however, use the Yule terminology, as can be noted from the Bibliography. If we have a set of equally spaced values, $Z_1, Z_2, \cdots, Z_n$, selected from a population with zero mean, the autocorrelation coefficient of lag L is

$$r_L = \frac{SZ_iZ_{i+L}}{\sqrt{SZ_i^2SZ_{i+L}^2}}, \tag{2}$$

where i goes from 1 to $n-L$.[1] Most writers have preferred to use a definition in which the denominator is simply

$$\overset{n}{\underset{i=1}{S}} Z_i^2. \tag{3}$$

A symposium on autocorrelated time series analysis was held in 1946 under the auspices of the Royal Statistical Society. M. S. Bartlett [2] presented a general paper and Foster [4] and Cunningham and Hynd [3] presented papers on the use of autocorrelation methods in non-economic fields. J. W. Tukey at the 1951 Annual Meeting of the American Statistical Association proposed the use of the autocovariance (the numerator of r_L) in his method of spectrum analysis of time series.

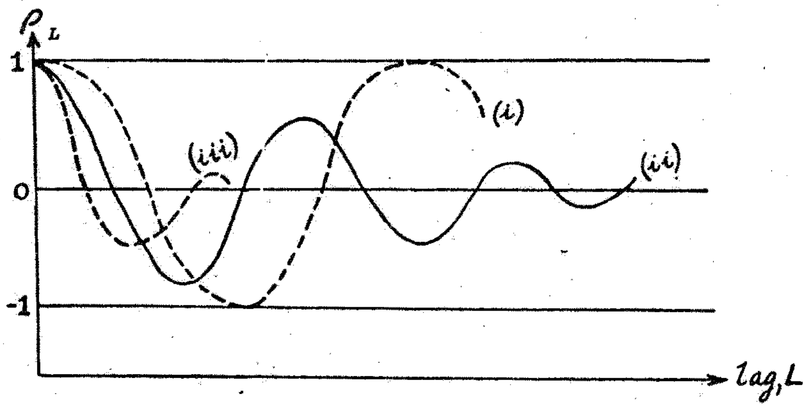
2. TESTS OF SIGNIFICANCE FOR AUTOCORRELATION

Yule [10] showed that the distribution of the correlation between two autocorrelated series tends to be U-shaped with a majority of the correlations near ±1. Bartlett [15] said that if the errors were autocorrelated, we could use the usual tests of significance of regression coefficients on a preliminary basis. If these coefficients were non-significant, accept the result; if they were significant, a test was needed which took account of the autocorrelation.

[1] S will be used to indicate summation over sample values.

One of the common methods of analyzing a single time series is *harmonic analysis*, in which the X's are sine and cosine terms. Fisher [18] presented a test of significance of the various amplitudes (the β's), in the restricted case of independent errors. Wilson [41] suggested that one compute successive lagged autocorrelation coefficients until the first non-significant one is reached; then use this lag (L) as an indication of the proportion of independent observations ($1/L$).

Three possible models are used to explain stationary trend-free time series data. Wold [8] indicated that the choice depended upon the relationship of successive true autocorrelation coefficients, ρ_L. These are usually displayed in a *correlogram*, as shown in the figure below.



(i) Repeated non-damped cycles: use harmonic analysis.
(ii) Damped correlations but with $|\rho| > 0$: use linear autoregression.
(iii) Damped correlations, with $\rho_L = 0$ for $L > m$: use the method of moving averages,

$$Y_t = \epsilon_t + \sum_{s=1}^{m} \gamma_s \epsilon_{t-s}. \tag{4}$$

Tintner [7] also discusses these methods in detail. Bartlett [2] cautions about the use of empirical correlograms to determine the correct model because successive sample autocorrelation coefficients tend to be highly correlated.

The approximate test of amplitudes in harmonic analysis and the decision regarding a proper model depends upon a test of significance for autocorrelation. For this reason the author decided to work on the distribution of r_L in 1939. Because of the mathematical difficulties involved, it was decided to follow up a suggestion of Hotelling to use a circular definition

$$r_L' = \frac{SZ_iZ_{i+L}}{SZ_i^2}, \tag{5}$$

where i goes from 1 to n, and $Z_{n+k} = Z_k$.

In 1941, the author studied the distribution for normal Z's when the population mean was zero, and, in 1942, the distribution for Z's which were deviations from the sample mean. Significance levels were computed for r_1', and for several cases of lags greater than one. The theory was simplified by the fact that

$$r' = \frac{\sum \lambda_i m_i}{\sum m_i},$$

where the m's are χ^2 variables with one degree of freedom, and the λ's are latent roots of the characteristic equation of the matrix of the coefficients in the numerator. Koopmans [22] reported on the distribution of r_1, as an estimate of ρ in the simple autoregressive model:

$$Y_t = \rho Y_{t-1} + \epsilon_t. \tag{6}$$

At the same time Dixon [16] was studying the moments of the distribution of r_1' and used Beta approximations to the exact distributions to obtain significance levels. T. W. Anderson [14] later showed that no test of the hypothesis $\rho = 0$ exists which is uniformly most powerful against alternatives of the Koopmans type.

Sometime before this, a problem involving autocorrelation came up in industrial quality control, in which the mean tended to creep up and down slightly on successive observations. In order to study the variation in the production process, von Neumann, *et al.* [35] suggested that the statistic

$$\frac{\delta^2}{2} = \overset{n-1}{\underset{i=1}{S}} (Z_{i+1} - Z_i)^2/2(n-1) \tag{7}$$

be used to estimate σ^2. Williams [40] and von Neumann [33, 34] studied the ratio δ^2/s^2, where $s^2 = SZ^2/n$ and $Z = Y - \bar{Y}$. Young [42] tabulated significance levels of a linear function of this ratio by use of an Incomplete Beta approximation. Hart [19, 20] tabulated probabilities by use of a series approximation suggested by R. H. Kent, We note that $\delta^2/s^2 = 2n(1 - r_c)/(n-1)$, where

$$r_c = \frac{\frac{1}{2}(Z_1^2 + Z_n^2) + \overset{n-1}{\underset{i=1}{S}} Z_i Z_{i+1}}{\overset{n}{\underset{i=1}{S}} Z_i^2}. \tag{8}$$

T. W. Anderson [14] showed that r_c could be used instead of r_1 to test the hypothesis $\rho = 0$ for Koopmans' model. T. W. Anderson transformed Hart's significance levels [20] to significance levels of r_c.

A non-parametric test for randomness by Wald and Wolfowitz [36] is based on the numerator of r_1' not corrected for the mean. Wallis and Moore [37] developed a series of non-parametric tests based on the signs of differences. Further contributions were made by Rubin [32], Madow [26] ,Hsu [21], Leipnik [25], Lehmann [23] and Quenouille [29, 30, 31].

If we let ϵ be the error vector and $\sigma^2\alpha$ its covariance matrix, dependent upon $\sigma^2, \rho_1, \rho_2, \cdots, \rho_{n-1}$, Lehmann and Stein [24] have shown that the best test statistic to test the hypothesis that all $\rho_i = 0$ is

$$\frac{\epsilon'\alpha^{-1}\epsilon}{\epsilon' I \epsilon}. \tag{9}$$

Whittle [39] used this method to test the null hypothesis that the data follow a first order moving average against the alternative that they follow an autoregressive scheme of first order, and vice versa.

T. W. Anderson and the author [13] derived the distribution of the circular autocorrelation coefficient for residuals from a fitted Fourier series. Significance levels were found and their use indicated. Exact distributions were possible because of the correspondence between the sine and cosine variables and the λ's in the distribution of the numerator.

Durbin and Watson [17] derived some approximate tests of autocorrelation of the successive residuals in least squares regression with fixed X's. Let the n successive least squares residuals be $Z_1, Z_2, \cdots, Z_n$. Durbin and Watson chose a modification of the von Neumann statistic,

$$d = \frac{\overset{n-1}{\underset{1}{S}} (Z_{i+1} - Z_i)^2}{\overset{n}{\underset{1}{S}} Z_i^2} = \frac{S(\Delta Z)^2}{SZ^2}, \tag{10}$$

to test for the existence of autocorrelation in the errors (ϵ's). We note

$$\frac{\delta^2}{s^2} = \frac{nd}{n-1},$$

so that that $d=2(1-r_c)$, where r_c was T. W. Anderson's statistic [14] to test the hypothesis $\rho=0$ for Koopman's model. It should be emphasized that the original von Neumann and T. W. Anderson statistics do not refer to deviations from a fitted regression; hence the Hart [20] and T. W. Anderson [14] significance levels cannot be used here. However, it would appear reasonable to expect that if we have a large positive autocorrelation, d should be near zero; and for a large negative autocorrelation, d should be near four.

Unfortunately an exact distribution could not be evaluated because the regression variables were not latent roots of the numerator matrix. Hence, only upper and lower bounds of the significance levels (d_U and d_L) could be computed. This was done for 5%, 2.5%, and 1% one-tailed tests, for $n=15$ (1) 40 (5) 100 and for $r=1$ (1) 5. It should be noted that d_U and d_L diverge more as r increases and also as n decreases.

In most cases, the experimenter desires a test of the null hypothesis against the alternative of positive correlation. Hence, one should expect a small value of d when the null hypothesis is false, and we should use the following testing procedure: If the computed value, d, is less than the tabulated value, d^*, the null hypothesis is rejected. On the other hand if the alternative hypothesis is negative correlation, one would expect a value of d near 4 when the null hypothesis is false. In this case we consider $d'=4-d$ and test d' against d^*, as above. Since only upper and lower bounds on the significance levels are available, we proceed as follows:

(i) If d (or d') is less than d_L, reject the null hypothesis.
(ii) If d (or d') is greater than d_U, do not reject.
(iii) If $d_L < d$ (or d') $< d_U$, the test is inconclusive.

If the experimenter wants a two-tailed test, he doubles the significance probability and proceeds as follows:

(i) If d or d' is less than d_L, reject.
(ii) If $d_U < d < 4-d_U$, do not reject.
(iii) Inconclusive otherwise.

An approximate procedure is available for large values of $(n-r-1)$, say greater than 40. In this case, $(1/4)d$ was transformed to a Beta distribution, as Dixon [16] did for r, with parameters p and q, where

$$p + q = \frac{E(d)[4 - E(d)]}{\sigma^2(d)} - 1$$

$$p = \tfrac{1}{4}(p + q)E(d).$$

An approximate test statistic is $F = [p(4-d)]/qd$ with $n_1 = 2q$ and $n_2 = 2p$ degrees of freedom. Or one can use Incomplete Beta tables. Durbin and Watson also present another approximation. Formulas for $E(d)$ and $\sigma^2(d)$ are presented in the 1951 article. Unfortunately, exact significance levels are really needed for small values of $(n-r-1)$, when d_L and d_U tend to be wide apart.

Durbin and Watson [17, 1951] also present methods of testing for autocorrelation with one- and two-way classification data and for curvilinear regression with equally spaced X's.

An example is presented for each of the three types of regression models. Short-cut methods of computing $S(\Delta Z)^2$ are presented for each case. Of course, SZ^2 is simply the error sum of squares. For example, with multiple or curvilinear regression, where $Y - \overline{Y}$ is estimated by

$$\sum b_i(X_i - \overline{X}_i), \qquad \Delta Z = \Delta Y - \sum b_i \Delta X_i.$$

Hence,

$$S(\Delta Z)^2 = S(\Delta Y)^2 + \sum_i \sum_j b_i b_j S(\Delta X_i - \Delta X_j) - 2 \sum_i b_i S(\Delta X_i \Delta Y).$$

Special formulas can be used for curvilinear regression, because of the orthogonal polynomials used in computing the regression coefficients.

Moran [28] presents an exact test of autocorrelation of the residuals, Z_i, when only one predictor is used. He uses the circular autocorrelation coefficient,

$$R_1 = \frac{SZ_i Z_{i+1}}{SZ_i^2},$$

where $Z_{n+1} = Z_1$, and gives formulas for $E(R_1)$ and $\sigma^2(R_1)$.

3. ESTIMATING REGRESSION COEFFICIENTS WHEN THE ERRORS ARE AUTOCORRELATED

To date most of the successful research on autocorrelation has been devoted to the problem of testing for its existence. All too little is known of what to do if the errors actually are autocorrelated. Aitken [1] first showed that if one knew the population covariance matrix for the ϵ's, he could transform the regression model so that the method of least squares would give efficient estimates of the β's. If the covariance matrix of the ϵ's is $\alpha\sigma^2$ and the regression model (in matrix form) is $Y = X\beta + \epsilon$, we premultiply this regression model by the non-singular matrix H, where

$$H\alpha H' = I,$$

and I is the $n \times n$ identity matrix. But even if α were known, the solution for H might be very difficult. However, if the ϵ's follow a first order autoregressive process with autocorrelation ρ and variance $\sigma^2/(1-\rho^2)$, the transformation is quite simple:

$$\epsilon_1^* = \sqrt{1-\rho^2}\,\epsilon_1, \quad \epsilon_i^* = \epsilon_i - \rho\epsilon_{i-1}, \quad \text{for} \quad 1 < i \leqq n.$$

The transformations for higher order autoregressive processes and for moving average processes are more complicated. A good explanation of this is given by Watson [54].

Allowing for the difficulty of making the transformation if α is known, the major defect is the lack of knowledge regarding the true value of α. Most time series are too short to enable one to derive good estimates of the parameters in α, or even to determine the type of process which is operating. A recent attempt to bypass the transformation problem when the ϵ's follow an autoregressive process was made by Champernowne [44]. He assumes that the model for the ϵ's is

$$\sum_{s=0}^{m} \gamma_s(\epsilon_{t-s} - \alpha) = \delta_t,$$

where the δ's are assumed normally and independently distributed with zero mean and variance σ^2. Champernowne presents the following results:

(i) Assuming the γ's are known, $\hat{\alpha}$ was determined as a weighted mean and $\hat{\sigma}^2$ as a weighted quadratic function of observed values of the ϵ's.
(ii) Assuming the γ's are known, estimates of and confidence limits for the regression coefficients are derived, both with α known and α unknown.
(iii) The results in (ii) are derived for $\alpha = 0$.
(iv) If the γ's are not known, the least-sqares estimates of the regression coefficients are not linear functions of the observed Y's and X's; hence, the usual χ^2 distribution theory does not hold exactly. A method involving the application of Bayes' Theorem was used in this case.
(v) A brief discussion is given of these problems when the X's also have disturbances.

Cochrane and Orcutt [46] indicate three principal reasons that the ϵ's in economic time series models tend to be positively autocorrelated:

(i) Faulty choice of the form of the regression model.
(ii) Omission of important variables from the model.
(iii) Use of incorrect variables or poor data.

They analyzed the sample residuals for a number of econometric studies, and found many significant autocorrelations, using von Neumann's statistic, δ^2/s^2. As indicated earlier, this statistic does not take account of the added correlation of the estimated residuals resulting from the necessity of estimating the regression coefficients; this defect becomes worse as the number of X's increase. In addition δ^2/s^2 does not take account of the autoregressive nature of many X-variables.

Cochrane and Orcutt also conducted some empirical sampling experiments to indicate the effects of autoregressive error processes on least squares regression analysis, [illegible] the following indicated results:

(i) The sample residuals tend to be biased towards randomness.
(ii) The variance of least squares estimates of the regression coefficients are very large if the errors are highly autocorrelated (in their example, $\rho \geqq .8$).
(iii) If the autocorrelations could be reduced to $\rho < .3$ or perhaps even $\rho < .5$, by use of a simple transformation, these variances appear to be close to those with random errors.
(iv) The removal of trend seems to be a crude but effective transformation in many cases.
(v) If sample residuals are used to estimate the error variance, σ^2, this estimate will be too small if the errors are positively correlated. This result can be proven exactly, see for example, Cochran [45].

Cochrane and Orcutt state that for many economic variables, it is a simple and practical procedure to analyze the first differences of the various series. If the original regression equation is

$$Y_t = \beta_0 + \sum_{i=1}^{r} \beta_i X_{it} + \epsilon_t,$$

the transformed equation will be

$$\Delta Y_t = \sum \beta_i \Delta X_{it} + \Delta\epsilon_t,$$

where $\Delta Z_t = Z_{t-1} - Z_t$. This would be the exact transformation for $\rho = 1$ in a first order autoregressive model, except that ρ must be less than 1 in order to avoid an explosive situation. However, the transformation should be reasonably good if ρ is near 1, and it is certainly very simple. If the sample residuals, after transforming the variables in this manner, are still highly autocorrelated, one might use the estimated autocorrelation coefficients to try a new transformation.

Stone [52] used the method of first differences advocated by Cochrane and Orcutt [46] to reanalyze his market demand data [51]. Stone uses the von Neumann statistic with the sample residuals to test for autocorrelation in the errors. He found the average autocorrelation for 13 analyses highly significant before transforming and almost equal to its expectation after transforming. It was interesting to note that the two sets of regression coefficients were not materially different.

Watson [54] has investigated the efficiencies and estimated variances of least-squares estimates of regression coefficients for fixed X's and tests of hypotheses concerning them, when an incorrect transforming model is used. General solutions of the following type are presented: bounds on the bias of the estimated variance, lower bound to the efficiency of the estimates of regression coefficients and some bounds on the significance points of the t- and F-tests. He then discusses the following special types of incorrect transformations:

(i) *Assumed and true error processes are both autoregressive.*
- (a) *Both are first order but an incorrect ρ_1 is used.* The greatest bias to the estimated variance is a downward bias when ρ is underestimated. This offers some justification for the use of the first difference transformation, which overestimates ρ. ρ is generally underestimated from sample residuals. However, we note low efficiencies of estimates of regression coefficients when ρ is overestimated unless ρ is nearly 1.
- (b) *True process is second order and assumed process is first order.* Results depend on how accurately one knows ρ_1 and on the magnitude of ρ_2.

(ii) *Assumed and true error processes are both moving averages.*
- (a) *Both are first order with incorrect ρ_1 used.* Results in (i) are reversed.
- (b) *True process is second order and assumed process is first order.* Indications are that an incorrect order is more serious for a moving average than for an autoregressive process.

(iii) *Assumed process is first order autoregressive and true process is first order moving average.* Even when ρ_1 is estimated correctly the bias in the variance can be appreciable and the efficiency quite low.

In all cases the true probabilities for 5% significance levels may be considerably different, the bounds being of the order of less than 1% to over 10% in many cases of what would appear to be only mildly inaccurate estimates.

Watson is rather pessimistic regarding the use of transforming devices to remove the effect of autocorrelation in least square analysis of time series data. However, he believes that more investigations need to be made of correlograms of residuals to see if a good analysis can be constructed on the basis of these correlograms. Quenouille [48] presents a test of the hypothesis that a sample was drawn from an autoregressive scheme of specified order and Wold [55] did the same for a moving average process. Similar tests are given by Bartlett and Diananda [43] and Walker [53]. However more efficient methods are needed, and especially we need to determine the proper process and order. After all, as Watson [54] remarks, one must use some kind of an analysis, and it is the duty of the statistician to find a good method, even if it is not the correct one.

A series of articles sponsored by the Indian Statistical Institute [47] describe the results of using empirical sampling methods to evaluate the usefulness of the Wold [55] and Quenouille [48] large sample tests for short series. Matthai and Kannan considered three different moving average models and S. R. Rao and Som two autoregressive models. Series of length 15 and 35 were used. It was shown that both large sample tests gave far too many significant results for the short series used. Quenouille's test showed that a second order autoregressive model would not fit third order moving average data; however, Wold's test indicated that a third order moving average model could be used even if the data were second order autoregressive. This may indicate that a moving average model of high order is more likely to represent a given set of data than is an autoregressive model. Or it may indicate that Quenouille's test is more powerful than Wold's in indicating the correct process. It was interesting to note that in both studies the correlogram was well estimated, if one knew the correct process. The third paper, by C. R. Rao, presents a sequential procedure for determining the number of sample autocorrelation coefficients needed to estimate the correlogram. Rao advocates the use of likelihood to discriminate between several possible models to represent a given set of data.

Sastry [49] used the above models and data to investigate the small sample bias in the estimates of the autocorrelation coefficients. He first compared definitions (2) and (3) and concluded that (2) was superior. However (3) is better for small lags and is certainly much easier to compute. In general small sample estimates have large biases, even for series of 100. The size of the bias depends on the type of model (it was much less for a second order autoregressive model than for the other four models) and on the values of the parameters in the model.

Sastry [49] also considers some theoretical results for comparing two series of autocorrelated variates, x and y. He presents the expected values of the means, variances, variances of the means, and covariances of x and y, and some higher moments for normal variates. He proposes this new statistic to test the hypothesis that $E(x) = \mu$:

$$t' = \frac{\sqrt{\left\{(n-1) - \frac{2}{n}\sum_k (n-k)\rho_k\right\} n(\bar{x}-\mu)^2}}{\sqrt{\left\{1 + \frac{2}{n}\sum_k (n-k)\rho_k\right\} S(x-\bar{x})^2}}$$

with

$$f' = \left\{n - 1 - \frac{2}{n}\sum_k (n-k)\rho_k\right\}$$

degrees of freedom. Sastry does not indicate how useful t' will be when ρ_k must be estimated from the data. One can surely see that relatively unbiased estimates need to be obtained. And, most important for regression analysis, he presents the expected values and variances of estimates of the parameters in $E(y)=\alpha+\beta x$.

4. FURTHER COMMENTS ON AUTOREGRESSIVE MODELS

Although the main topic of this paper is a discussion of regression analysis with fixed X's, some references on the use of autoregressive models will be included. These models were first discussed by Yule and have been used extensively by economists. The regression coefficients in these models are functions of the autocorrelation coefficients. Hurwicz [58] shows that least squares estimates of the parameters are biased in small samples. As indicated previously, Mann and Wald [6] showed that this bias approached zero as the sample size increased. Only large sample least-squares variances and covariances of the estimates of the parameters are available; hence, confidence limits for the parameters and predicted values are available only for large samples. Tintner [7] presents an example for a third-order process. Kendall [5, 59] gives further information on the use of least squares to estimate the parameters.

Bartlett [2] presents a method of estimation based on the concept of a continuous rather than a discrete process. Ghurye [57] has developed a method of using more of the autocorrelation coefficients in estimating the parameters. He introduces a superposed variation for each operation, so that the model is (assuming $\beta_0=0$):

$$(Y_t + \eta_t) = \sum_i \beta_i(Y_{t-i} + \eta_{t-i}) + \epsilon_t.$$

An extensive study of autoregressive analysis is presented by Orcutt [60].

Das [56] uses empirical sampling methods to measure the goodness of least squares methods for three equation economic models, in which one equation is quantity in terms of present prices and the others involve only lagged variables as predictors.

5. OMITTED TOPICS

The following topics, of importance in the analysis of time series, have not been discussed in detail.

(i) *The estimation of parameters in a multi-equation system.* For a discussion of this procedure, see for example Koopmans [62] and Klein [61]. An article by Orcutt and Cochrane [64] presents an empirical sampling study of the adverse effect of autocorrelation on the estimates of structural parameters in a multi-equation model. They concluded that, "Unless it it possible to specify something about the intercorrelations of the error terms in a set of relations and to choose approximately the correct auto-regressive transformation, a certain amount of skepticism is justified concerning the possibility of estimating structural parameters from aggregative time series of only twenty observations."

(ii) *Comparing two time series.* See Bartlett [15, 2], Orcutt and James [65] and Moran [63].

6. SUMMARY

Much research has been devoted to the distributions of various statistics used to test for the existence of autocorrelation of successive observations. Others have studied the problem of estimating parameters in various stochastic processes, such as autoregressive and moving average processes. A summary of this research is given in this paper.

Only recently has research been extended to the problem of testing for the existence of autocorrelated errors in regression models, such as

$$Y_t = \beta_0 + \sum_{i=1}^{r} \beta_i X_{it} + \epsilon_t, \qquad t = 1, 2, \cdots, n,$$

where the X's are fixed pred[illegible]rs and the ϵ's are normally distributed with equal variance. Durbin and Watson [17] present upper and lower bounds on the significance levels for making such tests. Moran [28] presents an exact test for $r=1$.

Too little information is available on the proper methods of estimating the β's when the ϵ's are autocorrelated. Aitken [1] indicated the exact method of transforming the regression variables when the autocorrelations were known. Champernowne [44] added to this general theory and presented a Bayesian method when the autocorrelations were not known.

Cochrane and Orcutt [46] used empirical sampling methods to indicate the effects of autocorrelated errors on the estimates of error and the β's. They showed that, in many cases, first differences of the Y's and X's would have a relatively uncorrelated error process. A series of articles in Sankhyā [47] have also used empirical sampling to indicate the large biases in testing and estimation procedures with small samples.

Watson [54] has shown the seriousness of using the wrong type of error process and incorrect estimates of the autocorrelations in transforming the regression variables. He concludes that the most fruitful research seems to be in utilizing more efficiently the estimates of the autocorrelations.

BIBLIOGRAPHY

1. *General*

[1] Aitken, A. C., "On least squares and linear combinations of observations," *Proceedings of the Royal Society of Edinburgh*, 55 (1935) 42–48.

[2] Bartlett, M. S., "On the theoretical specification and sampling properties of autocorrelated time series." *Journal of the Royal Statistical Society Supplement*, 8 (1946) 27–41.

[3] Cunningham, L. B. C. and Hynd, W. R. B., "Random processes in problems in air welfare," *Journal of the Royal Statistical Society Supplement*, 8 (1946) 62–85.

[4] Foster, G. A. R., "Some instruments for the analysis of time series and their applications to textile research," *Journal of the Royal Statistical Society Supplement*, 7 (1946) 42–61.

[5] Kendall, M. G., *The advanced theory of statistics*, Charles Griffin and Company, Limited, London, 1948.

[6] Mann, H. B. and Wald, A., "On the statistical treatment of linear stochastic difference equations," *Econometrica*, 11 (1943) 173–220.

[7] Tintner, G., *Econometrics*, John Wiley and Sons, Incorporated, New York, 1952.

[8] Wold, H., *A study in the analysis of stationary time series*, Almquist and Wiksells, Uppsala, 1938.

[9] Yule, G. U., "On the time-correlation problem," *Journal of the Royal Statiscal Society*, 84 (1921) 496–537.

[10] Yule, G. U., "Why do we sometimes get nonsense correlations between time series?", *Journal of the Royal Statistical Society*, 89 (1926) 1–64.

2. *Tests of Significance for Autocorrelation*

[11] Anderson, R. L., "Serial correlation in the analysis of time series," Unpublished Thesis, Library, Iowa State College, 1941.

[12] Anderson, R. L., "Distribution of the serial correlation coefficient," *Annals of Mathematical Statistics*, 13 (1942) 1–13.

[13] Anderson, R. L. and Anderson, T. W., "Distribution of the circular serial correlation coefficient for residuals from fitted Fourier series," *Annals of Mathematical Statistics*, 21 (1950) 59–81.

[14] Anderson, T. W., "On the theory of testing serial correlation," *Skandinavisk Aktuarietidskrift*, 31 (1948) 88–116.

[15] Bartlett, M. S., "Some aspects of the time-correlation problem in regard to tests of significance," *Journal of the Royal Statistical Society*, 98 (1935) 536–43.

[16] Dixon, W. J., "Further contributions to the problem of serial correlation," *Annals of Mathematical Statistics*, 15 (1944) 119–44.

[17] Durbin, J. and Watson, G. S., "Testing for serial correlation in least squares regression," *Biometrika*, 37 (1950) 409–28; 38 (1951) 159–78.

[18] Fisher, R. A., "Tests of significance in harmonic analysis," *Proceedings of the Royal Society Series A*, 125 (1929) 54–59.

[19] Hart, B. I., "Tabulation of the probabilities of the ratio of the mean square successive difference to the variance," *Annals of Mathematical Statistics*, 13 (1942a) 207–14.

[20] Hart, B. I., "Significance levels for the ratio of the mean square successive difference to the variance," *Annals of Mathematical Statistics*, 13 (1942b) 445–47.

[21] Hsu, P. L., "On the asymptotic distribution of certain statistics used in testing the independence between successive observations from a normal population," *Annals of Mathematical Statistics*, 17 (1946) 350–54.

[22] Koopmans, T., "Serial correlation and quadratic forms in normal variables," *Annals of Mathematical Statistics*, 13 (1942) 14–23.

[23] Lehmann, E. L., "On optimum tests of composite hypotheses with one constraint," *Annals of Mathematical Statistics*, 18 (1947) 473–94.

[24] Lehmann, E. L. and Stein, C., "Most powerful tests of composite hypotheses I Normal distributions," *Annals of Mathematical Statistics*, 19 (1948) 495–516.

[25] Leipnik, R. B., "Distribution of the serial correlation coefficient in a circularly correlated universe," *Annals of Mathematical Statistics*, 18 (1947) 80–87.

[26] Madow, W. G., "Note on the distribution of the serial correlation coefficient," *Annals of Mathematical Statistics*, 16 (1945) 308–10.

[27] Moore, G. H. and Wallis, W. A., "Time series significance tests based on signs of differences," *Journal of the American Statistical Association*, 38 (1943) 153–64.

[28] Moran, P. A. P., "A test for the serial independence of residuals," *Biometrika*, 37 (1950) 178–81.

[29] Quenouille, M. H., "Some results in the testing of the serial correlation coefficient," *Biometrika*, 35 (1948) 261–67.

[30] Quenouille, M. H., "Approximate tests of correlation in time series," *Journal of the Royal Statistical Society Supplement*, 11 (1949a) 68–84.

[31] Quenouille, M. H., "The joint distribution of serial correlation coefficients," *Annals of Mathematical Statistics*, 20 (1949b) 561–71.

[32] Rubin, H., "On the distribution of the serial correlation coefficient," *Annals of Mathematical Statistics*, 16 (1945) 211–15.

[33] von Neumann, J., "Distribution of the ratio of the mean square successive difference to the variance," *Annals of Mathematical Statistics*, 12 (1941b) 367–95.

[34] von Neumann, J., "A further remark on the distribution of the ratio of the

mean square successive difference to the variance," *Annals of Mathematical Statistics*, 13 (1942) 86–88.

[35] von Neumann, J., Kent, R. H., Bellinson, H. R., and Hart, B. I., "The mean square successive difference," *Annals of Mathematical Statistics*, 12 (1941a) 153–62.

[36] Wald, A. and Wolfowitz, J., "An exact test for randomness in the non-parametric case based on serial correlation," *Annals of Mathematical Statistics*, 14 (1943) 378–88.

[37] Wallis, W. A. and Moore, G. H., "A significance test for time series analysis," *Journal of the American Statistical Association*, 36 (1941) 401–09.

[38] Watson, G. S. and Durbin, J., "Exact tests of serial correlation using non-circular statistics," *Annals of Mathematical Statistics*, 22 (1951) 446–51.

[39] Whittle, P., *Hypothesis testing in time series*, Almqvist and Wiksells Boktryckeri A B, Uppsala, (1951).

[40] Williams, J. D., "Moments of the ratio of the mean square successive difference to the mean square difference in samples from a normal universe," *Annals of Mathematical Statistics*, 12 (1941) 239–51.

[41] Wilson, E. B., "Periodogram of American business activity," *Quarterly Journal of Economics*, 48 (1934) 375–417.

[42] Young, L. C., "On randomness in ordered sequences," *Annals of Mathematical Statistics*, 12 (1941) 293–300.

3. *Estimating Regression Coefficients When the Errors are Autocorrelated*

[43] Bartlett, M. S. and Diananda, P. H., "Extensions of Quenouille's test for autoregressive schemes," *Journal of the Royal Statistical Society Series B*, 12 (1950) 108–15.

[44] Champernowne, D. G., "Sampling theory applied to autoregressive sequences," *Journal of the Royal Statistical Society Supplement*, 10 (1948) 204–42.

[45] Cochran, W. G., "Some consequences when the assumptions for the analysis of variance are not satisfied," *Biometrics*, 3 (1947) 32–35.

[46] Cochrane, D. and Orcutt, G. H., "A sampling study of the merits of autoregressive and reduced form transformations in regression analysis," *Journal of the American Statistical Association*, 44 (1949) 356–72.

[47] Indian Statistical Institute, "The applicability of large sample tests for moving average and autoregressive schemes to series of short length—an experimental study." *Sankhyā*, 11 (1951) 217–72.

1. Matthai, A. and M. B. Kannan, "Moving averages."
2. Rao, S. R. and R. K. Som, "Autoregressive series."
3. Rao, C. R., "The discriminant function approach in the classification of time series."

[48] Quenouille, M. H., "A large-sample test for the goodness of fit of autoregressive schemes," *Journal of the Royal Statistical Society*, 110 (1948) 123–29.

[49] Sastry, A. S. R., "Bias in estimation of serial correlation coefficients," *Sankhyā*, 11 (1951a) 281–96.

[50] Sastry, A. S. R., "Some moments of moment statistics and their use in tests of significance in autocorrelated series," *Sankhyā*, 11 (1951b) 297–308.

[51] Stone, R., "The analysis of market demand," *Journal of the Royal Statistical Society*, 108 (1945) 1–98.
[52] Stone, R., "The analysis of market demand," *Review of the International Statistical Institute*, 10 (1948) 1–13.
[53] Walker, A. M., "Notes on a generalization of the large sample goodness of fit test for linear autoregression schemes," *Journal of the Royal Statistical Series B*, 12 (1950) 102–07.
[54] Watson, G. S., *Serial Correlation in Regression Analysis*, Unpublished Thesis, Library, North Carolina State College, also Institute of Statistics, Mimeo Series No. 49, (1951).
[55] Wold, H., "A large-sample test for moving averages," *Journal of the Royal Statistical Society, Series B*, 11 (1949) 297–305.

4. *Autoregressive Models*

[56] Das, A. C., "On the estimation of parameters in a recursive system," *Sankhyā*, 11 (1951) 273–80.
[57] Ghurye, S. G., "A method of estimating the parameters of an autoregressive time series," *Biometrika*, 37 (1950) 173–78.
[58] Hurwicz, L., "Least squares bias in time series," Chap. 15 in Koopmans, T. C., *Statistical Inference in Dynamic Economic Models*, John Wiley and Sons, Incorporated, New York, 1950.
[59] Kendall, M. G., "On autoregressive time series," *Biometrika*, 33 (1944) 105–22.
[60] Orcutt, G. H., "A study of the autoregressive anture of the time series used for Tinbergen's model of the economic system of the United States, 1919–32," *Journal of the Royal Statistical Society Supplement*, 10 (1948) 1–53.

5. *Other Topics*

[61] Klein, L. R., "The use of econometric models as a guide to economic policy," *Econometrica*, 15 (1947) 111–51.
[62] Koopmans, T. C., *Statistical Inference in Dynamic Economic Models*, John Wiley and Sons, Incorporated, New York, (1950).
[63] Moran, P. A. P., "Some theorems on time series," *Biometrika*, 34 (1947) 281–91; 35 (1948) 255–60.
[64] Orcutt, G. H. and Cochrane, D., "A sampling study of the merits of autoregressive and reduced form transformations in regression analysis," *Journal of the American Statistical Association*, 44 (1949) 356–72.
[65] Orcutt, G. H. and James, S. F., "Testing the significance of correlation between time series," *Biometrika*, 35 (1949) 397–413.

Autoregressive Modeling of Canadian Money and Income Data

CHENG HSIAO*

A sequential procedure based on Akaike's final prediction-error criterion and Granger's concept of causality to fit multiple autoregressions is suggested. The method not only allows each variable to enter the equation with a different time lag but also provides a reasonably powerful test of exogeneity or causality. The idea is applied to Canadian postwar money and income data. It is found that a bivariate feedback model for M1 and GNP and a one-way causal relation from GNP to M2 fit the data best. Diagnostic checks applied to our model seem to indicate the adequacy of our approach.

KEY WORDS: Final prediction error; Causality; Feedback.

1. INTRODUCTION

The purpose of this article is to study the strategy for fitting a vector autoregressive model. We are interested in fitting vector autoregressions as opposed to fitting econometric models based on supposed a priori knowledge or economic theory because on many occasions there are competing theories of economic activity. For example, the role of money supply is an area in which there is much disagreement among economists (Friedman 1970; Brunner and Meltzer 1964; Tobin 1970). Although theoretically we can construct a model based on different theories and carry out hypotheses testings to resolve the disputes, in practice, if the specification is incorrect, the estimates will be biased, and the model may become useless as a framework within which to do formal statistical tests.

To avoid infecting the model by spurious a priori constraints, Sims (1977) suggested an alternative strategy for empirical model building by treating all variables as joint dependent and fitting an unconstrained vector autoregression:

$$\underset{r \times 1}{y_t} = \sum_{\ell=1}^{M} \phi_\ell y_{t-\ell} + \underset{r \times 1}{u_t} , \qquad (1.1)$$

where y_t is a vector of r random variables and u_t is an $r \times 1$ vector of white-noise innovation term. There are several reasons for preferring such a formulation.

1. Under fairly general conditions a vector stationary time series admits an autoregressive representation (Masani 1966). A vector autoregressive model can be treated as unrestricted reduced form, thus offering an opportunity to drop the restrictions based on supposed a priori knowledge. Tests of economically meaningful hypotheses can still be executed at the second stage. For instance, Sims (1977) used quarterly, postwar time series for the United States and West Germany on money stock, real GNP, employment rate, price level, and import price index to illustrate that useful descriptive characteristics can be obtained by such a method.

2. The model can be used for controller design. Suppose some of the variables under consideration are government instruments and part are target variables. Then the model can be easily recast into a state space representation. Once a loss function is specified, say quadratic, the optimal values for the instruments over time can be resolved.

3. The formulation provides a reasonably powerful test of exogeneity of variables under consideration by testing the upper- or lower-block triangularity of autoregressive operator.

A major problem associated with the use of a vector autoregressive formulation, however, is in the determination of the order of the autoregressive process. The strategy Sims (1977) suggested that we emulate is that of frequency-domain time series theory in which what is being estimated is implicitly part of an infinite-dimensional parameter space. If we allow every variable to influence every other variable with the same length without restrictions, however, the number of parameters grows with the square of the number of variables and quickly exhausts degrees of freedom. For example, in a 10-variable system with the order of lags equal to 10, there will be no degrees of freedom left for the postwar quarterly time series data. Therefore, the maximum orders of lags of all the variables have to be small to limit the number of parameters to be estimated. But this often contradicts the observed phenomenon that it takes a long time for the influence of policy changes to make themselves felt.

Furthermore, by restricting the maximum order to be identical for all the variables, we end up with a large number of estimated coefficients with insignificant t statistics; thus, competing economic or behavioral

* Cheng Hsiao is Associate Professor, Department of Political Economy, and Research Associate, Institute for Policy Analysis, University of Toronto, Ontario, Canada M5S 1A1. Research support by the University of Toronto, Social Sciences and Humanities Research Council of Canada grant 410-78-0351-R1, and the National Science Foundation Grant SOC77-14944-A01 at the Institute for Mathematical Studies in the Social Sciences, Stanford University, is gratefully acknowledged. The author is indebted to C. W. [illegible] for computational assistance and wishes to thank S.E. Fienberg and J. Carr for helpful suggestions.

© Journal of the American Statistical Association
September 1979, Volume 74, Number 367
Applications Section

hypotheses often cannot be distinguished. This empirical phenomenon is further complicated by the fact that on many occasions the acceptance or rejection of the null hypothesis depends critically on the order of lags chosen (Hsiao 1978).

In addition to these problems, the distribution theory on which tests are based is asymptotic. For many of the hypotheses tested, the degree of freedom in the asymptotic χ^2 distribution for the likelihood ratio test is not a different order of magnitude from the degrees of freedom left in the data after fitting the model. Therefore, interpretation of the tests is difficult, because the tests will then be highly sensitive to nonnormality and different, reasonable-looking asymptotically equivalent formulas for the test statistic may give very different apparent significance levels for the same data.

To alleviate partially the problems associated with the general strategy suggested by Sims (1977) for estimating profligately parameterized macromodels, we propose a strategy for reducing the number of parameters to be estimated. Specifically, we wish to allow a variable to depend on a subset of variables under consideration and to allow each variable to enter the equation with different lags so that not only the number of parameters to be estimated can be reduced but also the influence of each variable can be felt at different time lags. The strategy is based on Akaike's (1969a,b) final prediction error (FPE) criterion and on Granger's (1969) definition of causality.

In Section 2 we discuss the general decision procedure for fitting vector autoregressive model. Section 3 applies the method to Canadian money and income data. Conclusions are in Section 4.

2. A DECISION PROCEDURE FOR FITTING VECTOR AUTOREGRESSIVE MODEL

For simplicity, we assume that the vector stationary time series **y** consists of two components $\{y, x\}$ only. Thus, we have a bivariate autoregressive model:

$$y_t = \psi_{11}(L)y_t + \psi_{12}(L)x_t + u_t \ , \qquad (2.1)$$

$$x_t = \psi_{21}(L)y_t + \psi_{22}(L)x_t + v_t \ , \qquad (2.2)$$

where $\psi_{ij}(L) = \sum_{l=1}^{M_{ij}} \psi_{ijl}L^l$, L is the lag operator, $Ly_t = y_{t-1}$, and the u_t, v_t are zero mean white-noise innovations with constant covariance matrix,

$$E\begin{pmatrix} u_t \\ v_t \end{pmatrix}(u_s \ \ v_s) = \delta_{ts}\Omega \ . \qquad (2.3)$$

We note that least squares can be applied to each equation and the estimates are consistent and asymptotically normally distributed. Therefore, we may ignore the correlation in the innovation for the moment and apply Akaike's (1969a) FPE criterion to each equation to determine the order of lags in ψ_{ij}. The FPE is defined as the (asymptotic) mean squared prediction error,

$$\text{FPE of } y_t = E(y_t - \hat{y}_t)^2 \ , \qquad (2.4)$$

where $\hat{y}_t$ is the predictor of y_t,

$$\hat{y}_t = \hat{\psi}_{11}^{m}(L)y_t + \hat{\psi}_{12}^{n}(L)x_t + \hat{a} \ . \qquad (2.5)$$

The superscripts m and n denote the order of lags in $\psi_{11}(L)$ and $\psi_{12}(L)$. The $\hat{\psi}_{11}^{m}(L)$, $\hat{\psi}_{12}^{n}(L)$, and $\hat{a}$ are least squares estimates when we treat the observations from $-M+1$ to 0 as fixed $\{t: t = -M+1, \ldots, 0, 1, \ldots, T\}$, $m, n \leq M$. Akaike (1969a,b) defines the estimate of FPE in this case by

$$\text{FPE}_y(m, n) = \frac{T+m+n+1}{T-m-n-1} \cdot \sum_{t=1}^{T} (y_t - \hat{y}_t)^2/T \ . \qquad (2.6)$$

The second term of the product term on the right side of the equality can be considered as a measure of modeling error and the first term as a measure of estimation error. The criterion tries to balance the risk resulting from the bias when a lower order is selected and the risk resulting from the increase of variance when a higher order is selected by choosing the specification that gives the smallest FPE. Doing so is equivalent to choosing the specification based on F test with varying significance level (Hsiao 1978).[1]

If M is the maximum order for ψ_{ij}, then one way to select the order of $\psi_{ij}(L)$ is to let each $\psi_{ij}(L)$ vary between 0 and M, which means that there will be $(M+1)^2$ combinations of $\psi_{11}(L)$ and $\psi_{12}(L)$ for y equation. If we let $M = 14$, we will have to compute 225 FPE's. To reduce the computation burden to less than $3M$, we make use of the definition of causality suggested by Granger (1969), even though there is not agreement regarding its appropriateness (see Zellner 1978).

Let $\bar{X}_t = \{x_s: s < t\}$ and $\bar{X}_t^* = \{x_s: s \leq t\}$, and similarly define $\bar{Y}_t$ and $\bar{Y}_t^*$. Denote by $\sigma^2(y_t|A)$ the mean squared error of the minimum mean squared prediction error of y_t given information set A. Granger's (1969) definition of causality and feedback are

Definition 1 (Causality): If $\sigma^2(y_t|\bar{Y}, \bar{X}) < \sigma^2(y_t|\bar{Y})$, we say that x is causing y, denoted by $x \Rightarrow y$.

Definition 2 (Feedback): If $\sigma^2(y_t|\bar{Y}, \bar{X}) < \sigma^2(y_t|\bar{Y})$ and $\sigma^2(x_t|\bar{Y}, \bar{X}) < \sigma^2(x_t|\bar{X})$, we say that feedback is occurring, denoted by $x \Leftrightarrow y$.

Definition 3 (Instantaneous Causality): If $\sigma^2(y_t|\bar{Y}, \bar{X}^*) < \sigma^2(y_t|\bar{Y}, \bar{X})$, we say that instantaneous causality of (x to y) is occurring.

Given these definitions, Caines and Chan (1975), Granger (1969), Pierce and Haugh (1977), and Sims (1972) have shown that $y \not\Rightarrow x$ is equivalent to $\psi_{21}(L) = 0$. Furthermore, if y does not cause x instantaneously (equivalently, x does not cause y instantaneously), Ω is diagonal. That $\psi_{21}(L) = 0$, however, is sufficient to

---

[1] The probability of selecting too low an order using Akaike's method approaches zero very quickly as sample size increases. Although it does not approach zero, the probability of selecting too high an order decreases fairly quickly. I think the cost of overfitting is less than the cost of underfitting in this kind of analysis. For a detailed analysis of the asymptotic distribution and risk of Akaike's method, see Shibata (1976).

establish the traditional endogenous-exogenous classification of variables, since if $\psi_{21}(L) = 0$, by premultiplying (2.1) and (2.2) by the unique upper triangular factor of Ω, $P(P\Omega P' = I)$, we can rewrite the autoregressive representation of the joint stationary time series in a standard dynamic simultaneous equation form. The current values of endogenous variables, y, depend on lagged endogenous variables, current and lagged exogenous variables, x; the exogenous variables x are generated by an independent process; and the innovations between y equations and x are uncorrelated. Thus, we shall ignore the notion of instantaneous causality or, more appropriately, the contemporaneous correlation in the innovations for system identification. Neither shall we attempt to identify models based on endogenous-exogenous classification of variables because doing so will complicate the matter of the order identified for $\psi_{ij}(L)$ by the premultiplication of upper (or lower) triangular factor P.

Combining the definition of causality and FPE criterion, we suggest the following sequential procedure for identifying bivariate autoregressive model:

1. Take y as the only output of the system. Determine the order of the one-dimensional autoregressive process for y by using the FPE criterion.

2. Assume x as the manipulated variable that controls the outcome of y. Use the FPE criterion to determine the lag order of x, say n, assuming that the order of the lag operator of y is the one specified in step 1, say s.

3. To check whether lagged value of y might pick up the effects of lagged x when y is treated as a one-dimensional autoregressive process, we let the order of lag operator ψ_{12} be fixed at n and let the order of lag operator ψ_{11} vary from 0 to s. Compute the corresponding FPE's. Choose the order of ψ_{11} that gives the smallest FPE (conditional on the order of ψ_{12} being n), say m. The m may or may not be equal to s.[2]

4. Compare the smallest FPE's of steps 1 and 3. If the former $(\text{FPE}_y(s, 0))$ is less than the latter, $(\text{FPE}_y(m, n))$, a one-dimensional autoregressive representation for y is used. If the converse is true, we say $x \Rightarrow y$, and the optimal model for predicting y is the one, including m lagged y and n lagged x.

5. Repeat steps 1 to 4 for the x process, treating y as the manipulated variable.

6. Put together all single-equation specifications to identify the system. Because the sequential procedure may bias the joint nature of the process and the single-equation approach is equivalent to ignoring the effect of possible correlations within the components of innovations, diagnostic checks are recommended to examine the adequacy of our model specification.[3] This can be carried out by treating the specification of the system as the maintained hypothesis and performing likelihood ratio tests by deliberately overfitting or underfitting the model.[4]

The procedure for fitting a bivariate autoregressive model can also be generalized to fitting a multivariate autoregressive model. For instance, assume $\mathbf{y}$ consists of (y, x, z). First repeat step 1 by taking y as the only output, then replace step 2 onward by the following:

2′. Assume other variables as the manipulated variables that control the outcome of y. Rank the importance of these manipulated variables in this equation. This rank may be based on a priori information (conjecture?) or on some supplementary statistical information such as the coherence between the input and the output. The rank is used to establish the precedence of introducing the manipulated variables sequentially in the later steps.

3′. Suppose x is more important than z in the y equation. We then assume that y is the only output and x is the only manipulated variable. Use the FPE criterion to determine the lag orders of x, say n, assuming that the order of the lag operator of y is the one specified in step 1, say s. Then compare this FPE with the smallest FPE of the one-dimensional autoregressive process. If the former $(\text{FPE}_y(s, n, 0))$ is greater than the latter $(\text{FPE}_y(s, 0, 0))$, we say $x \not\Rightarrow y$, and the equation consists of s lagged y only. If the converse is true, we say $x \Rightarrow y$ and tentatively identify the model, the order of ψ_{11} being s and that of ψ_{12} being n.

4′. Introduce z as the additional manipulated variable. Use the FPE criterion to determine the lag order of z, say r, assuming the orders of the lag operators of y and x to be the ones specified in step 3′ (s and n, respectively, if $x \Rightarrow y$). Compare this FPE with the smallest FPE obtained in the previous step. If the former $(\text{FPE}_y(s, n, r))$ is greater than the latter $(\text{FPE}_y(s, n, 0))$, we say $z \not\Rightarrow y$. If the converse is true, we say $z \Rightarrow y$ and **tentatively** identify the y equation as consisting of s lagged y, n lagged x, and r lagged z.

5′. There is the possibility that the order of the first introduced lag operators might be too high because of omitted variables effects. To check this, we first let the orders of the lag operators introduced last (in this case, ψ_{12} and ψ_{13}) be fixed at the orders specified in step 4′ (n and r) and let the order of lag operator ψ_{11} vary from 0 to s. Compute the corresponding FPE's and choose the order of ψ_{11} (say m) that gives the smallest FPE, where m may or may not equal s. Then fix the orders of the lag operators ψ_{11} and ψ_{13} at m and r and let the order of ψ_{12} vary from 0 to n. Compute the corresponding FPE's, and choose the order of ψ_{12} that gives the smallest FPE, say q. Again, q may or may not be equal to n. Thus, the

---

[2] Of course, a further reduction in parameters can be achieved by using McClave's (1975, 1978) suggestion of fitting a subset autoregression. Here we shall adopt the convention that if a higher-order coefficient is nonzero, then all the lower-order ones will be nonzeros.

[3] No further residual analysis is needed in this case, because the FPE formulas are derived under the assumption that the residuals are white noises.

[4] If the model fails its diagnostic tests, then we may have to try all possible combinations ψ_{ij} as indicated in the beginning of this section. I have yet to come across such a case in my limited experience.

optimal model so identified for predicting y is the one including m lagged y, q lagged x, and r lagged z.[5]

6′. Repeat steps 1 and 2′ to 5′ for the x and z processes, treating other variables as manipulated variables.

7′. Put together all single-equation specifications as the maintained hypothesis and perform diagnostic checks.

This sequential procedure would reduce considerably the computational burden of finding the optimum lag for ψ_{ij}. For a system of p variables, a complete procedure involves letting each ψ_{ij} vary between 0 and M. There will be $(M + 1)^p$ combinations of ψ_{ij} for the ith equation. In the case in which $p = 3$, $M = 14$, we will have to compute 3,375 FPE's for each of these three equations. Using the method suggested before, the computation for each equation can be reduced to a maximum of $[p + (p - 1)](M + 1)$ FPE's (here 75). Furthermore, in standard practice, the number of FPE's computed can be subtantially less. The limited experiments I have conducted with this approach indicate that it is feasible.

3. AUTOREGRESSIVE MODELING OF CANADIAN MONEY AND INCOME DATA

In this section we test the feasibility of the methodology by fitting an autoregressive model to Canadian money and income data. The relationship between money and income has been much debated in the economic literature. On the one side are the monetarists, who view money as an independent source of economic disturbance. On the other side are the critics of this view, who say money is a passive adapter to business conditions with little independent influence.

We use seasonally adjusted Canadian quarterly money stock and nominal GNP from 1955I to 1977IV as listed in the Appendix. Both M1 and M2 are used as alternative measures of money stock variables. We first take the second differences of the logarithm of each variable to remove the trend. Such a prefiltering procedure may be viewed as an attempt to eliminate the longer movement common to all the series or an attempt to remove the effects of a third variable so we may concentrate on the cyclical movements between money stocks and income.

The FPE of treating each variable as a one-dimensional autoregressive process is presented in Table 1 with the maximum M assumed to be equal to 14. The smallest FPE fo. M1, M2, and GNP is 9, 12, and 6, respectively. We then assume that each of money and income is an output variable and treat the other as an input variable. Holding the order of the autoregressive operator of the controlled variable to the one specified before, we compute the FPE of the controlled variable by varying the order of lags of the manipulated variable from 1 to 14. The order that gives the smallest FPE is presented in Table 2. As we can see for the GNP and M1 pair, adding the other variable reduces the FPE of the output. For the GNP and M2 pair, however, treating M2 as input increases the FPE of the GNP equation. On the other hand, the FPE of the M2 is marginally reduced when GNP is used as input compared with when it is not used. Therefore, for M1 and GNP we choose

$$\begin{pmatrix}(1-L)^2\log \text{GNP}\\(1-L)^2\log \text{M1}\end{pmatrix}=\begin{pmatrix}\psi_{11}{}^6(L) & \psi_{12}{}^8(L)\\ \psi_{21}{}^4(L) & \psi_{22}{}^9(L)\end{pmatrix}\cdot\begin{pmatrix}(1-L)^2\log \text{GNP}\\(1-L)^2\log \text{M1}\end{pmatrix}+\begin{pmatrix}a\\b\end{pmatrix}+\begin{pmatrix}u\\v\end{pmatrix}. \quad (3.1)$$

For M2 and GNP we choose

$$\begin{pmatrix}(1-L)^2\log \text{GNP}\\(1-L)^2\log \text{M2}\end{pmatrix}=\begin{pmatrix}\psi_{11}{}^6(L) & 0\\ \psi_{21}{}^2(L) & \psi_{22}{}^{12}(L)\end{pmatrix}\cdot\begin{pmatrix}(1-L)^2\log \text{GNP}\\(1-L)^2\log \text{M2}\end{pmatrix}+\begin{pmatrix}a\\b\end{pmatrix}+\begin{pmatrix}u\\v\end{pmatrix}. \quad (3.2)$$

Their full information estimates are presented in Tables 3 and 4.

The alternative approach to autoregressive modeling is to let every variable influence every other variable with the same order of lags (equation (1.1)). In this case we use Akaike's (1970) multiple final prediction error

1. FPE of Fitting a One-Dimensional Autoregressive Process for M1, M2, and GNP

The Order of Lags	FPE of M1 × 10^{-4}	FPE of M2 × 10^{-4}	FPE of GNP × 10^{-4}
1	3.775	1.624	2.021
2	3.281	1.610	1.741
3	3.329	1.621	1.786
4	3.003	1.399	1.720
5	2.777	1.399	1.765
6	2.742	1.381	1.701
7	2.794	1.418	1.741
8	2.671	1.385	1.763
9	2.572	1.404	1.742
10	2.641	1.409	1.755
11	2.713	1.417	1.792
12	2.743	1.333	1.816
13	2.802	1.370	1.789
14	2.794	1.394	1.819

2. The Optimum Lags of the Manipulated Variable and the FPE of the Controlled Variable

Controlled Variable[a]	Manipulated Variable	The Optimum Lag of Manipulated Variable	FPE × 10^{-4}
M1 (9)	GNP	4	2.372
M2 (12)	GNP	3	1.332
GNP (6)	M1	8	1.589
GNP (6)	M2	1	1.716

[a] The numbers in parentheses indicate the order of autoregressive operator in the controlled variable.

[5] This step can often be omitted or simplified by checking whether there is a substantial change in the t statistics associated with the coefficients of $\psi_{12}{}^q(L)$ and $\psi_{13}{}^r(L)$. If not, we can usually omit this step without committing a serious error.

optimal model so identified for predicting y is the one including m lagged y, q lagged x, and r lagged z.[6]

6′. Repeat steps 1 and 2′ to 5′ for the x and z processes, treating other variables as manipulated variables.

7′. Put together all single-equation specifications as the maintained hypothesis and perform diagnostic checks.

This sequential procedure would reduce considerably the computational burden of finding the optimum lag for ψ_{ij}. For a system of p variables, a complete procedure involves letting each ψ_{ij} vary between 0 and M. There will be $(M+1)^p$ combinations of ψ_{ij} for the ith equation. In the case in which $p = 3$, $M = 14$, we will have to compute 3,375 FPE's for each of these three equations. Using the method suggested before, the computation for each equation can be reduced to a maximum of $[p + (p-1)](M+1)$ FPE's (here 75). Furthermore, in standard practice, the number of FPE's computed can be subtantially less. The limited experiments I have conducted with this approach indicate that it is feasible.

3. AUTOREGRESSIVE MODELING OF CANADIAN MONEY AND INCOME DATA

In this section we test the feasibility of the methodology by fitting an autoregressive model to Canadian money and income data. The relationship between money and income has been much debated in the economic literature. On the one side are the monetarists, who view money as an independent source of economic disturbance. On the other side are the critics of this view, who say money is a passive adapter to business conditions with little independent influence.

We use seasonally adjusted Canadian quarterly money stock and nominal GNP from 1955I to 1977IV as listed in the Appendix. Both M1 and M2 are used as alternative measures of money stock variables. We first take the second differences of the logarithm of each variable to remove the trend. Such a prefiltering procedure may be viewed as an attempt to eliminate the longer movement common to all the series or an attempt to remove the effects of a third variable so we may concentrate on the cyclical movements between money stocks and income.

The FPE of treating each variable as a one-dimensional autoregressive process is presented in Table 1 with the maximum M assumed to be equal to 14. The smallest FPE for M1, M2, and GNP is 9, 12, and 6, respectively. We then assume that each of money and income is an output variable and treat the other as an input variable. Holding the order of the autoregressive operator of the controlled variable to the one specified before, we compute the FPE of the controlled variable by varying the order of lags of the manipulated variable from 1 to 14. The order that gives the smallest FPE is presented in Table 2. As we can see for the GNP and M1 pair, adding the other variable reduces the FPE of the output. For the GNP and M2 pair, however, treating M2 as input increases the FPE of the GNP equation. On the other hand, the FPE of the M2 is marginally reduced when GNP is used as input compared with when it is not used. Therefore, for M1 and GNP we choose

$$\begin{pmatrix}(1-L)^2 \log \text{GNP} \\ (1-L)^2 \log \text{M1}\end{pmatrix} = \begin{pmatrix}\psi_{11}{}^6(L) & \psi_{12}{}^8(L) \\ \psi_{21}{}^4(L) & \psi_{22}{}^9(L)\end{pmatrix} \cdot \begin{pmatrix}(1-L)^2 \log \text{GNP} \\ (1-L)^2 \log \text{M1}\end{pmatrix} + \begin{pmatrix}a \\ b\end{pmatrix} + \begin{pmatrix}u \\ v\end{pmatrix}. \quad (3.1)$$

For M2 and GNP we choose

$$\begin{pmatrix}(1-L)^2 \log \text{GNP} \\ (1-L)^2 \log \text{M2}\end{pmatrix} = \begin{pmatrix}\psi_{11}{}^6(L) & 0 \\ \psi_{21}{}^2(L) & \psi_{22}{}^{12}(L)\end{pmatrix} \cdot \begin{pmatrix}(1-L)^2 \log \text{GNP} \\ (1-L)^2 \log \text{M2}\end{pmatrix} + \begin{pmatrix}a \\ b\end{pmatrix} + \begin{pmatrix}u \\ v\end{pmatrix}. \quad (3.2)$$

Their full information estimates are presented in Tables 3 and 4.

The alternative approach to autoregressive modeling is to let every variable influence every other variable with the same order of lags (equation (1.1)). In this case we use Akaike's (1970) multiple final prediction error

1. FPE of Fitting a One-Dimensional Autoregressive Process for M1, M2, and GNP

The Order of Lags	FPE of M1 × 10⁻⁴	FPE of M2 × 10⁻⁴	FPE of GNP × 10⁻⁴
1	3.775	1.624	2.021
2	3.281	1.610	1.741
3	3.329	1.621	1.786
4	3.003	1.399	1.720
5	2.777	1.399	1.765
6	2.742	1.381	1.701
7	2.794	1.418	1.741
8	2.671	1.385	1.763
9	2.572	1.404	1.742
10	2.641	1.409	1.755
11	2.713	1.417	1.792
12	2.743	1.333	1.816
13	2.802	1.370	1.789
14	2.794	1.394	1.819

2. The Optimum Lags of the Manipulated Variable and the FPE of the Controlled Variable

Controlied Variable[a]	Manipulated Variable	The Optimum Lag of Manipulated Variable	FPE × 10⁻⁴
M1 (9)	GNP	4	2.372
M2 (12)	GNP	2	1.332
GNP (6)	M1	8	1.589
GNP (6)	M2	1	1.716

[a] The numbers in parentheses indicate the order of autoregressive operator in the controlled variable

[6] This step can often be omitted or simplified by checking whether there is a substantial change in the t statistics associated with the coefficients of $\hat{\psi}_{12}{}^{r}(L)$ and $\hat{\psi}_{13}{}^{r}(L)$. If not, we can usually omit this step without committing a serious error.

3. Autoregressive Estimates of GNP and M1

Coefficients on Lag of \ Dependent Variable		GNP	M1
GNP	(−1)ª	−.683 (−6.334)	.239 (1.831)
	(−2)	−.554 (−4.202)	−.065 (−.410)
	(−3)	−.4350 (−3.086)	−.300 (−1.928)
	(−4)	−.446 (−3.120)	−.387 (−3.050)
	(−5)	−.155 (−1.146)	
	(−6)	−.283 (−2.642)	
	(−7)		
	(−8)		
M1	(−1)	.078 (.876)	−.693 (−6.645)
	(−2)	.185 (1.760)	−.655 (−5.453)
	(−3)	.119 (1.083)	−.368 (−2.681)
	(−4)	.210 (2.136)	−.705 (−5.485)
	(−5)	.315 (3.229)	−.584 (−4.294)
	(−6)	.071 (.688)	−.376 (−2.867)
	(−7)	.116 (1.257)	−.260 (−2.10)
	(−8)	.192 (2.182)	−.284 (−2.328)
	(−9)		−.284 (−2.687)
Standard Error of the Regression		.010	0.013

ª The numbers in parentheses are *t* statistics.

(MFPE) (defined as the generalized variance of the one-step prediction error) as a criterion to determine the order of the bivariate autoregressive process. Both M1-GNP and M2-GNP pairs have the smallest MFPE at the ninth order. Comparing this result with the result of our approach, we note that a likelihood ratio test of (3.1) against a ninth-order GNP and M1 bivariate process with nine degrees of freedom has a chi-squared value of 9.27, indicating the acceptance of (3.1) as the appropriate model. For variables M2 and GNP, a likelihood ratio test of diagonality of the autoregressive operator against a ninth-order bivariate process has a chi-squared value of 12.56 with 18 degrees of freedom, hence indicating the acceptance of the null hypothesis of no causal relationship between M2 and GNP. If each of these series is treated as a univariate autoregressive process, however, then Table 1 indicates that M2 should be a 12th-order and GNP a sixth-order autoregressive process.

To check further the adequacy of specifications (3.1) and (3.2), a sequence of likelihood ratio tests was carried out by deliberately overfitting and underfitting (3.1) and (3.2). The results are reported in Tables 5 through 8 and do not seem to indicate any serious problem with our specification. Although the chi-squared value for testing model 1 against (3.2) is somewhat low (Table 8), it is not a reason for alarm. As indicated in Hsiao (1978), Akaike's FPE takes a more generous attitude in including additional variables than the conventional 5 and 10 percent significance levels.

The somewhat stronger response of Canadian money stock variable to income changes may be due to the Bank of Canada's dual goals of trying to maintain appropriate exchange rates and appropriate credit conditions. During the sample period, Canada has shifted from floating exchange rate in the 1950's to pegged rate in the 1960's and back to floating rate again in the 1970's. In the

4. Autoregressive Estimates of GNP and M2

Coefficients on Lag of \ Dependent Variable		GNP	M2
GNP	(−1)ª	−.742 (−6.661)	.158 (1.939)
	(−2)	−.597 (−4.325)	.132 (1.628)
	(−3)	−.270 (−1.855)	
	(−4)	−.398 (−2.764)	
	(−5)	−.207 (−1.553)	
	(−6)	−.233 (−2.150)	
M2	(−1)		−.508 (−4.716)
	(−2)		−.452 (−3.607)
	(−3)		−.242 (−2.003)
	(−4)		−.781 (−6.642)
	(−5)		−.465 (−3.553)
	(−6)		−.520 (−3.801)
	(−7)		−.322 (−2.369)
	(−8)		−.506 (−3.828)
	(−9)		−.262 (−2.239)
	(−10)		−.303 (−2.576)
	(−11)		−.223 (−1.953)
	(−12)		−.287 (−2.654)
Standard Error of the Regression		.012	.009

ª The numbers in parentheses are *t* statistics.

5. *Likelihood Ratio Tests of (3.1) Against Higher-Order Autoregressive Process*

Maximum Order Fitted for	Model 1	Model 2	Model 3	Model 4
ψ_{11}	9	6	7	8
ψ_{12}	9	9	8	8
ψ_{21}	9	5	4	4
ψ_{22}	9	10	10	11
Degrees of Freedom	9	3	2	4
Likelihood Ratio Statistic	9.27	1.48	1.44	1.44

1950's, however, there was very little fluctuation in the U.S. price of the Canadian dollar. Except for a short period during which the U.S. price of the Canadian dollar increased from 93.2 cents from 1970I to 98.1 cents in 1970IV, Canadian monetary policy in the 1970's was geared to exchange-rate consideration (Courchene 1977). In referring to the possibility of allowing the exchange rate to settle where it might be, the Bank of Canada states: "It would only be so if the level of the exchange rate were of no consequence to the well-being of the economy, and this is clearly not the case. The exchange rate is a very important price in a country that trades with the outside world on the scale that Canada does.... it is not possible to ignore it, even when it floats" (1970, p. 9). Thus, except for the one-shot phenomenon in 1970, we might think that the bank was acting as if Canada had an administered exchange rate. In a small, open economy with managed exchange rates, the private sector can more easily adjust money holdings to incomes than in a virtually closed economy such as the United States, because if the Bank of Canada does not provide this desired level of nominal balances, the difference will flow in via the balance of payments. Furthermore, the Bank of Canada's approach to policy is very closely related to the British approach, which is documented in the well-known Radcliffe Report (1959) that "the authorities thus have to regard the structure of interest rates rather than the supply of money as the centrepiece of the monetary mechanism." Indeed, the Bank of Canada makes its position on this issue very clear in Governor Rasminsky's (1967) report: "However, it is a fact that we do not operate on the basis of a precise view about the trend, over some period, of total chartered bank deposits. We give priority in our thinking to the kind of credit conditions that seem to be appropriate in the prevailing circumstances." Insofar as the Bank of Canada primarily aims to regulate the structure of interest rates, movements in the money stock can be expected to respond more to movements in nominal income.

The fact that nominal income responds to changes in M1 but not M2 might be the result of the instability in M2 rather than a contradiction of the monetarist position of nominal income's being closely related to the broad aggregates of the financial assets. The term structure of deposit rates was less stable in Canada than in the United States. With the existence of a great number of close substitutes issued by near-banks as well as by other financial institutions, structural shifts in intermediation can therefore occur—the effect of which might merely be to alter the charter bank's share but not the total value of these interest-bearing liabilities. A concrete example of this occurred in early 1972. In this sense, M1 as being mainly dictated by the transactions motives might be a better proxy to the broad aggregates of financial assets than M2. In fact, this more stable relation between M1 and nominal GNP was also observed by the Bank of Canada and was one of the major reasons why the bank favored M1 as the appropriate definition of money (Bank of Canada 1974).

4. CONCLUSION

In this article we suggested an alternative procedure to fitting multivariate autoregressive processes. Compared with the standard procedure of letting every variable enter into the equation with the same length, the method has the advantage of reducing the number

7. *Likelihood Ratio Test of (3.2) Against Higher-Order Autoregressive Process*

Maximum Order Fitted for	Model 1	Model 2	Model 3
ψ_{11}	6	6	6
ψ_{12}	1	4	1
ψ_{21}	3	2	2
ψ_{22}	12	12	12
Degrees of Freedom	2	4	1
Likelihood Ratio Statistic	1.95	3.49	1.67

6. *Likelihood Ratio Tests of (3.1) Against Lower-Order Autoregressive Process*

Maximum Order Fitted for	Model 1	Model 2	Model 3
ψ_{11}	6	6	6
ψ_{12}	6	0	8
ψ_{21}	2	4	0
ψ_{22}	9	9	9
Degrees of Freedom	4	8	4
Likelihood Ratio Statistic	12.67*	21.94**	14.39**

* Significant at 5% level.
** Significant at 1% level.

8. *Likelihood Ratio Tests of (3.2) Against Lower-Order Autoregressive Process*

Maximum Order Fitted for	Model 1	Model 2	Model 3
ψ_{11}	6	6	6
ψ_{12}	0	0	0
ψ_{21}	1	0	2
ψ_{22}	12	12	12
Degrees of Freedom	1	2	3
Likelihood Ratio Statistic	2.60*	4.21*	10.32**

* Significant at 15% level.
** Significant at 1% level.

of parameters to be estimated while making room for the effect of a variable to be felt after long delay as usually witnessed in economics.

It is well known that an autoregressive-moving average (ARMA) specification is more parsimonious than a purely autoregressive specification. The identification of an ARMA process is by no means trivial, however, even though we know that the maximum likelihood criterion can be used. The maximum likelihood estimation of a multivariate ARMA process is extremely costly. Recently, Granger and Newbold (1977) suggested an ingenious method for identification of a bivariate ARMA process. Apart from the problems mentioned in their book, the computation is still heavy. Yet if we know the process involving one-way causality or if the variables are unrelated, the computational burden can be substantially reduced by fitting a transfer function model or a univariate ARMA model using the method developed by Box and Jenkins (1970) and their associates. So if a more parsimonious specification is the aim, the autoregressive specification suggested here can still be used for the first-stage model identification. Least squares can then be applied with little computation cost.

The approach is applied to Canadian money and income data. In the course of fitting the multivariate autoregressive process, we found that between M1 and GNP a bivariate feedback model fit the data best; between M2 and GNP a one-way causal model from income to money performed better. The result is similar, but not identical, to the earlier result obtained by Barth and Bennett (1974) by using Sims's (1972) method for 1959-72 sample periods. In their study, feedback between M1 and income was found, but M2 and income showed no relationship. Apart from the different sample periods covered, this difference might be reconciled by the size and the power of the test. Our test is a likelihood ratio test when no a priori constraints are imposed on the model. The choice of significance level is based on an explicit optimality criterion rather than on the ad hoc 5 or 10 percent significance level. The method used by Barth and Bennett (1974) is not a likelihood ratio test. Neither is their method immune from the problem of serial correlation in the residuals (Granger and Newbold 1974; Pierce and Haugh 1977).

APPENDIX—THE DATA SET

The nominal GNP data are those reported by Statistics Canada, "National Income and Expenditure Accounts" (Cat. No. 13-001, CANSIM No. D40252). The source of M1 data is Bank of Canada Review, "Currency and Demand Deposits" (CANSIM B1609). For M2 data I have adopted the M2 series developed by Data Resources of Canada Inc. for its quarterly Canadian model.

	GNP	M1	M2
1955I	27,244	4,498.921	10,540.340
II	28,140	4,505.307	[illegible]
III	[illegible]	[illegible]	[illegible]
IV	29,588	4,715.184	[illegible]
1956I	30,936	4,680.165	10,905.190
II	31,600	4,715.116	11,014.092
III	32,496	4,708.097	11,047.270
IV	33,200	4,646.447	11,130.667
1957I	33,232	4,595.608	11,154.227
II	33,364	4,609.884	11,188.706
III	33,992	4,566.913	11,190.707
IV	33,464	4,633.743	11,319.019
1958I	33,776	4,755.419	11,566.619
II	34,804	4,953.435	11,935.421
III	35,008	5,169.048	12,429.856
IV	35,520	5,297.765	12,808.187
1959I	36,140	5,250.933	12,872.176
II	36,644	5,177.551	12,878.064
III	37,116	5,164.804	12,875.971
IV	37,484	5,099.308	12,698.657
1960I	38,564	5,134.487	12,704.452
II	37,852	5,225.318	12,801.631
III	38,432	5,252.491	12,915.736
IV	38,588	5,341.248	13,176.262
1961I	38,288	5,442.060	13,492.801
II	39,444	5,412.759	13,665.900
III	40,004	5,552.388	14,009.525
IV	40,848	5,636.783	14,275.307
1962I	42,280	5,639.611	14,367.421
II	42,116	5,651.840	14,589.296
III	43,240	5,644.082	14,572.603
IV	44,072	5,844.950	14,823.329
1963I	44,656	5,898.588	15,117.275
II	45,284	6,051.910	15,460.418
III	45,992	6,070.066	15,614.788
IV	47,980	6,103.871	15,810.436
1964I	49,016	6,238.130	16,062.765
II	49,604	6,326.641	16,320.225
III	50,812	6,350.325	16,514.137
IV	51,688	6,424.917	16,859.937*
1965I	53,404	6,533.178	17,423.528
II	54,416	6,695.585	18,054.971
III	56,064	6,841.400	18,613.219
IV	57,572	6,877.323	18,976.243
1966I	60,192	7,019.164	19,259.228
II	61,492	7,123.188	19,550.070
III	62,308	7,244.355	[illegible]
IV	63,320	7,427.842	20,264.657
1967I	64,776	7,705.250	21,008.778
II	66,432	7,869.525	21,616.761
III	66,586	7,978.059	22,479.391
IV	67,892	8,060.411	23,562.654
1968I	69,600	8,024.632	23,718.447
II	70,652	8,011.822	24,576.204
III	73,524	8,413.767	[illegible]
IV	75,508	8,534.140	26,409.126
1969I	77,728	8,739.887	27,495.965
II	78,084	8,935.103	27,609.926
III	80,460	8,868.640	27,425.962
IV	82,348	8,912.005	27,712.829
1970I	83,944	8,963.569	27,820.959
II	[illegible]	[illegible]	[illegible]
III	86,680	9,132.144	[illegible]
IV	87,236	9,227.995	[illegible]
1971I	89,808	9,606.875	31,457.183
II	[illegible]	[illegible]	[illegible]
III	[illegible]	10,406.100	[illegible]
IV	95,752	10,831.577	35,118.619

	GNP	M1	M2
1972I	100,584	11,111.811	36,765.461
II	104,044	11,356.617	38,764.352
III	106,064	11,817.585	39,958.385
IV	110,244	12,315.742	40,946.528
1973I	116,652	12,761.954	42,011.389
II	120,392	13,202.293	43,312.753
III	124,572	13,615.371	44,808.365
IV	122,624	13,767.463	47,323.924
1974I	139,656	14,227.200	50,093.558
II	145,320	14,911.534	52,117.397
III	150,164	14,667.350	54,253.188
IV	153,560	14,748.277	56,511.734
1975I	157,328	15,632.949	59,242.715
II	161,740	16,175.896	60,782.654
III	168,732	16,846.700	63,738.104
IV	173,980	17,878.904	66,338.143
1976I	182,744	17,716.054	68,693.811
II	190,172	17,792.310	72,238.405
III	191,592	18,143.953	74,543.816
IV	195,600	18,222.515	77,300.262
1977I	201,204	18,742.929	80,021.261
II	204,160	19,161.165	83,481.546
III	210,780	19,665.498	85,867.813
IV	214,712	20,138.906	87,911.290

[*Received August 1978. Revised April 1979.*]

REFERENCES

Akaike, H. (1969a), "Statistical Predictor Identification," *Annals of the Institute of Statistical Mathematics*, 21, 203–217.

——— (1969b), "Fitting Autoregressions for Prediction," *Annals of the Institute of Statistical Mathematics*, 21, 243–247.

——— (1970), "Autoregressive Model Fitting for Control," *Annals of the Institute of Statistical Mathematics*, 22, 163–180.

Bank of Canada (1970), Annual Report, Ottawa.

——— (1974), Annual Report, Ottawa.

Barth, J.R., and Bennett, J.T. (1974), "The Role of Money in the Canadian Economy: an Empirical Test," *Canadian Journal of Economics*, 7, 305–311.

Box, G.E.P., and Jenkins, G.M. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Brunner, K., and Meltzer, A.H. (1966), "A Credit Market Theory of the Money Supply and an Explanation of Two Puzzles in U.S. Monetary Policy," Essays in Honor of Macro Fanno, Padua, Italy.

Caines, P.E., and Chan, C.W. (1975), "Feedback Between Stationary Stochastic Process," *The Institute of Electrical and Electronics Engineers, Transactions on Automatic Control*, AC–20, 498–508.

Courchene, T.J. (1977), *Money, Inflation, and the Bank of Canada: An Analysis of Canadian Monetary Policy From 1970 to Early 1975*, C.D. Howe Research Institute.

Friedman, M. (1970), "A Theoretical Framework for Monetary Analysis," *Milton Friedman's Monetary Framework*, ed. R.J. Gordon, Chicago: Chicago University Press, 1–62.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–438.

Granger, C.W., and Newbold, P. (1974), "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111–120.

——— (1977), *Forecasting Economic Time Series*, New York: Academic Press.

Hsiao, C. (1978), "Autoregressive Modelling and Money-Income Causality Detection," Technical Report No. 262, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Masani, P. (1966), "Recent Trends in Multivariate Prediction Theory," in *Multivariate Analysis I*, ed. P.R. Krishnaiah, New York: Academic Press, 351–382.

McClave, J. (1975), "Subset Autoregression," *Technometrics*, 17, 213–220.

——— (1978), "Estimating the Order of Autoregressive Models: The Maxχ^2 method," *Journal of the American Statistical Association*, 73, 122–128.

Pierce, D.A., and Haugh, L.D. (1977), "Causality in Temporal Systems: Characterizations and Survey," *Journal of Econometrics*, 5, 265–293.

Radcliffe Report (1959), Committee on the Working of the Monetary System, London: Her Majesty's Stationery Office.

Rasminsky, L. (1967), "Central Banking in the Canadian Financial System," opening address by the Governor of the Bank of Canada to the 20th International Banking Summer School, Queen's University.

Shibata, R. (1976), "Selection of the Order of an Autoregressive Model by Akaike's Information Criterion," *Biometrika*, 63, 117–126.

Sims, C.A. (1972), "Money, Income and Causality," *American Economic Review*, 62, 540–555.

——— (1977), "Macro-Economics and Reality," Fisher-Schultz Lecture, Discussion Paper No. 77–91, Center for Economic Research, University of Minnesota.

Tobin, J. (1970), "Money and Income: Post Hoc Ergo Propter Hoc?" *Quarterly Journal of Economics*, 84, 301–317.

Zellner, A. (1978), "Causality and Econometrics," paper presented at a joint University of Rochester–Carnegie-Mellon University Conference.

Distribution of the Residual Cross-Correlation in Univariate ARMA Time Series Models

A. IAN McLEOD*

Cross-correlations between univariate autoregressive moving average (ARMA) time series residuals are useful in the examination of relationships between time series (Pierce 1977a) and in the identification of dynamic regression models (Haugh and Box 1977). In this article, the asymptotic distribution of these residual cross-correlations is derived, and its application to the problem of testing for lagged relationships in the presence of instantaneous causality is discussed. Some results of a simulation study to investigate the accuracy of the asymptotic variances and covariances of the residual cross-correlations in finite samples are reported.

KEY WORDS: ARMA time series; Granger causality; Model identification; Relationships between time series; Residual cross-correlations.

1. INTRODUCTION

One approach to the problem of the elucidation of the relationship between two time series is to examine the cross-correlation function of the residuals of univariate models fitted to the two series. Interestingly enough, this approach seems to have first been suggested by Fisher (1921) with polynomial trend models. Of course, now it is understood that stochastic time series models, such as the ARMA model, are more realistic and perform better in applications such as forecasting (Box and Jenkins 1970, Ch. 1). Some recent applications of the univariate ARMA residual cross-correlation approach are mentioned in Sections 1.2 and 3. In this article, the large-sample distribution of the residual cross-correlations in univariate ARMA models is derived, and the finite sample accuracy of the derived asymptotic variances and covariances of the residual cross-correlations is investigated by simulation. An application to the problem of testing for lagged relationships in the presence of instantaneous causality is discussed and a brief economic example given.

1.1 Univariate ARMA Time Series Models

The theory and application of univariate ARMA models is discussed in a book by Box and Jenkins (1970).

Let $(w_{1,t}, w_{2,t})$, $-\infty < t < \infty$ be a discrete-time bivariate stationary Gaussian time series with mean zero. Suppose that $w_{h,t}$ can be represented as a univariate stationary and invertible ARMA time series of order (p_h, q_h):

$$\phi_h(B)w_{h,t} = \theta_h(B)a_{h,t} \qquad (1.1)$$

* A. Ian McLeod is Assistant Professor, Statistics and Actuarial Science Group, Department of Mathematics, The University of Western Ontario, London, Canada N6A 5B9. The author wishes to thank the editor, an associate editor, and a referee for criticisms and suggestions that significantly improved the presentation of this article. The author also would like to thank M.E. Thompson for reading an earlier version of this article.

where

$$\phi_h(B) = 1 - \phi_{h1}B - \ldots - \phi_{hp_h}B^{p_h} ,$$
$$\theta_h(B) = 1 - \theta_{h1}B - \ldots - \theta_{hq_h}B^{q_h} ,$$

B is the backshift operator ($Bw_{h,t} = w_{h,t-1}$), and $a_{1,t}$ and $a_{2,t}$ are the individual innovation or white-noise series. The innovations $(a_{1,t}, a_{2,t})$ are then a bivariate Gaussian time series with mean zero and autocovariance function

$$\begin{aligned}\gamma_{a_ha_h}(l) &= \langle a_{h,t}a_{h,t+l}\rangle \\ &= \sigma_h^2 , \quad \text{if } l = 0 , \quad h = 1, 2 , \\ &= 0 , \quad \text{if } l \neq 0 , \quad h = 1, 2 , \end{aligned} \qquad (1.2)$$

where $\langle\cdot\rangle$ denotes mathematical expectation and σ_h^2 is the individual innovation variance for the time series $w_{h,t}$. The cross-covariance function of $a_{1,t}$ and $a_{2,t}$ is defined by

$$\gamma_{a_1a_2}(l) = \langle a_{1,t}a_{2,t+l}\rangle , \quad l = 0, \pm1, \ldots , \qquad (1.3)$$

and it is assumed that

$$\sum_{l=-\infty}^{\infty} |l|\,|\gamma_{a_1a_2}(l)| < \infty . \qquad (1.4)$$

Given n observations, $w_{h,t}$, $t = 1, 2, \ldots, n$, from the time series, efficient univariate algorithms to estimate the model parameters $\beta_h = (\phi_{h1}, \ldots, \phi_{hp_h}, \theta_{h1}, \ldots, \theta_{hq_h})$ have been described by Box and Jenkins (1970), Ljung and Box (1976), McLeod (1977b), and other researchers.

1.2 Cross-Correlations in Univariate ARMA Models

A number of authors (see Haugh and Box 1977 and references therein) have advocated the use of the cross-correlation function of $a_{1,t}$ and $a_{2,t}$,

$$\rho_{a_1a_2}(l) = \gamma_{a_1a_2}(l)/(\sigma_1\sigma_2) , \quad l = 0, \pm1, \ldots, \qquad (1.5)$$

for elucidating the relationship between $w_{1,t}$ and $w_{2,t}$. To measure the strength of the relationship between $w_{1,t}$ and $w_{2,t}$, Pierce (1977a,b) has suggested the coefficient

$$R^2 = \sum_{l=-\infty}^{\infty} \rho_{a_1a_2}^2(l) . \qquad (1.6)$$

If $\rho_{a_1a_2}(l) \neq 0$ for some $l > 0$, Pierce and Haugh (1977) showed that $w_{1,t}$ is a useful predictor for $w_{2,t}$. Furthermore, if $\rho_{a_1a_2}(l) = 0$ for all $l < 0$ and $\rho_{a_1a_2}(l) \neq 0$ for some

© Journal of the American Statistical Association
December 1979, Volume 74, Number 368
Theory and Methods Section

$l > 0$, then Haugh (1972a) and Haugh and Box (1977) have shown how a dynamic regression of $w_{2,t}$ on $w_{1,t}$ can be identified by using the innovation cross-correlation function $\rho_{a_1a_2}(\cdot)$.

For any given parameter value, say $\tilde{\beta}_h$, the corresponding estimated innovation series $\tilde{a}_{h,t}$, $t = 1, \ldots, n$ may be directly calculated from (1.1) either by setting $w_{h,t}$ and $a_{h,t}$ for $t \leq 0$ equal to their expected values conditional on $w_{h,1}, \ldots, w_{h,n}$ as described in Box and Jenkins (1970) and Newbold (1974) or more approximately by setting $w_{h,t}$ and $a_{h,t}$ for $t \leq 0$ equal to their unconditional expected value of zero. Also, for any given $\tilde{\beta}_h$, the sample innovation cross-covariance and cross-correlation functions of $\tilde{a}_{1,t}$ and $\tilde{a}_{2,t}$ at lag l are defined by

$$\tilde{c}_{a_1a_2}(l) = n^{-1} \sum_{t=1}^{n-l} \tilde{a}_{1,t}\tilde{a}_{2,t+l} \tag{1.7}$$

and

$$\tilde{r}_{a_1a_2}(l) = \tilde{c}_{a_1a_2}(l)/[(\tilde{c}_{a_1a_1}(0)\tilde{c}_{a_2a_2}(0))]^{\frac{1}{2}} , \tag{1.8}$$

respectively. It is easily shown that the absolute error in $\tilde{r}_{a_1a_2}(l)$ from using either of these methods to calculate $\tilde{a}_{h,t}$, $h = 1, 2$, $t = 1, \ldots, n$ is $O(1/n)$. Hence, if the exact value of the model parameters were known to be β_1 and β_2, then the large-sample variances and covariances of the cross-correlations of the calculated innovations, $a_{1,t}$, $t = 1, \ldots, n$ and $a_{2,t}$, $t = 1, \ldots, n$, can be obtained from a formula of Bartlett (1966, p. 332). This formula yields

$$n \cdot \text{cov}(r_{a_1a_2}(l), r_{a_1a_2}(k)) = \rho_{a_2a_2}(k - l) + \rho_{a_1a_2}(k)\rho_{a_1a_2}(l)\left(\sum_{i=-\infty}^{\infty} \rho_{a_1a_2}{}^2(i) - 3\right) + \sum_{i=-\infty}^{\infty} \rho_{a_1a_2}(l - i)\rho_{a_1a_2}(k + i) . \tag{1.9}$$

Let $\hat{\beta}_h$ be a univariate asymptotically efficient estimate of β_h for $h = 1, 2$ and let $\hat{a}_{h,t}$, $t = 1, \ldots, n$ and $\hat{r}_{a_1a_2}(l)$, $l = 0, \pm 1, \ldots$ be the corresponding residuals and residual cross-correlations. Box and Pierce (1970) obtained the large-sample distribution of the residual autocorrelations in univariate ARMA time series models that can be shown to be equivalent to the large-sample distribution of the residual cross-correlations when $\rho_{a_1a_2}(0) = 1$ and $\rho_{a_1a_2}(l) = 0$, $l \neq 0$ (see Section 3). It follows that the large-sample covariances of the residual cross-correlations are not given by Bartlett's formula in the general case. Nevertheless, Haugh (1972a,b, 1976) showed that if the series $w_{1,t}$ and $w_{2,t}$ are independent (so $\rho_{a_1a_2}(l) = 0$, $l = 0, \pm 1, \ldots$), then the large-sample distribution of the residual cross-correlations is jointly normal with covariance matrix determined by

$$\begin{aligned} \text{cov}(\hat{r}_{a_1a_2}(l), \hat{r}_{a_1a_2}(k)) &= 1/n , \quad \text{if} \quad l = k , \\ &= 0 , \quad \text{if} \quad l \neq k . \end{aligned} \tag{1.10}$$

Several authors (Haugh and Box 1977; Pierce and Haugh 1977; Pierce 1977a) have remarked that it would be useful to know the distribution of the residual cross-correlations in the general case. This is derived in Section 2.

2. GENERAL RESULT

In this section, the asymptotic joint distribution of the residual cross-correlations in model (1.1) is derived. For any fixed $M \geq 0$, let

$$\hat{\mathbf{r}} = (\hat{r}_{a_1a_2}(-1), \ldots, \hat{r}_{a_1a_2}(-M) , \hat{r}_{a_1a_2}(0) , \hat{r}_{a_1a_2}(1), \ldots, \hat{r}_{a_1a_2}(M)) , \tag{2.1}$$

and let

$$\boldsymbol{\rho} = (\rho_{a_1a_2}(-1), \ldots, \rho_{a_1a_2}(-M), \rho_{a_1a_2}(0), \rho_{a_1a_2}(1), \ldots, \rho_{a_1a_2}(M)) . \tag{2.2}$$

Let $\mathbf{r}$ and $\hat{\mathbf{r}}$ denote the vector $\tilde{\mathbf{r}}$ when $\tilde{\beta} = \beta$ and $\tilde{\beta} = \hat{\beta}$, respectively. Thus, $\mathbf{r}$ and $\hat{\mathbf{r}}$ are vectors of innovation and residual cross-correlations. The derivation of the asymptotic joint distribution of $\hat{\mathbf{r}}$ is based on the use of a Taylor series linearization of $\hat{\mathbf{r}}$ as a function of $(\hat{\beta}_1, \hat{\beta}_2, \mathbf{r})$ and the asymptotic joint distribution of $(\hat{\beta}_1, \hat{\beta}_2, \mathbf{r})$.

Lemma 1: The distribution of $\hat{\mathbf{r}}$ does not depend on σ_1 or σ_2.

Remark 1: It follows from Lemma 1 that, without loss of generality, it can be assumed that $\gamma_{a_1a_1}(0) = \gamma_{a_2a_2}(0) = 1$.

The following lemma is useful for simplifying double summations of theoretical cross-covariances.

Lemma 2:

$$n^{-1} \sum_{k=1}^{n} \sum_{l=1}^{n} \gamma_{a_1a_2}(k - l) = \sum_{k=-\infty}^{\infty} \gamma_{a_1a_2}(k) + O(1/n) . \tag{2.3}$$

Proof: This follows from the assumption stated in (1.4).

Lemma 3:

$$r_{a_1a_2}(l) - \rho_{a_1a_2}(l) = c_{a_1a_2}(l) - \tfrac{1}{2}\gamma_{a_1a_2}(l)(c_{a_1a_1}(0) + c_{a_2a_2}(0)) + O_p(1/n) . \tag{2.4}$$

Proof: This follows from the Taylor series expansion of $\hat{r}_{a_1a_2}(l)$ as a function of $(\hat{c}_{a_1a_2}(l), \hat{c}_{a_1a_1}(0), \hat{c}_{a_2a_2}(0))$ about

$$(\gamma_{a_1a_2}(l), \gamma_{a_1a_1}(0), \gamma_{a_2a_2}(0))$$

and evaluated at

$$(c_{a_1a_2}(l), c_{a_1a_1}(0), c_{a_2a_2}(0)) .$$

Lemma 4: Let $\hat{\beta}_h$ be an asymptotically efficient estimate of β_h in the univariate ARMA model. Then,

$$\hat{\beta}_h - \beta_h = \mathbf{I}_h^{-1}\mathbf{S}_h + O_p(1/n) \tag{2.5}$$

where $\mathbf{I}_h$ is the large-sample information matrix per observation given by

$$\mathbf{I}_h = \begin{bmatrix} \gamma_{v_h v_h}(i - j) & \gamma_{v_h u_h}(i - j) \\ \gamma_{u_h v_h}(i - j) & \gamma_{u_h u_h}(i - j) \end{bmatrix} \begin{matrix} p_h \\ q_h \end{matrix} \tag{2.6}$$
$$\qquad p_h \qquad\qquad q_h$$

where the (i, j) entry in each partitioned matrix is indicated and the auxiliary time series $v_{h,t}$ and $u_{h,t}$ are defined by

$$\phi_h(B)v_{h,t} = -a_{h,t} \tag{2.7}$$

and

$$\theta_h(B)u_{h,t} = a_{h,t} . \tag{2.8}$$

$\mathbf{S}_h = (S_{h,1}, \ldots, S_{h,p_h+q_h})$ is the score function

$$S_{h,i} = -n^{-1} \sum_{t=1}^{n} a_{h,t} v_{h,t-i} \; , \qquad \text{if} \quad i = 1, \ldots, p_h \; ;$$

$$= -n^{-1} \sum_{t=1}^{n} a_{h,t} u_{h,t+p_h-i} \; , \qquad \text{if} \quad i = p_h + 1, \ldots, p_h + q_h \; . \quad (2.9)$$

Proof: The likelihood function of $\boldsymbol{\beta}_h$ in the univariate ARMA model is

$$\log L_h = -\tfrac{1}{2} \sum_{t=1}^{n} \dot{a}_{h,t}^2 + m_h + O(r^n) \; , \qquad (2.10)$$

where m_h is a rational function of the elements of $\boldsymbol{\beta}_h$ that does not depend on n and $0 < r < 1$ (McLeod 1977b). Also, it can be shown (Box and Jenkins 1970, p. 237) that

$$\partial \dot{a}_{h,t} / \partial \phi_{hi} = v_{h,t-i} \qquad (2.11)$$

and

$$\partial \dot{a}_{h,t} / \partial \theta_{hi} = u_{h,t-i} \; , \qquad (2.12)$$

where the partial derivative of $\dot{a}_{h,t}$ with respect to ϕ_{hi} evaluated at ϕ_{hi} is denoted by $\partial \dot{a}_{ht}/\partial \phi_{hi}$ (and similar notation is used throughout the article). It follows that

$$\partial \log L_h / \partial \beta_{h,i} = n S_{h,i} + O(1) \; . \qquad (2.13)$$

Also, it can be shown that

$$\partial \log L_h / (\partial \boldsymbol{\beta}_h \partial \boldsymbol{\beta}_h^T) = -n \mathbf{I}_h + O(\sqrt{n}) \; . \qquad (2.14)$$

The lemma now follows from the Taylor series expansion of $\partial \log L_h / \partial \boldsymbol{\beta}_h$ about $\hat{\boldsymbol{\beta}}_h$ evaluated at $\boldsymbol{\beta}_h$.

Remark 2: Lemmas 3 and 4 present linearizations that are useful for handling the asymptotics of expressions involving $\mathbf{r}$ and $\hat{\boldsymbol{\beta}}_h$.

Lemma 5: The asymptotic joint distribution of $\sqrt{n}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$, $\sqrt{n}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2)$ is normal with mean vector zero and covariance matrix

$$\mathbf{V} = \left[\begin{array}{c|c} \mathbf{I}_1^{-1} & \mathbf{I}_1^{-1}\mathbf{A}\mathbf{I}_2^{-1} \\ \hline \mathbf{I}_2^{-1}\mathbf{A}^T\mathbf{I}_1^{-1} & \mathbf{I}_2^{-1} \end{array}\right] \begin{array}{l} p_1 + q_1 \\ p_2 + q_2 \end{array} \qquad (2.15)$$
$$\qquad p_1 + q_1 \qquad p_2 + q_2$$

where

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}^{(v_1,v_2)} & \mathbf{A}^{(v_1,u_2)} \\ \hline \mathbf{A}^{(u_1,v_2)} & \mathbf{A}^{(u_1,u_2)} \end{array}\right] \begin{array}{l} p_1 \\ q_1 \end{array} \qquad (2.16)$$
$$\qquad p_2 \qquad q_2$$

where the (i, j) element of the submatrix $\mathbf{A}^{(c,d)}$ is

$$A_{ij}^{(c,d)} = \sum_{k=-\infty}^{\infty} \gamma_{a_1a_2}(k)\gamma_{cd}(k + i - j) + \gamma_{a_1d}(k - j)\gamma_{ca_2}(k + i) \; . \qquad (2.17)$$

Proof: The matrix $\mathbf{A}$ is obtained by straightforward calculation using Lemmas 2 and 4. The joint asymptotic normality of the estimates follows from Lemma 4 because any linear function of both $\mathbf{S}_1$ and $\mathbf{S}_2$ is the average of a series of martingale differences and so by the martingale central limit theorem (Billingsley 1961) is asymptotically normal.

Lemma 6: The asymptotic joint distribution of $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\mathbf{r}})$ is normal with mean vector $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\rho})$ and covariance matrix

$$\frac{1}{n}\left[\begin{array}{c|c} \mathbf{V} & -\boldsymbol{\Delta} \\ \hline -\boldsymbol{\Delta}^T & \mathbf{E} \end{array}\right] \begin{array}{l} p_1 + q_1 + p_2 + q_2 \\ 2M + 1 \end{array} \qquad (2.18)$$
$$\qquad p_1 + q_1 + p_2 + q_2 \qquad 2M + 1$$

where $\mathbf{V}$ is defined in Lemma 5, $\mathbf{E}/n$ is the large-sample c... riance matrix of $\mathbf{r}$ that is determined directly from Bartlett's formula (1.9), and the jth column, $j = 1, \ldots, 2M + 1$, of $\boldsymbol{\Delta}$ is $(\boldsymbol{\delta}_j^{(1)}, \boldsymbol{\delta}_j^{(2)})^T$ where

$$\boldsymbol{\delta}_j^{(h)} = - \text{[illegible]} = \mathbf{I}_h^{-1}(\mathbf{f}^{(v_h)}, \mathbf{f}^{(u_h)})^T \qquad (2.19)$$

where

$$h = 1, 2 \; ,$$
$$k = -j \; , \qquad \text{if} \quad j = 1, \ldots, M \; ,$$
$$= j - M - 1 \; , \quad \text{if} \quad j = M + 1, \ldots, 2M + 1 \; ,$$

$$f_i^{(c)} = \gamma_{ca_2}(k + i) + \sum_{l=-\infty}^{\infty} \gamma_{a_1a_2}(l)[\gamma_{a_1c}(k - i - l) - \gamma_{a_1a_2}(k)\gamma_{ca_2}(l + i)] \; , \quad \text{if} \quad c = v_1, u_1 \; ,$$

$$= \gamma_{a_1c}(k - i) + \sum_{l=-\infty}^{\infty} \gamma_{a_1a_2}(l)[\gamma_{a_2c}(l - k - i) - \gamma_{a_1a_2}(k)\gamma_{ca_1}(l + i)] \; , \quad \text{if} \quad c = v_2, u_2 \; . \qquad (2.20)$$

Proof: The computation of $\mathbf{f}^{(c)}$ follows directly from Lemmas 2, 3, and 4. The asymptotic joint normality is proved, as in Lemma 5.

Theorem: The asymptotic distribution of $\hat{\mathbf{r}}$ is normal with mean vector $\boldsymbol{\rho}$ and covariance matrix

$$\operatorname{var}(\hat{\mathbf{r}}) = (\mathbf{E} + \mathbf{X}\mathbf{V}\mathbf{X}^T - \mathbf{X}\boldsymbol{\Delta} - \boldsymbol{\Delta}^T\mathbf{X}^T)/n \qquad (2.21)$$

where $\mathbf{E}$, $\mathbf{V}$, and $\boldsymbol{\Delta}$ are defined in Lemma 6 and

$$\mathbf{X} = \left[\begin{array}{c|c|c|c} r_{v_1a_2}(-i, j) & r_{u_1a_2}(-i, j) & r_{v_2a_1}(i, j) & r_{u_2a_1}(i, j) \\ \hline r_{v_1a_2}(0, j) & r_{u_1a_2}(0, j) & r_{v_2a_1}(0, j) & r_{u_2a_1}(0, j) \\ \hline r_{v_1a_2}(i, j) & r_{u_1a_2}(i, j) & r_{v_2a_1}(-i, j) & r_{u_2a_1}(-i, j) \end{array}\right] \begin{array}{l} M \\ 1 \\ M \end{array} \qquad (2.22)$$
$$\qquad p_1 \qquad q_1 \qquad p_2 \qquad q_1$$

where the (i, j) element in each partitioned matrix is indicated;

$$r_{ca_k}(l, j) = \gamma_{ca_k}(l + j) - \tfrac{1}{2}\gamma_{a_ka_k}(l)\gamma_{ca_k}(j) \qquad (2.23)$$

where

$$c = u_h, v_h \; , \quad h, k = 1, 2 \; , \quad -M \le |l| \le M \; , \quad j = 1, 2, \ldots .$$

Proof: Let $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2)$, and let $\hat{\beta}_i$ denote the ith element of $\hat{\boldsymbol{\beta}}$. Then by expanding $\hat{r}_{a_1a_2}(i)$ in a Taylor series about β and evaluating $\hat{\beta} = \hat{\beta}$ it follows that

$$\hat{r}_{a_1a_2}(i) = r_{a_1a_2}(i) + \sum_{j=1}^{p_1+q_1+p_2+q_2} (\hat{\beta}_j - \beta_j)\partial \hat{r}_{a_1a_2}(i)/\partial\beta_j + 0_p(1/n) . \quad (2.24)$$

Consider the $M \times p_1$ submatrix of X corresponding to $i = 1, \ldots, M$ and $j = 1, \ldots, p_1$. The (i, j) element of this submatrix is

$$X_{ij} = \gamma_{v_1a_2}(j-i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\gamma_{v_1a_1}(j) . \quad (2.25)$$

It follows directly from (2.11) that

$$\frac{\partial \hat{r}_{a_1a_2}(-i)}{\partial\phi_j} = \frac{c_{v_1a_2}(j-i)}{[(c_{a_1a_1}(0)c_{a_2a_2}(0))]^{\frac{1}{2}}} - \frac{1}{2} r_{a_1a_2}(-i)\frac{c_{v_1a_1}(j)}{c_{a_1a_1}(0)} . \quad (2.26)$$

To determine the large-sample mean and variance of

$$r_{a_1a_2}(-i)c_{v_1a_1}(j)/c_{a_1a_1}(0) ,$$

note that by Lemma 3

$$r_{a_1a_2}(-i) = \gamma_{a_1a_2}(-i) + c_{a_1a_2}(-i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)[c_{a_1a_1}(0) + c_{a_2a_2}(0)] + 0_p(1/n) , \quad (2.27)$$

and also by a Taylor series expansion it can be shown that

$$c_{v_1a_1}(j)/c_{a_1a_1}(0) = \gamma_{v_1a_1}(j) + c_{v_1a_1}(j) - \gamma_{v_1a_1}(j)c_{a_1a_1}(0) + 0_p(1/n) . \quad (2.28)$$

Because the variances and covariances of the sample cross-covariances in (2.27) and (2.28) are $0(1/n)$, it follows that

$$\operatorname{var}(r_{a_1a_2}(-i)c_{v_1a_1}(j)/c_{a_1a_1}(0)) = 0(1/n) \quad (2.29)$$

and

$$\langle r_{a_1a_2}(-i)c_{v_1a_1}(j)/c_{a_1a_1}(0)\rangle = \gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_1}(j) + 0(1/n) . \quad (2.30)$$

Similarly, it can be shown that

$$c_{v_1a_2}(j-i)/([c_{a_1a_1}(0)c_{a_2a_2}(0)])^{\frac{1}{2}}$$

has variance $0(1/n)$ and mean equal to $\gamma_{v_1a_2}(j-i) + 0(1/n)$. Hence, it follows that

$$\operatorname{var}(\partial\hat{r}_{a_1a_2}(-i)/\partial\phi_j) = 0(1/n) \quad (2.31)$$

and that

$$\langle \partial \hat{r}_{a_1a_2}(-i)/\partial\phi_j\rangle = \gamma_{v_1a_2}(j-i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_1}(j) + 0(1/n) . \quad (2.32)$$

It follows from Chebyshev's inequality that

$$\partial\hat{r}_{a_1a_2}(-i)/\partial\phi_j = \gamma_{v_1a_2}(j-i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_1}(j) + 0_p(1/\sqrt{n}) . \quad (2.33)$$

In general, it can be shown that

$$\partial\hat{r}_{a_1a_2}(l)/\partial\beta_j = X_{ij} + 0_p(1/\sqrt{n}) \quad (2.34)$$

where $-M \le l \le M$, $j = 1, \ldots, p_1 + q_1 + p_2 + q_2$,

$$i = |l| , \quad \text{if } l < 0 ,$$
$$= M + l + 1 , \quad \text{if } l \ge 0 ,$$

and X_{ij} is the (i, j) element of X.

Because $(\hat{\beta}_j - \beta_j)$ is $0_p(1/\sqrt{n})$ it follows from the theorem of Mann and Wald (1943, Corollary 1) that

$$\hat{\mathbf{r}} = \mathbf{r} + \mathbf{X}(\hat{\beta} - \beta) + 0_p(1/n) . \quad (2.35)$$

The theorem now follows directly from Lemma 6 and a theorem given by Rao (1974, (2c.4.12)).

Remark 3: If the assumption of joint normality of $a_{t,h}$; $t = 1, \ldots, n$; $h = 1, 2$ is invalid, (2.21) will involve fourth-order cumulants.

Remark 4: It is not difficult to show that the theorem also applies to the case of two time series with nonzero means if the series are corrected for their sample means.

3. AN APPLICATION

The general result of Section 2 can be simplified in certain special cases. Many economic time series have the property that the largest residual cross-correlation is at lag zero (Pierce 1977a). In this section, the distribution of the residual cross-correlations when only the lag-zero innovation cross-correlation is nonzero is obtained and is used to derive a test for lagged relationships between economic time series.

3.1 Instantaneous Causality Only

Suppose the time series $w_{1,t}$ and $w_{2,t}$ are generated by the model (1.1) and the cross-correlation function between $a_{1,t}$ and $a_{2,t}$ is given by

$$\rho_{a_1a_2}(l) = \rho , \quad \text{if } l = 0 , \qquad (3.1)$$
$$= 0 , \quad \text{if } l \ne 0 .$$

Then the relationship between $w_{1,t}$ and $w_{2,t}$ may be said to be one of instantaneous causality only (Pierce and Haugh 1977, 1979).[1]

Bartlett's formula (1.9) for the large-sample variances and covariances of the sample cross-correlations of $a_{1,t}$ and $a_{2,t}$ yields

$$\operatorname{var}(r_{a_1a_2}(l)) = (1-\rho^2)^2/n , \quad \text{if } l = 0 , \qquad (3.2)$$
$$= 1/n , \quad \text{if } l \ne 0 ,$$

and

$$\operatorname{cov}(r_{a_1a_2}(l), r_{a_1a_2}(k)) = 0 , \quad \text{if } l \ne k ,\ l \ne -k \qquad (3.3)$$
$$= \rho^2/n , \quad \text{if } l = -k ,\ l \ne 0 .$$

Let

$$\hat{\mathbf{r}}^{(1)} = (\hat{r}_{a_1a_2}(-1), \ldots, \hat{r}_{a_1a_2}(-M)) \quad (3.4)$$

and

$$\hat{\mathbf{r}}^{(2)} = (\hat{r}_{a_1a_2}(1), \ldots, \hat{r}_{a_1a_2}(M)) . \quad (3.5)$$

[1] Empirical methods for detecting causality relationships between time series have been discussed by Granger (1969) and Pierce and Haugh (1977).

Then it follows from the theorem in Section 2 that

$$\text{var}(\hat{\mathbf{r}}^{(h)}) = \mathbf{P}_h/n \quad (3.6)$$

where

$$\mathbf{P}_h = 1_M - \rho^2 \mathbf{X}_h \mathbf{I}_h^{-1} \mathbf{X}_h^T , \quad h = 1, 2 , \quad (3.7)$$

where 1_M is the $M \times M$ identity matrix, $\mathbf{I}_h$ is defined in (2.7), and

$$\mathbf{X}_h = (-\pi_{h,i-j} | \psi_{h,i-j}) M \quad (3.8)$$
$$\qquad p_h \qquad q_h$$

where the (i, j) element in each partitioned matrix is indicated, $\pi_{h,l} = -\gamma_{v_h a_h}(-l)$ and $\psi_{h,l} = \gamma_{u_h a_h}(-l)$. The coefficients $\pi_{h,l}$ and $\psi_{h,l}$ are easily calculated recursively (Box and Jenkins 1970, pp. 132–134) by using the identities $1/\phi_h(B) = \sum \pi_{h,k} B^k$ and $1/\theta_h(B) = \sum \psi_{h,k} B^k$. The elements of the information matrix $\mathbf{I}_h$ may be calculated by solving a set of linear equations as in McLeod (1975, 1977a). Also, from the theorem of Section 2,

$$\text{cov}(\hat{\mathbf{r}}^{(h)}, \hat{r}_{a_1 a_2}(0)) = 0 , \quad h = 1, 2 \quad (3.9)$$

and

$$\text{cov}(\hat{\mathbf{r}}^{(1)}, \hat{\mathbf{r}}^{(2)}) = (\rho^2 \mathbf{J} + 2 1_M - \mathbf{P}_1 - \mathbf{P}_2 + \rho^2 \mathbf{X}_1 \mathbf{I}_1^{-1} \mathbf{A} \mathbf{I}_2^{-1} \mathbf{X}_2^T)/n \quad (3.10)$$

where $\mathbf{J}$ is the $M \times M$ matrix with 1's on the diagonal at right angles to the main diagonal and 0's elsewhere and

$$\mathbf{A} = \begin{bmatrix} \gamma_{v_1 v_2}(i-j) & \gamma_{v_1 u_2}(i-j) \\ \gamma_{u_1 v_2}(i-j) & \gamma_{u_1 u_2}(i-j) \end{bmatrix} \begin{matrix} p_1 \\ q_1 \end{matrix} . \quad (3.11)$$
$$\qquad p_2 \qquad q_2$$

Finally, the large-sample variance of the estimate $\hat{\rho} = \hat{r}_{a_1 a_2}(0)$, of ρ is

$$\text{var}(\hat{\rho}) = (1 - \rho^2)^2/n . \quad (3.12)$$

If $\rho = 0$, the asymptotic variances of the residual cross-correlations are all equal to $1/n$, but from (3.7) it can be shown that when ρ^2 is close to one the asymptotic variances of $\hat{r}_{a_1 a_2}(l)$ may be significantly less than $1/n$. In fact, when $\rho = 1$, the residual cross-correlations and residual autocorrelations have the same asymptotic distributions. McLeod (1977a, 1978) has shown that, for large n, any fixed $M \geq 1$ and $h = 1, 2$ the covariance matrix of the residual autocorrelations

$$(\hat{r}_{a_h a_h}(1), \ldots, \hat{r}_{a_h a_h}(M)) \quad (3.13)$$

is given by

$$(1_M - \mathbf{X}_h \mathbf{I}_h^{-1} \mathbf{X}_h^T)/n . \quad (3.14)$$

It also can be shown (McLeod 1977a) that the covariance matrix (3.14) is exactly equivalent to that derived by Box and Pierce (1970) (also see Durbin 1970).

3.2 Test for Lagged Relationships

To test the null hypotheses

$$H_0^{(1)}: \rho_{a_1 a_2}(-1) = \ldots = \rho_{a_1 a_2}(-M) = 0$$

or

$$H_0^{(2)}: \rho_{a_1 a_2}(1) = \ldots = \rho_{a_1 a_2}(M) = 0$$

against the simple negation of $H_0^{(1)}$ or $H_0^{(2)}$, respectively, when $\rho \neq 0$, the following test statistic is suggested:

$$\hat{Q}_M^{(h)} = n(\hat{\mathbf{r}}^{(h)})^T \hat{\mathbf{P}}_h^{-1} \hat{\mathbf{r}}^{(h)} , \quad (3.15)$$

where $\hat{\mathbf{P}}_h$ denotes the matrix $\mathbf{P}_h$ in (3.7) calculated by using $\hat{\beta}_h$ rather than β_h. If both $H_0^{(1)}$ and $H_0^{(2)}$ are true, $\hat{Q}_M^{(h)}$ will be asymptotically χ^2-distributed on M degrees of freedom and large values of $\hat{Q}_M^{(h)}$ will provide evidence against $H_0^{(h)}$. This test of $H_0^{(h)}$ may be compared with the test suggested by Pierce (1977a) that is based on the statistic

$$Q_M^{(h)} = n(\hat{\mathbf{r}}^{(h)})^T \hat{\mathbf{r}}^{(h)} . \quad (3.16)$$

If $\rho = 0$ and $H_0^{(1)}$ and $H_0^{(2)}$ are both true, then $Q_M^{(h)}$ is also $\chi^2(M)$ for large n. The example given in the following paragraph shows that the test based on $\hat{Q}_M^{(h)}$ may be more sensitive than that using $Q_M^{(h)}$ when $\rho \neq 0$.

From the results of Davies, Triggs, and Newbold (1977) and Ljung and Box (1978) on the portmanteau significance test of Box and Pierce (1970), it may be expected that both these tests using $Q_M^{(h)}$ and $\hat{Q}_M^{(h)}$ may considerably underestimate the true significance level if M is fairly large. In the case when $\rho = 0$, Haugh (1976) provided simulation evidence that

$$\text{var}(\hat{r}_{a_1 a_2}(l)) \doteq (n - l)/n^2 \quad (3.17)$$

and suggested a modified test using this result. A general alternative approach would be to use shrinked residual cross-correlation estimates, such as

$$\tilde{r}_{a_1 a_2}(l) = \hat{r}_{a_1 a_2}(l)([(n - |l|)/n])^{\frac{1}{2}} . \quad (3.18)$$

Estimated standard deviations of the residual cross-correlations may be obtained by using estimated values of ρ, β_1, and β_2 in (3.7) and (3.12).

3.3 Example

Haugh (1976, p. 383) found that the first differences of two quarterly interest rate time series could be modeled by

$$w_{1,t} = .069 + (1 + .55B)\hat{a}_{1,t} \quad (3.19)$$

and

$$(1 - .76B + .39B^2) w_{2,t} = \hat{a}_{2,t} , \quad (3.20)$$

where $w_{h,t}$ and $\hat{a}_{h,t}$ are, respectively, the first differences and estimated innovations of series h for $h = 1, 2$. Also, in this example, $n = 71$ and $\hat{\rho} = .64$. The residual cross-correlations and their estimated standard errors calculated from (3.7) and (3.12) are shown in Table 1.

Note, that in Table 1, $\hat{r}_{a_1 a_2}(1)$ is significant at 5 percent, although it is not significant at 5 percent when compared with the benchmark standard deviation of $n^{-\frac{1}{2}} = .119$. The statistics for testing $H_0^{(1)}$ and $H_0^{(2)}$ have the following values:

h	1	2
$\hat{Q}_4^{(h)}$	1.34	14.06
$Q_4^{(h)}$	1.28	8.85

1. Residual Cross-Correlations and Estimated Standard Deviations

l	−4	−3	−2	−1	0	1	2	3	4
$\hat{r}_{a_1 a_2}(l)$	.04	.10	.00	−.08	.64	.20	−.13	−.26	.01
Estimated standard deviation	.118	.117	.113	.100	.070	.096	.102	.109	.117

Thus, $H_0^{(1)}$ is not significant at 5 percent for both tests. $H_0^{(2)}$ is significant at 5 percent if the test statistic $\hat{Q}_4^{(2)}$ is used, and it is not significant at 5 percent if the less-sensitive test based on $Q_4^{(2)}$ is used.

4. SIMULATION STUDY

A simulation study was done to examine the accuracy of the asymptotic variances and covariances of the residual cross-correlations in the case of two first-order autoregressions,

$$(1 - \phi_1 B) w_{1,t} = a_{1,t} \tag{4.1}$$

and

$$(1 - \phi_2 B) w_{2,t} = a_{2,t} \ , \tag{4.2}$$

where $t = 1, \ldots, n$, $\rho_{a_1 a_2}(0) = \rho$, $\rho_{a_1 a_2}(l) = 0$ if $l \neq 0$, and $a_{1,t}$ and $a_{2,t}$ are Gaussian white noise with unit variance. The theoretical large-sample covariances of the residual cross-correlations are shown in Table 2.

A total of 225 models corresponding to the parameter settings $\phi_1, \phi_2 = 0, \pm.5, \pm.9$, $\rho = .3, .6, .9$ and $n = 50$, 200, 400 were included, and for each model 1,000 simulations were done. A multiplicative congruential random-number generator with modulus 2^{35} and multiplier 5^{15} (recommended by Conveyou and MacPherson 1967) was used in conjunction with the method of Marsaglia and Bray (1964) to generate independent normal pseudo-random numbers. The method of generating initial values of the time series can be quite important (McLeod and Hipel 1978). The following technique was used:

1. Initial values in the time series were generated by using the covariance matrix of $(w_{1,1}, w_{2,1})$.
2. A linear transformation was made on successive pairs of independent normal random numbers to obtain the simulated bivariate white-noise series $(a_{1,t}, a_{2,t})$, $t = 2, 3, \ldots, n$.
3. The remaining values of the time series were calculated recursively by using (4.1) and (4.2).

2. Asymptotic Covariances of the Residual Cross-Correlations in Two First-Order Autoregressions With Instantaneous Causality Only

l,k	$n \cdot \mathrm{cov}(\hat{r}_{a_1 a_2}(l), \hat{r}_{a_1 a_2}(k))$
$l = 0$, any k	$\delta_{0,k}(1 - \rho^2)^2$
$l,k < 0$	$\delta_{l,k} - \rho^2 \phi_1^{\lvert l+k \rvert - 2}(1 - \phi_1^2)$
$l,k > 0$	$\delta_{l,k} - \rho^2 \phi_2^{l+k-2}(1 - \phi_2^2)$
$l < 0, k > 0$	$\rho^2 \delta_{-l,k} - \rho^2 \phi_1^{k-l-2}(1 - \phi_1^2) - \rho^2 \phi_2^{k-l-2}(1 - \phi_2^2)$
	$+ \rho^4 \phi_1^{\lvert l \rvert - 1} \phi_2^{k-1}(1 - \phi_1^2)(1 - \phi_2^2)/(1 - \phi_1 \phi_2)$

NOTE: $\delta_{l,k} = 1$, if $l = k$
$= 0$, if $l \neq k$.

The parameters ϕ_1 and ϕ_2 were estimated by using the lag-one sample autocorrelation coefficients and the residual cross-correlation vector, $\hat{\mathbf{r}}$, defined in (2.1) with $M = 2$ was calculated. For each model, the sample covariance matrix of $\hat{\mathbf{r}}$ was calculated by correcting the sample second moment of $\hat{\mathbf{r}}$ by its sample mean.

Let $\tilde{\mathbf{C}}$ denote the sample estimate, based on 1,000 simulations, of the covariance matrix of the 5×1 vector $\hat{\mathbf{r}}$, and let $\mathbf{C}$ be the corresponding asymptotic covariance matrix of $\hat{\mathbf{r}}$ determined from Table 2. Then, assuming that $\mathbf{C}$ provides a good finite-sample approximation, $\tilde{\mathbf{C}}$ has an asymptotic normal distribution (Anderson 1958, p. 75) with mean $\mathbf{C}$ and covariance matrix determined by

$$\mathrm{cov}(\tilde{C}_{ij}, \tilde{C}_{kl}) = (C_{ik}C_{jl} + C_{il}C_{jk})/1{,}000 \ , \tag{4.3}$$

where $\tilde{C}_{ij}$ and C_{ij} denote the (i, j) element of $\tilde{\mathbf{C}}$ and $\mathbf{C}$. Let $\tilde{\mathbf{c}}$ and $\mathbf{c}$ be the 15×1 vectors corresponding to the elements of $\tilde{\mathbf{C}}$ and $\mathbf{C}$ on or above the main diagonal taken in lexicographical order and let $\mathbf{\Xi}$ be the asymptotic covariance matrix of $\tilde{\mathbf{c}}$ determined by (4.3) and Table 2. If the covariance matrix of $\hat{\mathbf{r}}$ is well approximated by $\mathbf{C}$, T^2 should be $\chi^2(15)$, where

$$T^2 = (\tilde{\mathbf{c}} - \mathbf{c})^T \mathbf{\Xi}^{-1} (\tilde{\mathbf{c}} - \mathbf{c}) \ . \tag{4.4}$$

Large values of T^2 will tend to indicate that the asymptotic covariances given in Table 2 are not valid. The test statistic T^2 was evaluated by using double-precision arithmetic and the Cholesky decomposition method (Maindonald 1976). For each of the 25 models corresponding to a fixed value of n and ρ, the number of values of T^2 significant at the 5 and 1 percent levels and the mean T^2 value were calculated and the results are shown

3. Summary of T^2 Tests

n	Summary Result	ρ .3	.6	.9
50	Mean	24.72	37.38	621.63
	>25.00*	9	20	25
	>35.58	5	16	25
200	Mean	15.12	19.43	57.70
	>25.00	3	5	24
	>30.58	1	2	20
400	Mean	14.76	16.70	25.80
	>25.00	2	2	11
	>30.58	0	2	7

* The 5% and 1% critical values are, respectively, 25.00 and 30.58.
NOTE: For each value of n and ρ there are 25 models. Each T^2 is calculated from 1,000 simulations of one model.

in Table 3. The main conclusion to be drawn from these results is that the accuracy of the asymptotic approximation depends not only on n but also on ρ. For example, if $|\rho| \leq .6$, the asymptotic approximation appears to be quite good if $n \geq 400$ (because the probability of four or more rejections at the 5 percent level is .24), but for larger values of $|\rho|$, larger values of n are required.

[*Received March 1977. Revised April 1979.*]

REFERENCES

Anderson, T.W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons.

Bartlett, M.S. (1966), *Stochastic Processes* (2nd ed.), Cambridge, England: Cambridge University Press.

Billingsley, Patrick (1961), "The Lindeberg-Levy Theorem for Martingales," *Proceedings of the American Mathematical Society*, 12, 788–792.

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

———, and Pierce, David A. (1970), "Distribution of the Residual Autocorrelations in Autoregressive-Integrated-Moving Average Time Series Models," *Journal of the American Statistical Association*, 65, 1509–1526.

Conveyou, R.R., and MacPherson, R.D. (1967), "Fourier Analysis of Uniform Random Generators," *Journal of the Association of Computing Machinery*, 14, 100–119.

Davies, N., Triggs, C.M., and Newbold, P. (1977), "Significance Levels of the Box-Pierce Portmanteau Statistic in Finite Samples," *Biometrika*, 64, 517–522.

Durbin, James (1970), "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors Are Lagged Dependent Variables," *Econometrica*, 38, 410–421.

Fisher, R.A. (1921), "Contribution to the Discussion of the Paper 'On the Time-Correlation Problem,' by G.U. Yule," *Journal of the Royal Statistical Society*, 84, 534–536.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–438.

Haugh, Larry D. (1972a), "The Identification of Time Series Interrelationships With Special Reference to Dynamic Regression," unpublished PhD thesis, Dept. of Statistics, University of Wisconsin, Madison.

——— (1972b), "Checking Time Series Interrelationships and Identifying Dynamic Regression Models for Short-Term Forecasting," *Proceedings of the American Statistical Association*, Business and Economic Statistics Section, 325–330.

——— (1976), "Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross-Correlation Approach," *Journal of the American Statistical Association*, 71, 378–385.

———, and Box, George E.P. (1977), "Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series," *Journal of the American Statistical Association*, 72, 121–130.

Ljung, G.M., and Box, G.E.P. (1976), "Maximum Likelihood Estimation in the Autoregressive Moving Average Model," Technical Report No. 476, Dept. of Statistics, University of Wisconsin, Madison.

——— (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

Maindonald, J.H. (1976), "Least Squares Programs—A Second Look," *The American Statistician*, 30, 202–203.

Mann, Henry B., and Wald, Abraham (1943), "On Stochastic Limit and Order Relationships," *Annals of Mathematical Statistics*, 14, 217–226.

Marsaglia, G., and Bray, T.A. (1964), "A Convenient Method for Generating Normal Variables," *Society for Industrial and Applied Mathematics Review*, 6, 260–264.

McLeod, A. Ian (1975), "Derivation of the Theoretical Autocovariance Function of Autoregressive-Moving Average Time Series," *Applied Statistics*, 24, 255–256.

——— (1977a), "Topics in Time Series and Intervention Analysis," unpublished PhD thesis, Dept. of Statistics, University of Waterloo.

——— (1977b), "Improved Box-Jenkins Estimators," *Biometrika*, 64, 531–534.

——— (1978), "On the Distribution of Residual Autocorrelations in Box-Jenkins Models," *Journal of the Royal Statistical Society*, 40, 296–302.

———, and Hipel, Keith W. (1978), "Simulation Procedures for Box-Jenkins Models," *Water Resources Research*, 14, 969–975.

Newbold, Paul (1974), "The Exact Likelihood Function for a Mixed Autoregressive-Moving Average Process," *Biometrika*, 61, 423–426.

Pierce, David A. (1977a), "Relationships—and Lack Thereof—Between Economic Time Series, With Special Reference to Money and Interest Rates," *Journal of the American Statistical Association*, 72, 11–21.

——— (1977b), "R^2-Measures for Time Series," Special Studies Paper, Federal Reserve Board, Washington, D.C.

———, and Haugh, Larry D. (1977), "Causality in Temporal Systems: Characterizations and a Survey," *Journal of Econometrics*, 5, 265–293.

——— (1979), "The Characterization of Instantaneous Causality: A Comment," *Journal of Econometrics*, 10, 257–259.

Rao, C.R. (1974), *Linear Statistical Inference* (2nd ed.), New York: John Wiley & Sons.

An ARIMA-Model-Based Approach to Seasonal Adjustment

S. C. HILLMER and G. C. TIAO*

This article proposes a model-based procedure to decompose a time series uniquely into mutually independent additive seasonal, trend, and irregular noise components. The series is assumed to follow the Gaussian ARIMA model. Properties of the procedure are discussed and an actual example is given.

KEY WORDS: ARIMA model; Seasonal adjustment; Census X-11 program; Pseudospectral density function; Model-based decomposition; Canonical decomposition.

1. INTRODUCTION

Business and economic time series frequently exhibit seasonality—periodic fluctuations that recur with about the same intensity each year. It has been argued (c.f., Nerlove, Grether, and Carvalho 1979, p. 147) that seasonality should be removed from economic time series so that underlying "business cycles" can be more easily studied and current economic conditions can be appraised. Of the large number of seasonal adjustment procedures, the most widely used is the Census X-11 method described in Shiskin, Young, and Musgrave (1967). The X-11 program and other methods that have been empirically developed tend to produce what their developers feel are desirable seasonal adjustments, but their statistical properties are difficult to assess from a theoretical viewpoint. Recently, there has been considerable interest in developing model-based procedures for the decomposition and seasonal adjustment of time series (see, e.g., the work of Grether and Nerlove 1970; Cleveland and Tiao 1976; Pierce 1978, 1980; Box, Hillmer, and Tiao 1978; Tiao and Hillmer 1978; and Burman 1980). Following this line of work and motivated in part by the considerations in the X-11 program, this article proposes a model-based approach that decomposes a time series into seasonal, trend, and irregular components.

We suppose that an observable time series at time t, Z_t, can be represented as

$$Z_t = S_t + T_t + N_t. \tag{1.1}$$

where S_t, T_t, and N_t are unobservable seasonal, trend, and noise components. It may be the case that a more accurate representation for Z_t would be as the product of S_t, T_t, and N_t. In that situation the model (1.1) would be appropriate for the logarithms of the original series. We assume that each of the components follows an ARIMA model,

$$\begin{aligned} \phi_S(B)S_t &= \eta_S(B)b_t \\ \phi_T(B)T_t &= \eta_T(B)c_t \\ \phi_N(B)N_t &= \eta_N(B)d_t, \end{aligned} \tag{1.2}$$

where B is the backshift operator such that $BS_t = S_{t-1}$; each of the pairs of polynomials $\{\phi_S(B), \eta_S(B)\}$, $\{\phi_T(B), \eta_T(B)\}$, and $\{\phi_N(B), \eta_N(B)\}$ have their zeros lying on or outside the unit circle and have no common zeros; and b_t, c_t, and d_t are three mutually independent white noise processes, identically and independently distributed as $N(0, \sigma_b^2)$, $N(0, \sigma_c^2)$, and $N(0, \sigma_d^2)$, respectively. Then it is readily shown that the overall model for Z_t is the ARIMA model

$$\varphi(B)Z_t = \theta(B)a_t, \tag{1.3}$$

where $\varphi(B)$ is the highest common factor of $\phi_S(B)$, $\phi_T(B)$, and $\phi_N(B)$, and $\theta(B)$ and σ_a^2 can be obtained from the relationship

$$\frac{\theta(B)\theta(F)\sigma_a^2}{\varphi(B)\varphi(F)} = \frac{\eta_S(B)\eta_S(F)\sigma_b^2}{\phi_S(B)\phi_S(F)} + \frac{\eta_T(B)\eta_T(F)\sigma_c^2}{\phi_T(B)\phi_T(F)} + \frac{\eta_N(B)\eta_N(F)\sigma_d^2}{\phi_N(B)\phi_N(F)}, \tag{1.4}$$

where $F = B^{-1}$. We also assume that the parameters in (1.3) are known. In practice a model for the observable series Z_t can be built from the data, and the estimated parameter values used as if they were the true values.

The ARIMA form has been found flexible enough to describe the behavior of many actual nonstationary and seasonal time series (Box and Jenkins 1970). There are situations in which such models by themselves may not be adequate; for example, a series describing employment may be dramatically affected by a strike and the model (1.3) does not cover such contingencies. However,

* S.C. Hillmer is Assistant Professor, School of Business, University of Kansas, Lawrence, KS 66044. G.C. Tiao is Professor of Statistics and Business, Department of Statistics, University of Wisconsin, Madison, WI 53706. This research was partially supported by the U.S. Bureau of the Census, Washington, D.C. under JSA No. 80-10; Army Research Office, Durham under grant No. DAAG29-78-G-0166; the ALCOA Foundation; and University of Kansas research allocation 3159-X0-0038. The article was completed when G.C. Tiao was on leave at the University of Chicago.

© Journal of the American Statistical Association
March 1982, Volume 77, Number 377
Theory and Methods Section

in these situations ARIMA models can frequently be modified to approximate reality; for instance, intervention analysis techniques described in Box and Tiao (1975) might be used to account for the effects of strikes and other exogenous events.

Given the observable Z_t and the structure in (1.1), (1.2), and (1.3), the problem is to decompose Z_t into S_t, T_t, and N_t. Our approach is as follows: (a) We first impose restrictions on $\phi_S(B)$ and $\phi_T(B)$ for the component models (1.2) based in part on considerations in the Census X-11 program. (b) A model for Z_t is derived from observable data. (c) A principle is adopted that uniquely specifies the component models in a manner consistent with the imposed restrictions and the model derived for Z_t. (d) Given the component models, known signal extraction methods are applied to decompose Z_t into (estimates of) the components. Properties of the procedure are explored and an illustration using an actual time series is presented.

2. DECOMPOSITION WHEN THE COMPONENT MODELS ARE KNOWN

If in (1.1) the stochastic structures of S_t, T_t, and N_t in (1.2) are known, then estimates of S_t and T_t can be readily obtained (see, e.g., Whittle 1963 and Cleveland and Tiao 1976). Specifically, Cleveland and Tiao have shown that, when all the zeros of $\phi_S(B)$, $\phi_T(B)$, and $\phi_N(B)$ are on or outside of the unit circle, the minimum mean squared estimates of the seasonal and trend components S_t and T_t are, respectively,

$$\hat{S}_t = W_S(B)Z_t \quad \text{and} \quad \hat{T}_t = W_T(B)Z_t, \tag{2.1}$$

where

$$W_S(B) = \frac{\sigma_b^2}{\sigma_a^2}\frac{\varphi(B)\varphi(F)\eta_S(B)\eta_S(F)}{\theta(B)\theta(F)\phi_S(B)\phi_S(F)}$$

and

$$W_T(B) = \frac{\sigma_c^2}{\sigma_a^2}\frac{\varphi(B)\varphi(F)\eta_T(B)\eta_T(F)}{\theta(B)\theta(F)\phi_T(B)\phi_T(F)}.$$

Because in practice the S_t, T_t, and N_t series are unobservable, it is usually unrealistic to assume that the component models in (1.2) are known. As a result, the weight functions $W_S(B)$ and $W_T(B)$ cannot be determined and the values $\hat{S}_t$ and $\hat{T}_t$ cannot be calculated. We can, however, get an accurate estimate of the model (1.3) from the observable Z_t series. Consequently, it is of interest to investigate to what extent a known model for Z_t will determine the models for the component series.

3. PROPERTIES OF SEASONAL AND TREND COMPONENTS

It is well known that the Census X-11 procedure may be approximated by a linear filter (for instance see Young 1968 and Wallis 1974). One important feature of the X-11 filter weights for the trend and the seasonal components is that the weights applied to observations more removed from the current time period decrease. This feature was incorporated into the X-11 program probably because of the belief that the trend and seasonal components of many series change over time; consequently the information about the current trend or seasonal is contained in the values of Z_t close to current time. Therefore, in developing a decomposition procedure we should allow for evolving trend and seasonal components.

Stochastic Trend

Economic data often exhibit underlying movements that drift over time. While *locally* such movements might be adequately modeled by a polynomial in time, a fixed polynomial time function is clearly inappropriate over the *entire* time span. Thus a stochastic trend model is needed, and we assume that the trend component, T_t, follows the nonstationary model

$$(1 - B)^d T_t = \eta_T(B)c_t, \tag{3.1}$$

where $\eta_T(B)$ is a polynomial in B of degree at most d and c_t are iid $N(0, \sigma_c^2)$. Box and Jenkins (1970, p. 14[illegible]) have shown that the minimum mean squared error forecast function of (3.1) is a polynomial time function of degree $(d - 1)$ whose coefficients are updated as the origin of forecast is advanced; therefore (3.1) can be regarded as a polynomial model with stochastic coefficients.

It is also of interest to consider the trend component in the frequency domain. Intuitively, the spectral density function of a trend component should be large for the low frequencies and small for higher frequencies. Since the model (3.1) is nonstationary, the spectral density function is strictly speaking not defined. However, we can define a pseudospectral density function (psdf) for (3.1) by

$$f_T(w) = \sigma_c^2\eta_T(e^{iw})\eta_T(e^{-iw})/(1 - e^{iw})^d(1 - e^{-iw})^d, \quad 0 \le w \le \pi. \tag{3.2}$$

Now the psdf (3.2) is infinite at $w = 0$ and very large for small w. This is consistent with what could be viewed as a stochastic trend component.

Stochastic Seasonal

A deterministic seasonal component S_t of period s would have the property that it repeats itself every s periods and that the sum of any s consecutive components should be a constant, that is,

$$S_t = S_{t-s} \quad \text{and} \quad U(B)S_t = c, \tag{3.3}$$

where $U(B) = 1 + B + \ldots + B^{s-1}$ and c is an arbitrary constant that can be taken as zero. Such a model, however, implies that the seasonal pattern is fixed over time. For business and economic time series, it seems reasonable to require that the seasonal component should be capable of evolving over time but that *locally* a rea[illegible] seasonal pattern should be preserved. In other words, $U(B)S_t$ should be random but cluster about zero. Consider the nonstationary model

$$U(B)S_t = \eta_S(B)b_t,$$

where $\eta_S(B)$ is a polynomial in B of degree at most $s - 1$ and b_t are iid $N(0, \sigma_b^2)$. That is, the consecutive moving sum of s components, $U(B)S_t$, follows a moving average model of order (at most) $s - 1$. It is readily shown that the forecasting function of (3.4) at a given time origin follows a fixed seasonal pattern of period s, but the pattern is updated as the origin is advanced. Also, $EU(B)S_t = E\eta_S(B)b_t = 0$. Thus, the model (3.4) preserves a local cyclical pattern but allows seasonality to evolve over time.

It is also informative to consider the psdf, $f_S(w)$, of the model in (3.4)

$$f_S(w) = \sigma_b^2 \frac{\eta_S(e^{iw})\eta_S(e^{-iw})}{U(e^{iw})U(e^{-iw})}. \tag{3.5}$$

It can be shown that $f_S(w)$ has the following properties: (a) $f_S(w)$ is infinite at the seasonal frequencies $w = 2k\pi/s$ for $k = 1, \ldots, [s/2]$, where $[x]$ denotes the greatest integer less than or equal to x; (b) $f_S(w)$ has relative minimum at $w = 0$ and near the frequencies $w = ((2k - 1)\pi)/s$ for $k = 2, \ldots, [s/2]$. Therefore, the psdf of (3.4) has infinite power at the seasonal frequencies and relatively small power away from the seasonal frequencies.

4. MODEL-BASED SEASONAL DECOMPOSITION

From considerations in the previous section, we require $\varphi(B)$ to contain the factor $U(B)$ before we impose a seasonal component S_t and to contain the factor $(1 - B)^d$ before we impose a trend component T_t for Z_t. We further require that in (1.2) the autoregressive polynomial of N_t, $\phi_N(B)$, has no common zeros with either $(1 - B)^d$ or $U(B)$, because otherwise it would imply the existence of additional seasonal and trend components that could then be absorbed into S_t and T_t. Thus, we shall suppose that in (1.3)

$$\varphi(B) = (1 - B)^d U(B)\phi_N(B), \tag{4.1}$$

where the three factors on the right side have no common zeros. In other words, knowing the model for Z_t and assuming that a decomposition is possible, the autoregressive polynomials of S_t, T_t, and N_t can be uniquely determined. Also, the relationship (1.4) becomes

$$\frac{\theta(B)\theta(F)}{\varphi(B)\varphi(F)}\sigma_e^2 = \frac{\eta_S(B)\eta_S(F)}{U(B)U(F)}\sigma_b^2 + \frac{\eta_T(B)\eta_T(F)}{(1 - B)^d(1 - F)^d}\sigma_c^2 + \frac{\eta_N(B)\eta_N(F)}{\phi_N(B)\phi_N(F)}\sigma_d^2. \tag{4.2}$$

The more difficult task is to determine the moving average polynomials and the innovation variances. Within the class of $\eta_S(B)$ and $\eta_T(B)$ whose degrees are at most $(s - 1)$ and d as required by (3.1) and (3.4), any choice of the three moving average polynomials $\eta_S(B)$, $\eta_T(B)$, and $\eta_N(B)$ and the three variances σ_b^2, σ_c^2, and σ_d^2 satisfying (4.2) will be called an *acceptable* decomposition because it is consistent with information provided by the model for the observed data Z_t.

We now give a necessary and sufficient condition for the existence of an acceptable decomposition. Assuming that $\varphi(B)$ takes the form (4.1), we may perform a unique partial fraction decomposition of the left side of (4.2) to yield

$$\frac{\theta(B)\theta(F)}{\varphi(B)\varphi(F)}\sigma_a^2 = \frac{Q_S(B)}{U(B)U(F)} + \frac{Q_T(B)}{(1 - B)^d(1 - F)^d} + \frac{Q_N(B)}{\phi_N(B)\phi_N(F)}, \tag{4.3}$$

where

$$Q_S(B) = q_{0S} + \sum_{i=1}^{s-2} q_{iS}(B^i + F^i),$$

$$Q_T(B) = q_{0T} + \sum_{i=1}^{d-1} q_{iT}(B^i + F^i),$$

and $Q_N(B)$ can be obtained by subtraction. The uniqueness in (4.3) results from the fact that the degrees of $Q_S(B)$ and $Q_T(B)$ are lower than the degrees of the corresponding denominator. Now for $0 \leq w \leq \pi$, let

$$\epsilon_1 = \min_w \frac{Q_S(e^{-iw})}{|U(e^{-iw})|^2},$$

$$\epsilon_2 = \min_w \frac{Q_T(e^{-iw})}{|1 - e^{-iw}|^{2d}}, \tag{4.4}$$

and

$$\epsilon_3 = \min_w \frac{Q_N(e^{-iw})}{|\phi_N(e^{-iw})|^2}.$$

We now show that an acceptable decomposition exists if and only if $\epsilon_1 + \epsilon_2 + \epsilon_3 \geq 0$.

Proof. By writing $B = e^{-iw}$, $0 \leq w \leq \pi$, each of the three terms on the right side of (4.2) is a psdf.

Since $\eta_S(B)$ is of degree at most $s - 1$ and $\eta_T(B)$ is of degree at most d, by comparing (4.2) with (4.3) we can write

$$\frac{|\eta_S(e^{-iw})|^2\sigma_b^2}{|U(e^{-iw})|^2} = \frac{Q_S(e^{-iw})}{|U(e^{-iw})|^2} + \gamma_1,$$

$$\frac{|\eta_T(e^{-iw})|^2\sigma_c^2}{|1 - e^{-iw}|^{2d}} = \frac{Q_T(e^{-iw})}{|1 - e^{-iw}|^{2d}} + \gamma_2, \tag{4.5}$$

and

$$\frac{|\eta_N(e^{-iw})|^2\sigma_d^2}{|\phi_N(e^{-iw})|^2} = \frac{Q_N(e^{-iw})}{|\phi_N(e^{-iw})|^2} + \gamma_3,$$

where γ_1, γ_2, and γ_3 are three constants such that $\gamma_1 + \gamma_2 + \gamma_3 = 0$. The constants γ_i provide a means to change from the initial partial fractions decomposition (4.3) to an acceptable decomposition if one exists. Thus, an acceptable decomposition implies and is implied by the fact that $\gamma_i + \epsilon_i \geq 0$ for $i = 1, 2, 3$ or equivalently that $\epsilon_1 + \epsilon_2 + \epsilon_3 \geq 0$.

From the previous discussion, when $\epsilon_1 + \epsilon_2 + \epsilon_3 \geq 0$, every set of γ_i's corresponds to a unique acceptable decomposition; thus a unique decomposition exists if and only if $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$. On the other hand, when $\epsilon_1 + \epsilon_2 + \epsilon_3 > 0$, there are an infinite number of ways of adding constants to the three terms on the right side of (4.3) to obtain acceptable decompositions.

5. A CANONICAL DECOMPOSITION

In the absence of prior knowledge about the precise stochastic structure of the trend and seasonal components, all of the information in the known model of Z_t, (1.3), about S_t and T_t is embodied in (4.2). However, when $\epsilon_1 + \epsilon_2 + \epsilon_3 > 0$, this information is not sufficient to uniquely determine the models for S_t and T_t. To perform seasonal adjustment of the data, an arbitrary choice must be made. Considering that the seasonal and trend components should be slowly evolving, it seems reasonable to extract as much white noise as possible from the seasonal and trend components subject to the restrictions in (4.2). Thus, we seek to maximize the innovation variance σ_d^2 of the noise component N_t. Therefore, we define the *canonical decomposition* as the decomposition that maximizes σ_d^2 subject to the restrictions in (4.2).

Properties of the Canonical Decomposition

In the following we denote the canonical seasonal component by $\tilde{S}_t$, the canonical trend component by $\tilde{T}_t$, and use the same convention when referring to the moving average polynomials and innovation variances of the canonical decompositions. We prove the following properties of the canonical decomposition in the appendix. (a) The canonical decomposition is unique. (b) It minimizes the innovation variances σ_b^2 and σ_c^2. (c) The polynomials $\tilde{\eta}_S(B)$ and $\tilde{\eta}_T(B)$ have at least one zero on the unit circle so that the models for $\tilde{S}_t$ and $\tilde{T}_t$ are noninvertible. (d) If $\hat{S}_t$ and $\hat{T}_t$ are any acceptable seasonal and trend components other than the canonical decomposition, then $\hat{S}_t = \tilde{S}_t + e_t$ and $\hat{T}_t = \tilde{T}_t + \alpha_t$, where e_t and α_t are white noise series. (e) The variance of $U(B)S_t$ is minimized for the canonical decomposition.

One may lend justification of the (arbitrary) choice of the canonical decomposition on the basis of these properties. In particular, property (b) is intuitively pleasing since the randomness in S_t arises from the sequence of b_t's and the randomness in T_t arises from the sequence of c_t's. Thus, minimizing σ_b^2 and σ_c^2 makes the seasonal and trend components as deterministic as possible while remaining consistent with the information in the observable Z_t series. Also, from property (d) any acceptable seasonal component can be viewed as the sum of the canonical seasonal and white noise. But $\tilde{S}_t$ is a highly predictable component that accounts for all of the seasonality in the original series and e_t is a completely unpredictable component. Thus, one might argue that the choice of an acceptable decomposition other than the canonical decomposition only produces a more confused seasonal component than necessary. Finally, property (e) is intuitively pleasing since $E[U(B)S_t] = 0$ and a small value for $\text{var}[U(B)S_t]$ will help ensure that the sum of s consecutive seasonal components remains close to zero.

6. APPLICATION TO SOME SPECIAL SEASONAL MODELS

We now illustrate the results in the preceding sections with the following three special cases of (1.3). These models have been frequently used in practice to fit seasonal data (see, e.g., Box and Jenkins 1970, and Tiao, Box, and Hamming 1975).

$$(1 - B^s)Z_t = (1 - \theta_2 B^s)a_t, \tag{6.1}$$

$$(1 - B)(1 - B^s)Z_t = (1 - \theta_1 B)(1 - \theta_2 B^s)a_t, \tag{6.2}$$

and

$$(1 - B^s)Z_t = (1 - \theta_1 B)(1 - \theta_2 B^s)a_t. \tag{6.3}$$

Without loss of generality, we assume that $\sigma_a^2 = 1$. For these models, the general approach is as follows. We first divide the denominator of the left side of (4.3) into the numerator to obtain $Q_N(B)$ and a remainder term $R(B)$; we then perform a partial fractions expansion of $R(B)/\varphi(B)\varphi(F)$ to obtain $Q_S(B)$ and $Q_T(B)$; and finally we find the minimum values ϵ_1, ϵ_2, and ϵ_3 in order to investigate whether an acceptable decomposition exists.

The model (6.1)

In this case, $d = 1$ and $\phi_N(B) = 1$. By partial fraction, (4.3) becomes

$$\frac{(1 - \theta_2 B^s)(1 - \theta_2 F^s)}{(1 - B^s)(1 - F^s)} = \frac{Q_S(B)}{U(B)U(F)} + \frac{Q_T(B)}{(1 - B)(1 - F)} + \theta_2, \tag{6.4}$$

where

$$Q_T(B) = \frac{1}{s^2}(1 - \theta_2)^2$$

and

$$Q_S(B) = (1 - \theta_2)^2 \left[1 - \frac{1}{s^2} U(B)U(F)\right] \Big/ (1 - B)(1 - F)$$
$$= \frac{1}{6s^2}(1 - \theta_2)^2 \times \left[\sum_{i=2}^{s-1}(i - 1)i(i + 1)(B^{s-i} + F^{s-i}) + (s - 1)s(s + 1)\right].$$

For the trend component, we see that $Q_T(e^{-iw})\,|1 - e^{-iw}|^{-2}$ is monotonically decreasing in w and

$$\epsilon_2 = \frac{1}{4s^2}(1 - \theta_2)^2.$$

For the seasonal component, it is easy to show that $Q_S(e^{-iw})\,|U(e^{-iw})|^{-2} \geq 0$ and has a local minimum at $w = 0$. Also, we conjecture that $w = 0$ is in fact the global minimum. This conjecture is verified analytically for $s \leq 7$ and numerically for s from 4 to 20. Assuming this is true for all s we find that

$$\min_w \left\{1 - \frac{1}{s^2}|U(e^{-iw})|^2\right\} \Big/ |1 - e^{-iw}|^2 \qquad (6.5)$$

$$= (s^2 - 1)/12$$

so that $\epsilon_1 = (1 - \theta_2)^2(s^2 - 1)/12s^2$. Since $\epsilon_3 = \theta_2$, for an acceptable decomposition to exist it is required that

$$\epsilon_1 + \epsilon_2 + \epsilon_3 = \theta_2 + \frac{(1 - \theta_2)^2}{4s^2} + \frac{(s^2 - 1)(1 - \theta_2)^2}{12s^2} \geq 0$$

or equivalently

$$\theta_2 \geq -\frac{(5s^2 - 2) + 2s\sqrt{6(s^2 - 1)}}{(s^2 + 2)}. \qquad (6.6)$$

Values of the lower bound of θ_2 for selected values of s are given in the following tabulation:

s	2	4	6
l.b.	−.1716	−.1170	−.1080
s	10	12	∞
−.1049	−.1035	−.1027	−.1010

Therefore, there are values of θ_2 for which the model (6.1) is not consistent with an additive decomposition as we have defined it; however, a value of $\theta_2 > -.1010$ will always lead to an acceptable decomposition.

When strict inequality is obtained in (6.6), there will be an infinite number of acceptable decompositions. The canonical decomposition corresponds to

$$\frac{\hat\sigma_s^2\hat\eta_s(B)\hat\eta_s(F)}{U(B)U(F)} = \frac{Q_S(B)}{U(B)U(F)} - \frac{s^2 - 1}{12s^2}(1 - \theta_2)^2$$

and (6.7)

$$\frac{\hat\sigma_T^2\hat\eta_T(B)\hat\eta_T(F)}{(1 - B)(1 - F)} = \frac{1}{4s^2}(1 - \theta_2)^2\frac{(1 + B)(1 + F)}{(1 - B)(1 - F)}.$$

The Model (6.2)

For this model, $d = 2$ and $\phi_N(B) = 1$. After some algebraic reduction, we find

$$\frac{(1 - \theta_1 B)(1 - \theta_2 B^s)(1 - \theta_1 F)(1 - \theta_2 F^s)}{(1 - B)(1 - B^s)(1 - F)(1 - F^s)} \qquad (6.8)$$

$$= \frac{Q_S(B)}{U(B)U(F)} + \frac{Q_T(B)}{(1 - B)^2(1 - F)^2} + \theta_1\theta_2,$$

where

$$Q_T(B) = \frac{(1 - \theta_1)^2(1 - \theta_2)^2}{s^2} \times \left\{1 + \left[\frac{\theta_2 s^2}{(1 - \theta_2)^2} + \frac{(s^2 - 4)}{12} + \frac{(1 + \theta_1)^2}{4(1 - \theta_1)^2}\right] \times (1 - B)(1 - F)\right\}$$

and

$$(1 - \theta_2)^{-2}(1 - B)^2(1 - F)^2Q_S(B)$$

$$= (1 - \theta_1)^2\left\{1 - \frac{1}{s^2}U(B)U(F)\right\} + \theta_1(1 - B)(1 - F)$$

$$- \left\{\frac{s^2 - 4}{12s^2}(1 - \theta_1)^2 + \frac{(1 + \theta_1)^2}{4s^2}\right\}(1 - B^s)(1 - F^s).$$

We now show that an acceptable decomposition exists if $\theta_2 \geq 0$.

Proof. First, setting $B = -1$ (or $w = \pi$ in $B = e^{-iw}$) in $Q_T(B)(1 - B)^{-2}(1 - F)^{-2}$, we have

$$\frac{Q_T(-1)}{16} = \frac{(1 - \theta_2)^2}{48s^2} \times \{(1 - \theta_1)^2(s^2 - 1) + 3(1 + \theta_1)^2\} \qquad (6.9)$$

$$+ \frac{\theta_2(1 - \theta_1)^2}{4} = C$$

say. The right side of (6.8) can now be written as

$$\frac{Q_S^*(B)}{U(B)U(F)} + \frac{Q_T^*(B)}{(1 - B)^2(1 - F)^2} + \theta_2\frac{(1 + \theta_1)^2}{4}, \qquad (6.10)$$

where

$$Q_T^*(B) = Q_T(B) - C(1 - B)^2(1 - F)^2$$

and

$$Q_S^* = Q_S(B) + U(B)U(F)\left\{C - \frac{\theta_2(1 - \theta_1)^2}{4}\right\}.$$

Also, it can be verified that

$$(1 - \theta_2)^{-2}(1 - B)^2(1 - F)^2Q_S^*(B)$$

$$= \frac{(1 - \theta_1)^2}{4}(1 + B)(1 + F) \times \left\{1 - \frac{1}{s^2}U(B)U(F) - \frac{s^2 - 1}{12s^2}(1 - B^s)(1 - F^s)\right\} \qquad (6.11)$$

$$+ \frac{(1 + \theta_1)^2}{4}(1 - B)(1 - F) \times \left\{1 - \frac{1}{4s^2}U(B)U(F)(1 - B)(1 - F)\right\}.$$

When $\theta_2 \geq 0$, one can readily show that $Q_T(e^{-iw})\,|1 - e^{-iw}|^{-4}$ is monotonically decreasing in w so that the second term in (6.10) is nonnegative for all w. Now, on the right side of the equation in (6.11), the second term with $B = e^{-iw}$ is clearly nonnegative for all w and, from (6.5), so is the first term. Thus, an acceptable decomposition exists and is given by (6.10)

The Model (6.3)

In this case, $d = 1$ and $\phi_N(B) = 1$. By partial fraction, we find

$$\frac{(1-\theta_1B)(1-\theta_1F)(1-\theta_2B^s)(1-\theta_2F^s)}{(1-B^s)(1-F^s)} = \frac{Q_S(B)}{U(B)U(F)} + \frac{Q_T(B)}{(1-B)(1-F)} + Q_N(B), \quad (6.12)$$

where

$$Q_T(B) = \frac{1}{s^2}(1-\theta_1)^2(1-\theta_2)^2,$$

$$(1-\theta_2)^{-2}(1-B)(1-F)Q_S(B) = \frac{(1+\theta_1)^2}{4}(1-B)(1-F) + (1-\theta_1)^2\left\{\frac{1}{4}(1+B)(1+F) - \frac{1}{s^2}U(B)U(F)\right\},$$

and

$$Q_N(B) = \theta_2(1-\theta_1B)(1-\theta_1F).$$

Noting that

$$\min_w Q_T(e^{-iw})\,|1-e^{-iw}|^{-2} = \frac{1}{4s^2}(1-\theta_1)^2(1-\theta_2)^2,$$

we can express the right side of (6.12) alternatively as

$$\frac{Q_S^*}{U(B)U(F)} + \frac{Q_T^*(B)}{(1-B)(1-F)} + Q_N^*(B), \quad (6.13)$$

where

$$Q_T^*(B) = Q_T(B) - \frac{1}{4s^2}(1-\theta_1)^2 \times (1-\theta_2)^2(1-B)(1-F),$$

$$Q_N^*(B) = Q_N(B) + \frac{1}{4s^2}(1+\theta_1)^2(1-\theta_2)^2,$$

and

$$(1-\theta_2)^{-2}(1-B)(1-F)Q_S^*(B) = (1-\theta_1B)(1-\theta_1F)\left\{1 - \frac{1}{s^2}U(B)U(F)\right\}.$$

Similar to the model (6.2), when $\theta_2 \geq 0$, all three terms in (6.13) are nonnegative for all w so that acceptable decompositions exist.

For the models (6.2) and (6.3), acceptable decompositions also exist for negative values of θ_2 near zero. The precise lower bounds are difficult to determine analytically. However, for these as well as for any model of the form (1.3) satisfying the condition (4.1), the existence of acceptable decompositions and the corresponding canonical form can always be determined by numerical methods. A computer program to determine the canonical component models and to compute the estimates $\hat{S}_t$, $\hat{T}_t$, and $\hat{N}_t$ is available on request.

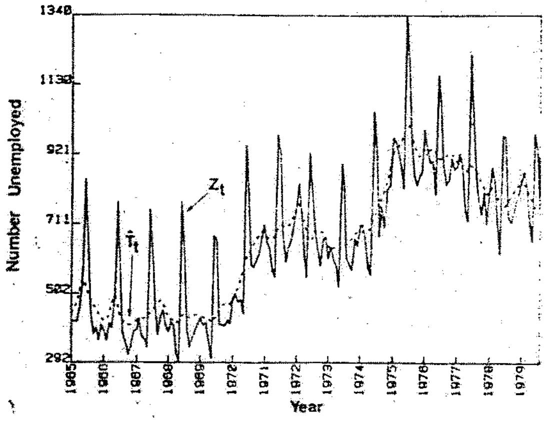


Figure 1. Monthly Unemployed Males Aged 16 to 19 (January 1971–August 1979) and the Estimated Trend Component Series

7. AN EXAMPLE

We now apply the model-based decomposition procedure to the monthly series of U.S. unemployed males aged 16 to 19 from January 1965 to August 1979, obtained from the Bureau of Labor Statistics. The series is a component used in constructing the monthly unemployment index.

The series is plotted in Figure 1. The variability of the series appears relatively constant over time; thus we decided to model the series in the original metric. It is found that the data can be adequately represented by the model (6.2) with

$$s = 12, \quad \hat{\theta}_1 = \underset{(.075)}{.313}, \quad \text{and} \quad \hat{\theta}_2 = \underset{(.035)}{.817}. \quad (7.1)$$

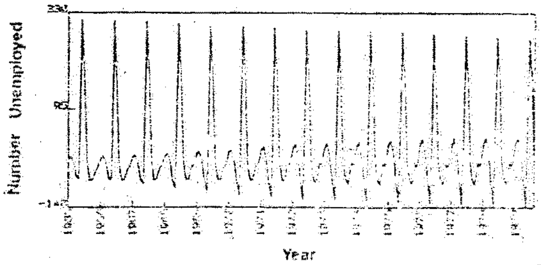


Figure 2. Estimated Seasonal Component Series for the Unemployed Males Data

Table 1. Weight Function for Estimating the Seasonal Component: Unemployed Males Data

Lag j	w_j											
0–11	.085	−.007	−.008	−.008	−.008	−.008	−.008	−.007	−.007	−.007	−.007	−.007
12–23	.076	−.007	−.007	−.007	−.006	−.006	−.006	−.006	−.006	−.006	−.006	−.006
24–35	.062	−.006	−.005	−.005	−.005	−.005	−.005	−.005	−.005	−.005	−.005	−.005
36–47	.051	−.005	−.004	−.004	−.004	−.004	−.004	−.004	−.004	−.004	−.004	−.004

with the standard errors of the parameter estimates given in parentheses below the estimates.

Assuming the estimates in (7.1) are the true values, we computed the corresponding canonical decomposition and, from (2.3), the associated weights for the estimates of the seasonal and trend components. These weights are given in Tables 1 and 2 from the center through lag 47. In both cases the remaining weights can be obtained by using the equation $w_j = .313\ w_{j-1} + .817\ w_{j-12} - .256\ w_{j-13}$. We observe that the weights associated with the seasonal component die out slowly and span a large number of years. This is in contrast to the weights associated with the standard Census X-11 program whose weights die out in about three years (see, e.g., Wallis 1974). We note that the rate at which the weight in the model-based approach decreases is primarily determined by the value of the parameter $\hat{\theta}_2 = .817$, which is determined from the original series.

The estimated trend component $\hat{T}_t$ is shown in Figure 1 and the estimated seasonal component $\hat{S}_t$ is plotted in Figure 2. We make the following observations. (a) The estimated trend component appears to capture the basic underlying movements of the series. (b) The seasonal component seems to have been adequately removed by the model-based decomposition. (c) The estimated seasonal component varies around a zero level and it is slowly changing over time. Therefore, for this particular series it appears that the model-based seasonal adjustment procedure has led to intuitively pleasing results.

8. DISCUSSION

In this article, we have proposed a model-based procedure to decompose a time series uniquely into mutually independent seasonal, trend, and irregular noise components. The method can be readily extended to models other than the ones discussed. For example, when $s = 12$, the autoregressive part of the seasonal component need not be $U(B)$, but can be any product of the factors $(1 + B)$, $(1 + B^2)$, $(1 + B + B^2)$, $(1 - B + B^2)$, $(1 + \sqrt{3}B + B^2)$, and $(1 - \sqrt{3}B + B^2)$. Also, the trend component may be augmented into a "trend-cycle" component by allowing the autoregressive part to take the form $(1 - B)^d \phi_T^*(B)$, where $\phi_T^*(B)$ has all its zeros lying on the unit circle (but distinct from $B = 1$ and those of the seasonal component). The possibilities are unlimited, depending on the form of the known model of Z_t and the nature of the problem.

Finally, we remark here that in illustrating the decomposition procedure with the models (6.1) to (6.3), in each case the values of θ_2 are restricted essentially to be nonnegative to yield acceptable decompositions. While we have rarely seen in practice a negative estimate of θ_2, it is conceivable that this could happen. One possible explanation for a negative θ_2 is that the white noise b_t and c_t for the seasonal and trend components are correlated. As an extreme example of the model (6.1) with $s = 2$, suppose the component models are

$$(1 + B)S_t = (1 - B)b_t,$$

$$(1 - B)T_t = (1 + B)c_t, \tag{8.1}$$

and

$$N_t \equiv 0.$$

The reader can readily verify that if $\sigma_b^2 = \sigma_c^2$ and b_t and c_t are perfectly positively correlated, then $\theta_2 = -1$. Thus, by allowing the component models to be dependent, we could increase the range of the models of Z_t for which acceptable decompositions exist. This seems to be an interesting topic for further study.

APPENDIX

In this appendix we sketch the proof of the properties of the canonical decomposition given in Section 5. Upon multiplying each expression in (4.5) by the denominator on the left side of the corresponding equation, we obtain

Table 2. Weight Function for Estimating the Trend Component: Unemployed Males Data

j	w_j											
0–11	.318	.212	.072	.025	.014	.010	.008	.008	.007	.005	.001	−.012
12–23	−.021	−.012	.001	.005	.006	.006	.006	.006	.006	.004	.001	−.009
24–35	−.018	−.010	.001	.004	.005	.005	.005	.005	.005	.004	.001	−.008
36–47	−.014	−.008	.001	.003	.004	.004	.004	.004	.004	.003	.001	−.006

$$
\begin{aligned}
|\eta_S(e^{-iw})|^2 \sigma_b^2 &= Q_S(e^{-iw}) + \gamma_1 |U(e^{-iw})|^2 \\
&= f_S(w, \gamma_1), \\
|\eta_T(e^{-iw})|^2 \sigma_c^2 &= Q_T(e^{-iw}) + \gamma_2 |1 - e^{-iw}|^{2d} \\
&= f_T(w, \gamma_2), \\
|\eta_N(e^{-iw})|^2 \sigma_d^2 &= Q_N(e^{-iw}) + \gamma_3 |\phi_N(e^{-iw})|^2 \\
&= f_N(w, \gamma_3).
\end{aligned} \tag{A.1}
$$

Using a result of Hannan (1970, p. 137), we have that

$$
\begin{aligned}
\sigma_b^2(\gamma_1) &= \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f_S(w, \gamma_1)\, dw\right\} \\
\sigma_c^2(\gamma_2) &= \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f_T(w, \gamma_2)\, dw\right\} \\
\sigma_d^2(\gamma_3) &= \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f_N(w, \gamma_3)\, dw\right\}
\end{aligned} \tag{A.2}
$$

Now in (A.1), $f_N(w, \gamma_3)$ does not depend on γ_3 if $\phi_N(e^{-iw}) = 0$ and is otherwise strictly increasing in γ_3; thus σ_d^2 is maximized when $\gamma_3 = \epsilon_1 + \epsilon_2$. From the restrictions that $\gamma_1 + \gamma_2 + \gamma_3 = 0$ and $\gamma_i + \epsilon_i \geq 0$, $i = 1, 2, 3$, we have that for the canonical decomposition $\gamma_1 = -\epsilon_1$ and $\gamma_2 = -\epsilon_2$. Therefore, the canonical decomposition is unique and furthermore, from (A.1) and (A.2), the innovation variances $\sigma_b^2(\gamma_1)$ and $\sigma_c^2(\gamma_2)$ are minimized for the canonical decomposition. In addition, if we take $\gamma_1 = -\epsilon_1$ and $\gamma_2 = -\epsilon_2$ in (A.1), both $f_S(w, -\epsilon_1)$ and $f_S(w, -\epsilon_2)$ are zero for some $0 \leq w \leq \pi$ implying that $\bar{\eta}_S(B)$ and $\bar{\eta}_T(B)$ are not invertible.

If we let

$$f_{\bar{S}}(w) = Q_S(e^{-iw}) \,|\, U(e^{-iw})\,|^{-2} - \epsilon_1$$

denote the psdf of $\bar{S}_t$ and let $f_{\tilde{S}}(w)$ denote the psdf of any other acceptable decomposition $\tilde{S}_t$, then it follows that

$$f_{\tilde{S}}(w) = f_{\bar{S}}(w) + \sigma_e^2 \tag{A.3}$$

with $\sigma_e^2 > 0$. Equation (A.3) implies $\tilde{S}_t = \bar{S}_t + e_t$, where e_t is white noise with variance σ_e^2.

Finally, from (4.5) the variance of $U(B)S_t$ is

var$[U(B)S_t]$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} [Q_S(e^{-iw}) + \gamma_1 \,|\, U(e^{-iw})\,|^2]\, dw. \tag{A.4}$$

It is evident that (A.4) is minimized when γ_1 is made as small as possible or $\gamma_1 = -\epsilon_1$, the value corresponding to the canonical decomposition.

[*Received October 1980. Revised June 1981.*]

REFERENCES

BOX, G.E.P., HILLMER, S.C., and TIAO, G.C. (1978), "Analysis and Modeling of Seasonal Time Series," in *Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, U.S. Department of Commerce, 309.

BOX, G.E.P., and JENKINS, G.M. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

BOX, G.E.P., and TIAO, G.C. (1975), "Intervention Analysis With Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70.

BURMAN, J.P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society*, Ser. A, 143, 321.

CLEVELAND, W.P., and TIAO, G.C. (1976), "Decomposition of Seasonal Time Series: A Model for the Census X-11 Program," *Journal of the American Statistical Association*, 71, 581.

GRETHER, D.M., and NERLOVE, M. (1970), "Some Properties of 'Optimal' Seasonal Adjustment," *Econometrica*, 38, 682.

HANNAN, E.J. (1970), *Multiple Time Series*, New York: John Wiley.

NERLOVE, M., GRETHER, D.M., and CARVALHO, J.L. (1979), *Analysis of Economic Time Series*, New York, Academic Press.

PIERCE, D.A. (1978), "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality Are Present," in *Seasonal Analysis of Economic Time Series*, ed. Arnold Zellner, U.S. Department of Commerce, 242.

——— (1980), "Data Revisions With Moving Average Seasonal Adjustment Procedures," *Journal of Econometrics*, 14, 95.

SHISKIN, J., YOUNG, A.H., and MUSGRAVE, J.C. (1967), "The X-11 Variant of Census Method II Seasonal Adjustment Program," Technical Paper 15, Bureau of the Census, U.S. Dept. of Commerce.

TIAO, G.C., and HILLMER, S.C. (1978), "Some Consideration of Decomposition of a Time Series," *Biometrika*, 65, 497.

TIAO, G.C., BOX, G.E.P., and HAMMING, W. (1975), "A Statistical Analysis of the Los Angeles Ambient Carbon Monoxide Data 1955–1972," *Journal of the Air Pollution Control Association*, 25, 1130.

WALLIS, K.F. (1974), "Seasonal Adjustment and the Relations Between Variables," *Journal of the American Statistical Association*, 69, 18.

WHITTLE, P. (1963), *Prediction and Regulation*, New York: D. Van Nostrand.

YOUNG, A.H. (1968), "Linear Approximations to the Census and BLS Seasonal Adjustment Methods," *Journal of the American Statistical Association*, 63, 445.

Econometrica, Vol. 44, No. 6 (November, 1976)

ON THE INSENSITIVITY OF THE AUTOREGRESSIVE MOVING AVERAGE REPRESENTATIONS OF SOME AUSTRALIAN QUARTERLY TIME SERIES

BY JOHN MCDONALD[1]

Economic researchers are rarely able to conduct surveys or design experiments to obtain evidence with which to assess theories or hypotheses but must rely on information, such as the national income accounting data, compiled by the government bureau of statistics. The bureau revises its national income estimates as more information becomes available or as a result of changes in methods of estimation or minor changes in definitions or classifications. The purpose of this paper is to show that the correlation structures and the autoregressive moving average representations of a number of Australian quarterly time series extracted from the income accounts are relatively insensitive to data revision. The same is true of the cross correlation functions between the "pre-whitened" series.

1. INTRODUCTION

THE NATIONAL INCOME ACCOUNTS are used widely in applied econometric work—indeed the time series which can be extracted from the accounts provide the basic source of data on the major macroeconomic aggregates. The Australian income and expenditure estimates are prepared from a very wide range of statistical information. This information is mainly derived from censuses and surveys undertaken by the Australian Bureau of Statistics or as a by-product of government administration processes. Some of the information is available quickly, but some information only becomes available after a delay of several years subsequent to the period to which it relates. Since the statistics are required as soon as possible, the Bureau does not wait until all the relevant information is available. Initial estimates are therefore subject to subsequent revisions as more information becomes available or as a result of changes in the method of estimation or minor changes in definitions or classifications.[2]

[1] I would like to thank Peter Monk, Tom Osborne, and Eva Åker for providing excellent research assistance. The A.R.G.C. provided financial support for the project. The responsibility for any errors is, of course, my own.

[2] Although we may hope that the revised estimates are less subject to error than the initial estimates, this may not be the case. The method of construction of the national income aggregates is extremely complicated. As time elapses, new information reveals weaknesses relating to certain elements of the method of construction, and revised estimates are produced. If all such weaknesses were corrected for, we could argue that the resulting figures would be less subject to error. However, it is by no means clear that a more accurate aggregate figure is produced by correcting any specific weakness, unless of course we can be confident that all individual errors have the same sign; i.e., the national income statistician faces his own particular version of the theory of the second best.

There is indeed evidence to suggest that the revision process may not result in national income and expenditure component estimates that are more consistent with the conceptual framework underlying the national accounts. The balancing item in the Australian seasonally unadjusted quarterly accounts, called the Statistical Discrepancy, can be thought of as an index of the overall inconsistency in the accounts. In McDonald and Monk [13] we have shown that quarter to quarter changes in the Discrepancy figures are not insignificant in size as compared with changes in the major aggregates, such as GDP, and that the Discrepancy figures exhibit marked seasonal characteristics. These undesirable features persist as revision takes place. In view of these considerations, in this paper, I have not taken the view that the most revised figures are necessarily the most accurate. My analysis is based on the safer premise that the revision process reveals some measurement errors.

Footnote continued next page

Revisions to quarterly or annual figures can be quite large and can lead to changed assessments of the state of the economy at a particular moment in time. Thus Rush and Harrison in [22] calculated that for the three years 1962–63, 1964–65, and 1969–70 the size of revisions to G.N.P. and consumption were in excess of the respective planned government budget deficit, and when Australia revalued relative to the United States dollar in December 1971, the latest available annual balance of trade figure was of the opposite sign and only one-sixth of its revised magnitude. Observations of this nature have relevance to the question of whether or not the economy can be finely tuned. In this paper, however, rather than consider the effect of revisions to estimates for particular quarters or years, I have attempted to assess the effect of revisions on a sequence of observations of the macroeconomic aggregates. Little is known of the impact of revealed measurement error on the characteristics of the Australian national income time series as a whole.

Most economic researchers estimating relationships between variables ignore the fact that the data are subject to measurement error. Nevertheless, there is evidence that the choice of data can significantly modify inferences. Thus Rush and Harrison have shown that estimation in a simple 15 equation model of the Australian economy is more sensitive to the choice of data than the choice of the estimation method and Denton and Kuiper [5], Holden [9], and Denton and Oksanen [6], examining data relating to other countries, found that data revisions could have a significant influence on inferences in structural models.[3]

A problem relating to the empirical studies of Rush and Harrison and the other authors is that their models contain specification errors and the model construction was not subject to the "tender loving care" that Howrey, Klein, and McCarthy [10] consider so important.[4] Since the results from the studies are clearly particular to the model formulation this is unfortunate. It may be argued that the relationships considered are, because of misspecifications, not very interesting and it

Tl extent of revision and the degree of reliability of the figures depends in a large part on the range and quality of the basic statistical data. The Bureau is clearly in a better position to assess the quality of these data and its method of construction of the national income aggregates than most users of the accounts. Unfortunately the Bureau has not published a detailed assessment of this type. Although the 1972–73 issue of Australian National Accounts [1, p. 19] gives a few examples of the sort of problem which can arise in relation to reliability and revision, more information on particular series is needed. Thus a reliability grading of the type produced by the C.S.O. for the United Kingdom national income accounts would be of value to users. An attempt to fill this gap in our knowledge has been made by Rush and Harrison who in [22] calculated a reliability ranking for seven seasonally unadjusted current priced Australian national income components by analyzing the impact of revision on the component estimates for particular quarters.

[3] Given that the choice of the data set may give somewhat different regression results the question arises as to which data will give the most satisfactory results. It is interesting that, even if as revisions take place the error in measurement is reduced, it could still be advantageous to use the sequence of first estimates data rather than data that has been subject to more but an unequal number of revisions. This is, I think, a somewhat counter-intuitive result. Nevertheless, it may be argued that using equally revised data is preferable if the pattern of data revisions relating to each time period is similar. The pattern of revisions for the independent and dependent variables need not be the same for the result to be true. On this and its practical importance see McDonald [11].

[4] Thus, for instance, Rush and Harrison [22, p. 23] state "no claims are made about the adequacy of the model being used to fully reflect the structural behaviour of the Australian economy" and again "No attempt was made to try to improve on poor equations such as that for the change in the value of non-arm stocks, as the bounds to such improvements are limitless and resources limited."

may be that in a correctly specified model the results would be different. Thus, for instance, if some estimates are sensitive to revisions this may be due to a high degree of multicollinearity between the explanatory variables resulting from an error in specification.

Certainly a large measure of judgement is involved in structural equation formulation. With respect to almost all economic relations there is not one well-established theory or hypothesis but a whole host of conflicting ideas. The ideas themselves are often based on individual maximizing behavior in a certain world and they are frequently applied in a cavalier manner to relationships involving large aggregates. Theory rarely gives a clear guide to the model-builder. Generally speaking the theory does not provide an unambiguous statement on the list of explanatory variables, on the functional form of the relationship, on the nature of the lags, or the structure of the disturbances. Consequently in this paper I have decided to focus on simple models of univariate time series and simple bivariate relationships. The specification of the univariate models and bivariate relationships considered requires some but not a great deal of subjective judgement. I have also worked with the most basic type of data—that is seasonally unadjusted current priced data—thus avoiding difficulties associated with different methods of seasonal adjustment and revised price indices.

The first stage of the analysis was a comparison of the correlation structure of the initial and revised time series. I then employed the model building approach developed and popularized by Box and Jenkins [2] to estimate autoregressive moving average (ARMA) models of the processes generating both initial and revised series. I also examined the cross correlation structures relating initial and revised estimates and the effect of revisions on the establishment, by means of cross correlation analysis, of simple relationships between the national income aggregates. For most series examined the correlation structures and model estimates were relatively insensitive to data revisions.[5]

2. THE DATA

Five seasonally unadjusted current priced national income time series were examined. They were the expenditure account estimate of gross domestic product;

[5] The correlation structures and ARMA models summarize the internal dependence characteristics of the time series. Zellner and Palm [23] have examined the relationship between the ARMA representations of processes and dynamic simultaneous equation models. In [23] they suggest that the ARMA representations of series may provide useful guidance on the structural specification of econometric models.

ARMA models have also been used to forecast future values of the time series. Some investigators, e.g., Nelson [15] and Naylor, Seaks, and Wichern [14] have reported that ARMA model forecasts compare favorably with the forecasts generated by econometric models. For comment on this see Howrey, Klein, and McCarthy [10] and Pagan [18]. ARMA forecasts can also be used in combination with other forecasts. Thus Granger and Newbold [7 and 16] have shown that forecasts generated by econometric models or by statistical methods may be improved by combining with ARMA forecasts, and Cooper and Nelson [4] and Pagan [18] have suggested that ARMA models may be used to forecast the exogenous variables of econometric models which are not policy variables under the control of the government.

The "pre-whitened" cross correlations between the variables can be used as an aid to model formulation (on this see Pierce [20], Pagan [18], and Granger and Newbold [8]), and can be used to develop multivariate Box-Jenkins forecasting and control models.

exports; imports; private final consumption expenditure; and wages, salaries, and supplements. The data relate to the period from the September quarter of 1958–59 to the June quarter of 1972–73, a period of 15 years. For each variable a series of first estimates, a series of estimates which had been subject to a few revisions and a series which had been subject to several revisions were constructed. The first two series were obtained by searching through a large number of issues of *Quarterly Estimates of National Income and Expenditure*, but most of the observations of the last series were obtained from a single supplement of that publication. This series probably corresponds most closely to the series that are most often used in research work. I will refer to the series as the "first estimates", "revised estimates", and "supplement estimates" series, respectively.

3. ANALYSIS OF THE TIME SERIES

3.1 *Model Identification, Estimation, and Diagnostic Checking*

Each of the 15 time series was graphed and the autocorrelation functions and partial autocorrelation functions of the series and the series suitably differenced were calculated. I then constructed ARMA models for each series employing the Box-Jenkins methodology. On this see Box and Jenkins [2]. I will use the GDP supplement estimates series to briefly describe the procedures.

Table I gives the autocorrelation and partial autocorrelation functions of the series and the series differenced with lags of one period, four periods, and one and four periods. Taking first differences will remove a linear trend in the series, differencing with a lag of four periods removes an additive seasonal component, and differencing with lags of both one and four periods removes both linear trend and seasonal components from the series. The autocorrelation and partial autocorrelation functions suggest that once the trend and seasonal components have been removed by differencing the resulting series can be represented by a fourth order moving average scheme; that is, the series z_t, $t = 1 \dots 60$, can be represented by the model

$$(1) \qquad (1 - B)(1 - B^4)z_t = (1 - \theta_4 B^4)a_t$$

where B is the backward shift operator, θ_4 is the moving average parameter, and a_t is a white noise source. I estimated the moving average parameter by nonlinear least squares, estimating the pre-period disturbances by the backcasting method. For details of this and other methods see Nicholls, Pagan, and Terrell [17] and Pagan and Byron [19]. The estimate of θ_4 was 0.42, and its estimated standard error, 0.15. I checked the model in several ways. First, I estimated the parameters in the model

$$(1 - \phi_1 B)(1 - \phi_4 B^4)z_t = (1 - \theta_4 B^4)a_t.$$

As expected ϕ_1 and ϕ_4 were both very close to unity, taking values of 0.999 and 0.96, respectively, and the moving average estimate was consistent with that obtained in model (1).

TABLE I

G.D.P. SUPPLEMENT ESTIMATES SERIES: AUTOCORRELATION FUNCTIONS (A.C.F.) AND PARTIAL AUTOCORRELATION FUNCTIONS (P.A.F.)

Lag	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Series $z_t, t = 1 \ldots 60$															
a.c.f.	.91	.85	.79	.76	.69	.64	.59	.57	.49	.45	.40	.38	.32	.27	.23
p.a.f.	.91	.15	−.01	.18	−.25	.05	.02	.06	−.18	.02	.03	.02	−.16	.02	−.01
Series $(1 - B)z_t, t = 1 \ldots 59$															
a.c.f.	−.31	−.33	−.25	.87	−.27	−.31	−.21	.80	−.26	−.30	−.20	.75	−.25	−.28	−.17
p.a.f.	−.31	−.47	−.77	.57	.21	.14	.02	.18	.11	.09	−.10	.01	−.06	−.06	−.05
Series $(1 - B^4)z_t, t = 1 \ldots 56$															
a.c.f.	.80	.67	.51	.46	.47	.52	.47	.38	.32	.25	.23	.19	.16	.13	.14
p.a.f.	.80	.08	−.11	.15	.23	.15	−.16	−.13	.11	−.04	−.07	−.12	.03	.07	.09
Series $(1 - B)(1 - B^4)z_t$, $t = 1 \ldots 55$															
a.c.f.	−.13	.03	−.13	−.29	−.05	.12	.18	−.00	−.02	−.02	.05	−.17	.13	−.12	−.23
p.a.f.	−.13	.02	−.13	−.34	−.15	.08	.15	−.09	−.09	.09	.23	−.19	.04	−.06	−.26

The autocorrelation function of the residuals from the estimated model (1) exhibited a minor but not significant peak at lag-1, the coefficient being 0.23, but for the rest the function was flat and the χ^2 statistic (see Box and Pierce [3]) took a value of 10.7 which is not significant at the 50 per cent level. The cumulative periodogram of the residuals also suggested that the fitted model (1) was adequate. In addition, overfitting tests, allowing for a constant and for a first order moving average parameters as well as the fourth order scheme, indicated that the additional parameters were not significant.

3.2. *The Sensitivity of the Models with Respect to Data Revisions*

In Table II the main results for the five macro variables series are summarized. For exports, imports, and wages, salaries, and supplements, the graphs of the series and autocorrelation and partial autocorrelation functions constructed from the first, revised, and supplements estimates series were very similar indeed. The analysis of the autocorrelation and partial autocorrelation functions of the three imports series suggested the model

$$w_t - c = a_t, \quad \text{where} \quad w_t = (1 - B)z_t;$$

that is the series followed random walks about a trend. The preferred models for the exports series were

$$(1 - \phi_4 B^4)(w_t - c) = a_t, \quad \text{where} \quad w_t = (1 - B)z_t$$

although the moving average representation

$$w_t - c = (1 - \theta_4 B^4)a_t, \quad \text{where} \quad w_t = (1 - B)z_t$$

performed almost as well. The wages, salaries, and supplements series autocorrelation functions suggested the model

$$(1 - B)(1 - B^4)z_t = (1 - \theta_5 B^5)a_t.$$

In relation to this model it is worth remembering that the double differencing of the series affects the lag-5 correlations.

The parameter estimates given in Table II indicate that, although the wages, salaries, and supplements estimates are a little more variable than one might expect from the estimated standard errors, the ARMA representations of these three variables are relatively insensitive to data revisions.

Although the GDP series produced similar graphs, autocorrelation, and partial autocorrelation functions for the undifferenced series, the Box-Jenkins identification procedure resulted in a slightly more complicated model being selected for the first estimates series than for the other two series.

The autocorrelation function of the first estimates series $(1 - B)(1 - B^4)z_t$, $t = 1 \ldots 55$, has peaks at both lag-1 and lag-4 with values of -0.32 and -0.34, respectively, suggesting the model

$$(2) \qquad (1 - B)(1 - B^4)z_t = (1 - \theta_1 B)(1 - \theta_4 B^4)a_t.$$

TABLE II

MODEL ESTIMATES FOR THE IMPORTS; EXPORTS; WAGES, SALARIES, AND SUPPLEMENTS; GDP; AND FINAL CONSUMPTION EXPENDITURE SERIES[a]

Imports Series: Model $w_t - c = a_t$, where $w_t = (1 - B)z_t$.

	c
First estimates series	15.7 (7.1)
Revised estimates series	15.7 (6.2)
Supplement estimates series	16.3 (6.6)

GDP Series: Model $(1 - B)(1 - B^4)z_t = (1 - \theta_4 B^4)a_t$.

	θ_4
First estimates series	.41 (.13)
Revised estimates series	.42 (.15)
Supplement estimates series	.42 (.15)

Exports Series: [illegible] $\phi_4 B^4)(w_t - c) = a_t$, where $w_t = (1 - B)z_t$.

	c	ϕ_4
First estimates series	25.6 (12.3)	.30 (.13)
Revised estimates series	25.4 (10.9)	.28 (.13)
Supplement estimates series	25.9 (11.6)	.32 (.13)

GDP Series: Model $(1 - B)(1 - B^4)z_t = (1 - \theta_1 B)(1 - \theta_4 B^4)a_t$.

	θ_1	θ_4
First estimates series	.36 (.13)	.39 (.14)
Revised estimates series	.15 (.14)	.44 (.14)
Supplement estimates series	.18 (.14)	.43 (.15)

Wages, Salaries, and Supplements: Model $(1 - B)(1 - B^4)z_t = (1 - \theta_5 B^5)a_t$.

	θ_5
First estimates series	.47 (.12)
Revised estimates series	.62 (.10)
Supplement estimates series	.74 (.09)

Final Consumption Expenditure Series: Model $(1 - B)(1 - B^4)z_t = (1 - \theta_5 B^5)a_t$.

	θ_5
First estimates series	−.02 (.15)
Revised estimates series	.15 (.15)
Supplement estimates series	.41 (.14)

[a] Estimated standard [illegible]

In Table II the moving average parameter estimates are exhibited. They are both clearly significant. For the revised and supplement estimates series model identification suggested model (1). Although for these series the autocorrelation function of the differenced series $(1 - B)(1 - B^4)z_t, t = 1 \dots 55$, exhibited peaks at both lag-1 and lag-4, the lag-1 coefficients were not significant. If the model specification (2) is used at the estimation stage for these two series, the coefficient estimates are positive but only just bigger than their estimated standard errors. Even so, Box-Jenkins practitioners who are not greatly concerned with parsimonious parameterization may prefer model specification (2) for these series, and they would obtain fairly similar model estimates for all three series. If the first estimates series is used to estimate the moving average parameter in model (1), the estimate is very similar to those obtained from the revised and supplement estimates series. However, the autocorrelation function of the residuals has a clearly significant peak at lag-1 indicating model inadequacy.

The three final consumption expenditure series contain seasonal and trend components but once these were removed by differencing the resulting autocorrelation functions for the first and revised series were flat with χ^2 values based on 15 autocorrelations of 14.7 and 15.9, respectively. The supplement estimates series χ^2 value was only 14.7 but there was a significant peak at lag-5. Consequently the slightly more complicated model

$$(3) \qquad (1 - B)(1 - B^4)z_t = (1 - \theta_5 B^5)a_t$$

was identified for this series and the moving average parameter when estimated was clearly significant.

If model (3) is used at the estimation stage for the first estimates series, an estimate very close to zero is obtained for the moving average parameter. For the revised estimates series the moving average parameter estimate is positive but only as large as the estimated standard error. Estimates of the constant were very small indeed so the preferred model for these series was

$$(4) \qquad (1 - B)(1 - B^4)z_t = a_t.$$

3.3. *Cross Correlation Analysis*

Even if the first, revised, and supplement estimates series have similar internal structures, it is possible that one series lags another. In order to investigate this possibility I examined the cross correlation structure of the series. Two uncorrelated series may nevertheless give spuriously large cross correlations as a result of a high degree of autocorrelation within the series so the estimated ARMA models were used to "pre-whiten" the series before the correlation coefficients were computed. For the imports series three cross correlation functions, relating the first with the revised series, the first with the supplements series, and the revised with the supplements series, were computed. For these functions the lag zero coefficients were, respectively, 0.91, 0.90, and 0.99. All other lag coefficient estimates were not significantly different from zero. This suggests that one series does not lag another. This result was also obtained with the other four macroeconomic variables.

TABLE III

FINAL CONSUMPTION EXPENDITURE AND GDP: ESTIMATED CROSS CORRELATION FUNCTIONS AFTER "PRE-WHITENING"

	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10
Cross correlation function between final consumption expenditure and GDP first estimates series	.08	−.27	.06	−.08	.02	.25	.04	.05	.16	−.20	.40	−.19	−.03	.09	.25	.28	−.06	−.21	.04	−.23	.08
Function between revised estimates series	−.10	.10	.24	−.09	.18	−.16	.02	−.02	.14	−.13	.42	−.18	−.13	−.18	.27	.03	.06	−.08	.15	.10	−.22
Function between supplement estimates series	−.13	.08	.28	.02	−.05	−.13	.01	.03	.12	.03	.39	−.04	−.24	.04	.19	.15	.09	−.11	.20	−.01	−.16

Lastly I examined the effect of revisions on the estimated cross correlation functions relating the five macro aggregates. The series were "pre-whitened" and then the cross correlation functions relating all the time series were calculated. For example, Table III gives the cross correlation functions between the final consumption expenditure and GDP series. The functions exhibit similar characteristics for all three sets of data. This was the case with all the functions that were computed.

4. CONCLUSIONS

All the series examined could be adequately represented by simple ARMA forms once the series had been differenced appropriately. The GDP, final consumers expenditure, and wages, salaries, and supplements series exhibited linear trend and additive seasonal patterns. The imports and exports series only required differencing to remove a trend component although the exports series did exhibit some seasonal characteristics.

For the three variables: wages, salaries, and supplements; exports; and imports; application of Box-Jenkins methods to the preliminary and revised estimates series led to the same model specifications. Moreover the parameter estimates were quite similar. For GDP, a slightly more complicated model was preferred for the first estimates series than for the other series. For final consumption expenditure a somewhat more complicated model was obtained from the supplement estimates series than for the other two. Even so for these variables the resulting estimated models had many characteristics in common and if the same model specification is imposed on the preliminary and revised series fairly similar estimates are obtained.

On the whole the ARMA representations of the time series seem to be fairly insensitive with respect to data revisions. This is encouraging for researchers and suggests that revealed measurement errors are not so gross that they seriously affect attempts to construct simple time series models. A similar conclusion was reached from an analysis of United Kingdom data (see McDonald [12]). Two points should be emphasized, however. First, the series examined are all seasonally unadjusted, current priced data series. Second, revisions may only reveal some measurement errors in the data. It is still possible that both initial and later estimates are subject to unknown errors which result in incorrect inferences.

Flinders University of South Australia

Manuscript received June, 1975; revision received October, 1975.

REFERENCES

[1] *Australian National Accounts, National Income and Expenditure*, Canberra: Australian Bureau of Statistics.

[2] BOX, G. E. P., AND G. M. JENKINS: *Time Series Analysis Forecasting and Control*. San Francisco: Holden Day, 1970.

[3] Box, G. E. P., and D. A. Pierce: "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 64 (1970), 1509–1526.

[4] Cooper, J. P., and C. R. Nelson: "The Ex-Ante Prediction Performance of the St. Louis and F.R.B.–M.I.T.–Penn. Econometric Models." Report No. 7236, Center for Mathematical Studies in Business and Economics, University of Chicago, 1974.

[5] Denton, F. T., and J. Kuiper: "The Effect of Measurement Errors on Parameter Estimates and Forecasts: A Case Study Based on Canadian Preliminary National Accounts," *Review of Economics and Statistics*, 47 (1965), 198–206.

[6] Denton, F. T., and E. H. Oksanen: "A Multi-Country Analysis of the Effect of Data Revisions on an Econometric Model," *Journal of the American Statistical Association*, 67 (1972), 286–291.

[7] Granger, C. W. J., and P. Newbold: "Economic Forecasting—The Atheist's Viewpoint." Paper presented to the Conference on the Modelling of the U.K. Economy. London: London Graduate School of Business Studies, 1972.

[8] ———: "Spurious Regression in Econometrics," *Journal of Econometrics*, 2 (1974), 111–120.

[9] Holden, K.: "The Effect of Revisions to Data on Two Econometric Studies," *Manchester School*, 37 (1969), 23–36.

[10] Howrey, E. P., L. R. Klein, and M. D. McCarthy: "Notes on Testing the Predictive Performance of Econometric Models," *International Economic Review*, 15 (1974), 366–383.

[11] McDonald, J.: "Errors in Economic Time Series: Some Implications for Estimation in Econometric Models," *The Economic Record*, 50 (1974), 278–286.

[12] ———: "An Analysis of the Significance of Revisions to Some Quarterly U.K. National Income Time Series," *Journal of the Royal Statistical Society*, Series A, 138 (1975), 242–256.

[13] McDonald, J., and P. Monk: "An Analysis of the Statistical Discrepancy in the Australian National Accounts," *The Australian Journal of Statistics*, 17 (1975).

[14] Naylor, T. H., T. G. Seaks, and D. W. Wichern: "Box-Jenkins Methods: An Alternative to Econometric Models," *International Statistical Review*, 40 (1972), 123–137.

[15] Nelson, C. R.: "The Prediction Performance of the F.R.B.–M.I.T.–Penn. Model of the U.S. Economy," *American Economic Review*, 62 (1972), 902–917.

[16] Newbold, P., and C. W. J. Granger: "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," *Journal of the Royal Statistical Society*, Series A, 137 (1974), 131–165.

[17] Nicholls, D. F., A. R. Pagan, and R. D. Terrell: "The Estimation and Use of Models with Moving Average Disturbance Terms: A Survey," *International Economic Review*, 16 (1975), 113–134.

[18] Pagan, A. R.: "Applications of Time Series Analysis to Econometrics." Paper prepared for the C.S.I.R.O. Conference, Sydney, 1975.

[19] Pagan, A. R., and R. P. Byron: "A Synthetic Approach to the Estimation of Models with Autocorrelated Disturbance Terms," Australian National University, 1974.

[20] Pierce, D. A.: "Relationships—and the Lack Thereof—Between Economic Time Series, with Special Reference to Money, Reserves and Interest Rates," Federal Reserve Board, 1974.

[21] *Quarterly Estimates of National Income and Expenditure*. Canberra: Australian Bureau of Statistics.

[22] Rush, J. R., and S. B. M. Harrison: "Data Revisions and Some Econometric and Policy Consequences," Paper presented at the 45th A.N.Z.A.S. Congress, section 24: Economics, Perth, 1973.

[23] Zellner, A., and F. Palm: "Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, 2 (1974), 17–54.

O. D. Anderson, ed., FORECASTING

THE IMPLICATIONS FOR ECONOMIC FORECASTING OF TIME SERIES MODEL BUILDING METHODS

P. NEWBOLD and G. V. REED
Department of Mathematics and Economics (resp.), University of Nottingham

Traditional econometric forecasting methodology is based on the construction of what is often a highly complex behavioural model. In addition, there is generally a heavy judgmental content in the forecasts actually produced. Traditional methods are described in some detail and it will be noted that, when the forecasting ability of econometric models has been objectively evaluated, the results have often been rather disappointing. It will be suggested that one possible explanation is that, while the models generally take account of the structures imposed by economic theory in a fairly sophisticated way, their treatment of time series structure is almost invariably rudimentary. The time series model building methods which derive from the approach of Box and Jenkins would seem to be particularly relevant in econometrics. The uses which have been made of these techniques will be discussed, and possible extensions outlined.

1. Introduction

The development of time series model building methods, greatly stimulated by the work of Box and Jenkins (1970), has in the last few years profoundly influenced the thinking of research workers over a wide range of fields. The object of the present paper is to assess the impact of time series methods on economic model building and forecasting, both in retrospect and prospect. It is not our intention to argue the relative merits of various schools of thought or to set up a contest between methodologies which, in our view, should sensibly be regarded as allies. On the contrary, we feel that all research workers in this area have common goals – the understanding, explanation and prediction of economic behaviour. It is quite clear that time series model building methods have had a considerable impact on economic forecasters and are likely to become even more influential in the next few years as their lessons are absorbed.

Whatever the specific techniques subsequently employed, it can be argued that any quantitative model building exercise, where the facility of collecting data from designed experiments is not available, can be viewed as consisting of the following stages:

(i) the choice of a class of models that the investigator is prepared to contemplate. It may not be necessary, however, to specify analytically every member of this class before proceeding to the next stage,

(ii) the selection, at least as a starting point, of a specific model from the general class,

(iii) the estimation, based on information available to the investigator, of the unknown parameters of the chosen model,

(iv) the checking of the adequacy of the fitted model,

(v) the use of the model. Here we will restrict attention to forecasting, while acknowledging that economic models are frequently built for other purposes.

It is a matter of historical fact that, over the years, those engaged in economic forecasting have followed two distinct paths, which can be conveniently labelled as "classical econometric" and "time series". Certainly these paths appear to be coming closer together, and in one or two places the distinctions are blurred, but nevertheless for convenience of exposition we will examine and contrast these two approaches as separate philosophies. This should not be taken as approval of their separation in practical work. Indeed, we are greatly encouraged by the number of applied projects which contain elements of both philosophies. Thus, while in the next five sections of this paper we attempt to compare the "pure" or "idealised" forms of these approaches, we would be delighted to see in the future as full a merger as possible between them.

Briefly the time series approach, in its idealised form, is one where the model builder relies most heavily on the data at every stage of the analysis. Economic theory might only be appealed to for suggestions as to which variables could be relevant. The classical econometric approach appeals to, and rests heavily on, economic theory wherever possible. The evidence provided by data is often only of major importance at the parameter estimation stage.

2 Class of models

In the pure time series approach a specific model is chosen from a general class solely on the evidence of the available data. In principle current and past values of a group of time series could be inter-related in a vast number of ways. However, in order to have any reasonable chance of distinguishing between alternative models, using data series of length likely to occur in practice, it is necessary to put some restrictions on the functional forms of the models contemplated; time series analysts generally, at least as a starting point, restrict attention to linear models. Time series model building is "data based" rather than "theory based". The models achieved do not result from behavioural hypotheses, but from their ability to fit well a particular data set. In such circumstances it seems sensible to try to represent the underlying system under study with as few parameters as possible – Box and Jenkins (1970) call this the "principle of parsimonious parameterisation" – and to employ a class of models whose members describe a rich variety of structures without being over-prodigal in their use of parameters. Finally, it has been noted that many practically occurring time series tend to follow rather smooth paths and, in particular, do not appear to have a fixed mean through time. This form of non-stationary behaviour is particularly common in economic time series, and was noted from a spectral analysis point of view by Granger (1966). Very often in such circumstances either the first or higher order differences of such series appear to be stationary, in which case the time series analyst prefers to work with the differenced form, since for the undifferenced series the underlying smoothness tends to dominate all else, making other characteristics difficult to identify.

Taking these considerations into account, a number of model structures have proved valuable in the practical representation of time series behaviour. The simplest possible situation is that in which the value at time t, X_t, of a series is related to previous values $X_{t-1}, X_{t-2}, \ldots$. The class of models

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d X_t = (1 - \theta_1 B - \cdots - \theta_q B^q) a_t \tag{2.1}$$

is called ARIMA(p, d, q) – Autoregressive Integrated Moving Average. Here p, d and q are non-negative integers, B is a backshift operator on the index of the time series, defined so that $BX_t \equiv X_{t-1}$ and by repeated application, $B^j X_t \equiv X_{t-j}$, the ϕ_j's and θ_j's are fixed parameters and a_t is a "white noise" series, i.e. it has zero mean and constant variance for all t, with a_t and a_s uncorrelated for all $t \neq s$. (It is assumed here and throughout that any non-zero mean or trend has been removed from the differenced series $x_t = (1 - B)^d X_t$.) In (2.1), it is assumed that the roots of the polynomial equations in B

$$(1 - \phi_1 B - \cdots - \phi_p B^p) = 0, \qquad (1 - \theta_1 B - \cdots - \theta_q B^q) = 0$$

all have modulus greater than unity. The restriction on the first equation ensures stationarity of $(1 - B)^d X_t$ while that on the second guarantees uniqueness of representation. This is by far the most commonly used class of univariate time series models. Clearly it relates X_t linearly to $X_{t-j}, j \geqslant 1$, though an element of non-linearity can be introduced by replacing X_t by the Box and Cox (1964) transform

$$X_t^{(\lambda)} = \frac{X_t^\lambda - 1}{\lambda} \qquad \lambda \neq 0,$$

$$= \log X_t \qquad \lambda = 0,$$

where λ is now an extra parameter in the model. The use of this transformation in time series model building is described and illustrated by Wilson in the discussion of Chatfield and Prothero (1973) and by Box and Jenkins (1973), and some notes on the implications of its use for model building and forecasting are given in Granger and Newbold (1976a). Finally, it should be noted that the class (2.1) can be extended to deal with series exhibiting periodic seasonal behaviour, though this topic is outside the scope of the present paper.

Frequently, of course, the analyst has information on several related time series. The simplest case of this kind is where current and past values of a series X_2 provide information useful in forecasting future values of X_1, but the converse is not true. It might then be said, loosely, that causality runs in one direction only. (Granger (1969) would term this "empirical apparent causality within a restricted information set"). Box and Jenkins (1970), to model such relationships, use the "transfer function–noise" class

$$x_{1,t} = \frac{\omega_0 + \omega_1 B + \cdots + \omega_r B^r}{1 - \delta_1 B - \cdots - \delta_s B^s} x_{2,t} + \frac{1 - \theta_1 B - \cdots - \theta_q B^q}{1 - \phi_1 B - \cdots - \phi_p B^p} a_t \tag{2.2}$$

where $x_{1,t}$ and $x_{2,t}$ are appropriately differenced versions of the original series $X_{1,t}$ and $X_{2,t}$ and a_t is again white noise. In addition to the restrictions above on the θ_j's and ϕ_i's it is assumed that the roots of the polynomial equation in B, $(1 - \delta_1 B - \cdots - \delta^s B^s) = 0$, all have modulus greater than one. This restricts us to situations in which the influence of distant values of X_2 on X_1 becomes negligible. The model (2.2) can be extended in an obvious way to deal with the situation where X_1 is "caused by" $X_2, X_3, \ldots, X_k$.

A more general bivariate situation is where causality runs in both directions, in which case "feedback" is said to be present. Granger and Newbold (1976b, 1977) consider an obvious extension of (2.2)

$$x_{1,t} = \frac{\omega_{1,0} + \omega_{1,1}B + \cdots + \omega_{1,r_1}B^{r_1}}{1 - \delta_{1,1}B - \cdots - \delta_{1,s_1}B^{s_1}} x_{2,t} + \frac{1 - \theta_{1,1}B - \cdots - \theta_{1,q_1}B^{q_1}}{1 - \phi_{1,1}B - \cdots - \phi_{1,p_1}B^{p_1}} a_{1,t}$$

$$x_{2,t} = \frac{\omega_{2,0} + \omega_{2,1}B + \cdots + \omega_{2,r_2}B^{r_2}}{1 - \delta_{2,1}B - \cdots - \delta_{2,s_2}B^{s_2}} x_{1,t} + \frac{1 - \theta_{2,1}B - \cdots - \theta_{2,q_2}B^{q_2}}{1 - \phi_{2,1}B - \cdots - \phi_{2,p_2}B^{p_2}} a_{2,t} \tag{2.3}$$

where $a_{1,t}$ and $a_{2,t}$ are white noise series with

$$E(a_{1,t}a_{2,\tau}) = 0 \quad \text{for all } t \neq \tau.$$

A perfectly general linear representation can be obtained under either of the following restrictions

(i) $\omega_{1,0} = \omega_{2,0} = 0$,

(ii) $E(a_{1,t}a_{2,t}) = 0$,

and *one* of $\omega_{1,0} = 0$, $\omega_{2,0} = 0$ holds.

For many practical model building purposes it has been found more convenient to work with (ii).

In principle, a very general multivariate time series framework is provided by an analogue of (2.1). Let $x_1, x_2, \ldots, x_k$ be suitably differenced forms of $X_1, X_2, \ldots, X_k$. A general linear time series model is given by

$$(I - \Phi_1 B - \cdots - \Phi_p B^p)x_t = (I - \Theta_1 B - \cdots - \Theta_q B^q)a_t \tag{2.4}$$

where $x_t' = (x_{1,t}, \ldots, x_{k,t})$, the Φ_i's and Θ_j's are $k \times k$ matrices of coefficients and a_t is a vector white noise process, i.e.

$$\begin{aligned} E(a_t a_\tau') &= \Sigma \text{ (positive definite)} && \text{if } t = \tau, \\ &= 0 \text{ (null matrix)} && \text{if } t \neq \tau. \end{aligned}$$

This form was first studied in detail by Quenouille (1957). It should be added, however, that the ability to write down such models is no guarantee of their practical usefulness. If model structure is to be determined on the evidence of the data alone, it is difficult to see how large systems involving several series can be reliably constructed. Unless the relationships involved happen to be particularly simple, adoption of a purely time series approach is likely to restrict the analyst to dealing with models involving only a very few series.

In building econometric models, the analyst's chief concern is to explore formulations which are in accord with, and as far as possible derived from, some economic theory. However, theory may suggest as relevant variables which have not been, or are not capable of being, adequately measured. This, of course, imposes a further restriction, since the analyst is constrained to work with available data and may therefore need to resort to the use of various proxy measures. For example, expectations about future price levels may be hypothesised as a determinant of actual inflation (see Cagan (1956)), but some weighted average of current and past price levels would typically be used as a proxy for expectations. Again, it might be felt that the political colour of the Government in power influences the rate of wage inflation. Klein and Ball (1959) introduce a "political factor" taking the value zero when a Labour Government is in office and unity otherwise. In studies of the long run growth of income, technological progress is obviously important, but its rate cannot be directly measured (for an attempt to estimate this rate through time see Solow (1957)) and frequently time trends are introduced into models as proxies.

The simplest type of economic model hypothesises the dependence of a variable Y on k other variables $X_1, X_2, \ldots, X_k$. If the relationship can be assumed to be linear, the familiar multiple regression model,

$$Y_t = \alpha + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + U_t \tag{2.5}$$

where U_t is an error term, is typically used. More generally, economic theory frequently suggests that variables of interest are simultaneously determined. A simultaneous equation econometric model jointly determines M "endogenous" variables $Y_1, Y_2, \ldots, Y_M$ in terms of k "exogenous" or "pre-determined" variables $X_1, X_2, \ldots, X_k$. The exogenous variables are assumed to influence, but not be influenced by, the endogenous. The general linear form of the model is

$$\left.\begin{array}{l} \beta_{11} Y_{1t} + \cdots + \beta_{1M} Y_{Mt} + \gamma_{11} X_{1t} + \cdots + \gamma_{1k} X_{kt} + U_{1t} = 0 \\ \vdots \\ \beta_{M1} Y_{1t} + \cdots + \beta_{MM} Y_{Mt} + \gamma_{M1} X_{1t} + \cdots \gamma_{Mk} X_{kt} + U_{Mt} = 0 \end{array}\right\} \quad (2.6)$$

where the U_{jt} are error terms and each equation is normalised by setting one of the β_{ij} coefficients in it equal to -1. Typically, economic theory will postulate a structural form of the type (2.6) with some a priori information on the parameters – for example, that certain of them are zero.

The contrast between the kind of models contemplated by time series analysts and by econometricians appears to be very marked indeed, though as we shall see a substantial part of this difference is illusory. A great deal of emphasis is placed by time series analysts on the timing of relationships and the need to get the time lags right, using the evidence of past data as a guide. Econometricians too have long been conscious of the importance of time lags and, indeed, economic theory frequently suggests the presence of such lags. A good deal of research effort has been spent by economists on attempts to justify theoretically the introduction of lags into their models. An early example of this kind is Brown (1952), who examines a number of relationships involving lags between consumption C and income Y, the simplest of which is

$$C_t = \alpha + \beta_1 Y_t + \beta_2 C_{t-1} + U_t.$$

These are justified in general terms as reflecting "habit persistence" in consumers. Cagan (1956) proposes a model in which the quantity Y_t depends, not on X_t, but on expected future values of X_t. These expectations, X_t^*, are assumed to adapt through time according to

$$X_t^* - X_{t-1}^* = (1 - \lambda)(X_t - X_{t-1}^*) \quad 0 < \lambda < 1. \tag{2.7}$$

It is then assumed that

$$Y_t = \alpha + \beta X_t^* + U_t$$

where U_t is an error term. Combining (2.7) and (2.8) gives

$$Y_t = \alpha(1 - \lambda) + \beta(1 - \lambda)X_t + \lambda Y_{t-1} + U_t - \lambda U_{t-1} \tag{2.9}$$

a form first studied, under the assumption that U_t is white noise, by Koyck (1954). It is well known (see, for example, Muth (1960)) that the quantity X_t^* is the optimal quadratic loss forecast of $X_{t+1}, X_{t+2}, \ldots$ linear in $X_t, X_{t-1}, X_{t-2}, \ldots$ provided the process X_t is generated by the first order integrated moving average model.

$$X_t - X_{t-1} = a_t - \lambda a_{t-1}.$$

However, there is no God-given rule that an economic time series is generated by such a process, and, if it is not, the "expectations" assumed here will be sub-optimal forecasts.

This formulation was further studied by Nerlove (1956, 1958) who also gave an additional justification for the introduction of lags. This is the "partial adjustment model", in which the desired value, Y_t^*, of the dependent variable depends on X_t according to

$$Y_t^* = \alpha + \beta X_t + U_t. \tag{2.10}$$

It is then hypothesised that only a proportion $(1 - \gamma), 0 < \gamma < 1$, of the desired adjustment $(Y_t^* - Y_{t-1})$ is actually made between time $(t - 1)$ and time t, so that

$$Y_t - Y_{t-1} = (1 - \gamma)(Y_t^* - Y_{t-1})$$

Combining these two equations then gives

$$Y_t = \alpha(1 - \gamma) + \beta(1 - \gamma)X_t + \gamma Y_{t-1} + (1 - \gamma)U_t \tag{2.11}$$

This formulation was somewhat refined to take into account costs of adjustment by Eisner and Strotz (1963).

Our intention here has been to give a flavour of the rationale advanced by econometricians in formulating models involving lags. Many other hypotheses, frequently of a more sophisticated character than these just described, have been proposed. Other works concerned with the formulation and estimation of lag models include Jorgenson (1966), Griliches (1967), Zellner and Geisel (1970) and Dhrymes (1971). However, rather than discuss these further, it seems more pertinent for our present purposes to ask whether "theories" of this kind do any more than *explain* how lags *might* arise, or suggest that one ought to be on the look-out for a lagged relationship of *some* kind. We have yet to be convinced that any theory has succeeded, in a given instance, in convincingly arguing for a *specific* lag model – that is, for specific values of r and s in (2.2). Indeed we question whether it is reasonable to expect economic theory to be completely definitive in this respect. We doubt even more strongly whether economic theory has anything of value to tell us about the time series structure of the error terms in (2.5) and (2.6), since these presumably represent the parts of the dependent variables that theory has failed to explain. In principle, then, there is nothing to prevent even the most classical econometrician from contemplating a wide range of alternative lag and error structures without breaking faith with the need to conform to some economic theory.

It can be argued that, in a sense, an appeal to economic theory allows the construction of much larger models than the modest sized structures to which the time series analyst is constrained by the requirement that model selection be determined by the data. Indeed, many macro-economic models suggested by theory are very large, and Klein (1971a) has argued that, in order to adequately describe the behaviour of a national economy, a minimum of fifty equations is required. It will be argued later that this proposition begins to leak quite dramatically if time series problems are to be taken into account.

Economic theory frequently suggests that non-linear relationships are appropriate, and the econometrician, unlike the time series analyst, rarely restricts himself to linear models. In fact virtually all modern macro-economic forecasting models contain many non-linearities in the variables (see, for example, Ball et al. (1975)).

3. Model selection

Time series analysts frequently refer to model selection as the "identification problem". Unfortunately, in the literature of econometrics this term has a different, though not unrelated, meaning. The econometric identification problem, a very full treatment of which is given by Fisher (1966), is concerned with the fact that a particular underlying structure may generate a multiplicity of observationally equivalent models of a particular type. For example, given the joint distribution of the endogenous variables conditional on the exogenous, the parameters of the model (2.6) are not uniquely determined. Econometricians typically resolve this problem by appealing to economic theory for the imposition of deterministic constraints on the parameters of the model. Of course, when considering multivariate models such as (2.4), time series analysts meet a similar problem, and a solution is given by Hannan (1969). More recently Hannan (1971) and Hatanaka (1975) have investigated the problem in relation to econometric models with sophisticated lag and time series error structures.

In attempting to select a single model from a broad general class, using only data, the time series analyst encounters what could be regarded with some justification as one of the most fascinating problems in Statistics. By its nature the problem does not admit a deterministic solution or an approach through the conventional statistical hypothesis testing framework. Rather, it seems inevitable that the analyst must resort to judgment and to his ingenuity in designing tools to aid that judgment. The first step is to search for properties which distinguish different members of the class of models. For example, taking two of the simplest processes of the class (2.1),

$$X_t - \phi_1 X_{t-1} = a_t \tag{3.1}$$

and

$$X_t = a_t - \theta_1 a_{t-1} \tag{3.2}$$

and denoting the autocorrelations of a time series as

$$\rho_k = \text{Corr.}(X_t, X_{t-k}),$$

it is easily shown that if X_t is generated by (3.1)

$$\rho_k = \phi_1^k$$

while for (3.2)

$$\begin{aligned} \rho_k &= -\theta/(1+\theta^2) \quad & k &= 1, \\ &= 0 & k &= 2, 3, 4, \ldots \end{aligned}$$

Thus the pattern of the autocorrelations is markedly different for the two processes, which could therefore readily be distinguished from one another if these autocorrelations were known. However, the analyst must make do with sample estimates. Provided that the behaviour of the sample quantities mimics sufficiently closely that of the corresponding population values, the analyst may be able to select a particular model with some confidence. Of course, this implies the need for fairly long time series, for otherwise the message from the data is likely to be obscure.

For single series Box and Jenkins (1970) describe and illustrate a model selection procedure based on the sample autocorrelations and their transforms, sample partial autocorrelations, of the series X_t and its first two or three differences. Their approach has been widely used in the last few years and found to be generally effective. For example, Reid (1969) and Newbold and Granger (1974) found that a forecasting procedure based on this approach typically outperformed, in empirical comparisons on real data, methods based on exponential smoothing where an underlying model structure was implicitly assumed a priori, rather than being deduced from the data.

Box and Jenkins (1970) also give a model selection method for the transfer function–noise models (2.2), while Jenkins (1974) considers the problem in relation to multivariate models of the type (2.4). A method for selecting feedback models of the form (2.3) is given by Granger and Newbold (1976b, 1977). All of these procedures have in common an element of "pre-whitening"; that is, when examining relationships between series, as a first step a univariate model (2.1) is fitted to at least one of the series. In subsequent analysis one works with the residuals from this fitted model rather than with the original series. It would be reasonable to expect, in multivariate model building, that sample correlations between $X_{1,t}$ and $X_{2,t-k}$, $k = \ldots, -1, 0, 1, \ldots$, would be of value. These certainly do contain information about quantities of interest, the ω's and δ's of (2.2) or (2.3) for example. Unfortunately this information is, for practical purposes, inextricably intertwined with information about the autocorrelations of the individual time series. Pre-whitening resolves this difficulty by first removing these autocorrelations, and, although the procedure is time consuming, it would appear to be an essential ingredient in multivariate model selection procedures. An illustration, for an extreme case, of the difficulties of interpretation which can arise when pre-whitening is not used is given by Box and Newbold (1971).

In traditional econometric model building the central feature of model selection is the choice, or development, of a behavioural theory capable of generating an estimable structure. In many areas there are competing economic hypotheses and frequently a particular one is chosen on non-statistical grounds, while in other studies alternative models have been compared in terms of their ability to fit or forecast data. For nested hypotheses this generally involves standard statistical tests, though, of course, a non-symmetry is imposed in this way, and the choice of significance level is arbitrary. A good recent example of this approach is given in Deaton (1974), where several mostly nested models of consumer demand are compared. Frequently, however, the competing hypotheses are non-nested and a good deal of econometric research effort has been spent on examination of procedures for distinguishing between such hypotheses (see, for example, Ramsey (1969, 1970), Dhrymes et al (1972) and Peseran (1974)). A number of empirical comparisons of alternative models have been made; for example, Jorgenson and Siebert (1968) and Jorgenson, Hunter and Nadiri (1970) compare models of investment behaviour, while Ramsey and Zarembka (1971) consider various specifications of a production function. The great difficulty in many situations of this kind is that, in statistical terms, the economy is a very bad design, so that frequently it is not possible to express with any great confidence a preference for one model over another. Econometricians have, for this reason, often been unable to shed much light on theoretical debates, though one of the most important objectives of econometric research is to do just that, rather than simply to produce a forecasting model. A further difficulty in appealing to theory to specify a structural model is that in some areas no really satisfactory theory has been developed. This can perhaps fairly be said of inflation, a field deserving a great deal more theoretical investigation.

It was noted in the previous section that some theoretical justification for imposing particular lag structures is often attempted, though in fact arguments of this kind have been criticised on economic theoretical grounds by Mundlak (1966, 1967). Moreover, the arguments used appear to imply some "natural interval" for decision taking, with virtually no rationale for doing so. For example, if a model of the form (2.9) or (2.11) is deemed appropriate for quarterly data, then it will not generally be so for annual data. This problem of temporal aggregation is examined in detail by Brewer (1973), whose solution depends on the univariate time series model generating the exogenous variable X_t, a factor not considered in classical econometric studies. Our own feeling is that while economic theory might suggest the possibility of particular lag structures it is rarely, if ever, definitive on this point. Theory would appear to have even less to say about the time series structure of errors, though contrary views have been expressed by econometricians. For example, Jorgenson and Siebert (1968) make a case for white noise errors on a priori grounds alone. We find arguments of this kind difficult to accept. Typically in applied econometric work a white noise error structure is assumed at this stage, though subsequent tests of this assumption are generally made. In fact, errors are often regarded by econometricians as a necessary elaboration, to be added on once a model structure has been developed in non-stochastic terms. Thus, Desai (1975) states: "While the model . . . was in deterministic terms, the obligatory stochastic terms are added to each structural equation". (It must be added that this particular paper goes on to incorporate sophisticated time series methods in the analysis that follows.)

We have seen that, if he is to use just data to select a model, the time series analyst is likely to need fairly large samples. On the other hand, if economic theory can be used to derive a model, no such restriction is imposed on the econometrician at this stage. However, it is often acknowledged by econometricians that economic theory alone is inadequate to determine particular lag structures, and an appeal to the data is frequently made for help. For example, we can write (2.2) as

$$x_{1,t} = (V_0 + V_1 B + V_2 B^2 + \ldots) x_{2,t} + u_t \quad (3.3)$$

where u_t is an error term and the weights V_j extend into the indefinite past. The econometrician may well be able to hypothesise that all these weights are of the same sign and will generally be such that, as j increases, they will eventually tend to zero and perhaps have a specific sum. He thus has in mind a particular "shape" for the lag distribution, and may then experiment with a few models of the type (2.2) which yield this general shape (see, for example, Jorgenson and Stephenson (1967).) A commonly used approach is due to Almon (1965), who assumes that lag structure can be well approximated by taking V_j as being the value at $z = j$ of a polynomial of degree u in z, i.e.

$$V_j = d_0 + d_1 j + d_2 j^2 + \cdots + d_u j^u \quad (3.4)$$

where the d's are unknown coefficients. Further, the V_j's in (3.3) are assumed to be zero after some specified cut-off point $j = J$. Hence knowledge of the d's in (3.4) completely specifies the lag structure used. The choice of appropriate values of u and J involves a good deal of trial and error, mainly through fitting alternative forms and comparing their corrected coefficients of multiple correlation $\bar{R}^2$. In fact it is straightforward to show that this class of models is simply a sub-set of (2.2).

The degree of sophistication involved in approaches to time series error specification in econometric work varies enormously. The typical treatment is to consider as the only alternative to the white noise specification the first order autoregression

$$u_t - \phi_1 u_{t-1} = a_t \quad (3.5)$$

where a_t is white noise. Hendry (1975) claims that this describes "poor average" practice, but if so the distribution must be one in which the mode and the "poor average" coincide. As we shall see in the next section, there are plenty of observations available to the poorer side of this mode. Of course, many econometricians do go much further than the simple structure (3.5). For example, Sargan (1964) and Wallis (1972) consider higher order autoregressive processes, while moving average error structures have been studied and used by Phillips (1966), Trivedi (1970, 1975) and Hendry and Trivedi (1972). The specification of time series error structure remains a topic of considerable research interest in econometrics, and it might be reasonable to conjecture that some of this interest has been stimulated by recent developments in time series analysis. For simultaneous equations models like (2.6), time series error specification becomes far more difficult. Hendry (1974) fits a six equations model in which the errors in the behavioural equations are allowed to follow a first order vector autoregressive process

$$u_t - \Phi_1 u_{t-1} = a_t$$

where a_t is a white noise vector. This is, of course, a special case of (2.4).

When models are fitted to economic data by time series methods, as for example in Wall et al (1975), the analysis generally suggests that differences (or perhaps percentage changes or changes in logarithms) be used to remove non-stationarity in the component series. By contrast, perhaps because much of economic theory is developed in static terms, many models suggested by theory are in terms of levels and are fitted by econometricians in this form. This suggests the possibility at least of a non-stationary time series error structure, so that a simple alternative to (3.5) might be

$$u_t - u_{t-1} = a_t - \theta_1 a_{t-1}$$

4. Parameter estimation

Perhaps not surprisingly parameter estimation is the stage at which the principles employed by time series analysts and classical econometricians are most similar. Broadly speaking, in both approaches the methods used involve either maximum likelihood, based on a Normality assumption, or some variant of least squares, which generally approximates the maximum likelihood estimators and yields coefficient estimates having desirable properties for a wider range of underlying distributions.

Likelihood functions for members of the univariate class of models (2.1) have been derived by Box and Jenkins (1970) and Newbold (1974), though in fact the full likelihood function is rarely maximised (For exceptions, see Wallis (1975) and Prothero and Wallis (1976).) The more usual procedure, described by Box and Jenkins (1970), is to minimise the sum of squares term (or at least a close approximation to it) appearing in the exponent of the likelihood function. The exact likelihood function for the transfer function – noise class of models (2.2) is given by Newbold (1973), though maximisation of this function constitutes a formidable computational problem for all but the very simplest members of the class and has not, as far as we know, been attempted in practice. The usual procedure is to employ the least squares algorithm of Box and Jenkins (1970) which is computationally simpler, but makes assumptions about "starting up" values required to generate the model. A similar algorithm for the multivariate model (2.4) is given by Wilson (1973).

The likelihood function for the simultaneous equations econometric model (2.6) can easily be derived, but maximisation of this function can be computationally burdensome for all but small models. In fact an enormous number of estimation procedures have been proposed for this particular problem, and many of these are described in standard econometrics texts (for example, Johnston (1972)). Hendry (1976) has shown that these various estimators can be viewed as approximations of the full maximum likelihood estimator. Typically, in applied econometric work, the burden of computation for the various estimation procedures increases dramatically with the size of the model, so that for larger models the methods chosen tend to be cruder approximations to maximum likelihood than for smaller models.

The differences in techniques used for parameter estimation are not very marked, depending more on the sizes of the models considered than anything else. The reason is that econometricians tend largely to use data-dependent methods at this stage. In principle, restrictions on coefficient values imposed by economic theory could be incorporated into econometric model estimation, perhaps through a formal Bayesian analysis, but this is rarely done. For example, economic theory frequently postulates particular signs for parameters, but this information is generally not incorporated into the estimation procedure.

It is common practice in econometrics to report, along with the estimated equation, the coefficient of multiple correlation. For example, if $\hat{\alpha}, \hat{\beta}_1, \ldots, \hat{\beta}_k$ denote the least squares estimates of the parameters in (2.5), with estimated residuals

$$e_t = Y_t - \hat{\alpha} - \hat{\beta}_1 X_{1t} - \cdots - \hat{\beta}_k X_{kt}$$

the coefficient of multiple correlation is

$$R^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \tag{4.1}$$

An alternative is to use the "corrected" (for degrees of freedom) coefficient

$$\bar{R}^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2/(n-k-1)}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2/(n-1)} . \qquad (4.2)$$

These measures are generally treated as descriptive statistics and are intended to convey a measure of the "strength" of the relationship between the dependent variable and the independent variables. It can be questioned, from a time series viewpoint, whether or not they do so. Suppose, for example, Y_t denotes the level of an economic time series and, as is very common, this series is highly autocorrelated. Then any regression such as

$$Y_t = \alpha + \beta_1 X_t + \beta_2 Y_{t-1} + U_t \qquad (4.3)$$

in which Y_{t-1} was included as a regressor would, using (4.1) or (4.2) as a criterion, appear to have very high explanatory power. However, this would not necessarily imply a strong relationship between the series Y and X, but could merely reflect the autocorrelation in Y_t. Moreover, any relationship not involving Y_{t-1} as a regressor must be either mis-specified or (at least in the population sense) have as high power in "explaining" Y_t as equations like (4.3). Thus it will be the rule rather than the exception, when the dependent variable is the level of an economic time series, to find apparently strong relationships between economic series, though this finding may not reflect reality. To a time series analyst, an alternative measure to (4.1) might be

$$R^{*2} = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n} \hat{a}_t^2}$$

where $\hat{a}_t$'s are the estimated residuals from a univariate time series model, of the form (2.1), fitted to the dependent variable. Use of definitions like this has persuaded time series analysts, for example Pierce (1977), that many economic relationships are very weak indeed.

Against this, the econometrician might argue that it is entirely possible for Y_t to be generated by a model like (2.5) where the X_{jt}'s are truly exogenous variables and the error term U_t is stationary. The high autocorrelation in Y_t

might then be simply a result of its dependence on the highly autocorrelated exogenous variables. This is indeed so, but it remains to consider the null case – where all the β_j's in (2.5) are zero. In such circumstances, since Y_t is highly autocorrelated, its variation about its mean can be very well described by its relationship to $Y_{t-1}, Y_{t-2}, \ldots$ Hence it is hardly reasonable to ascribe the whole of this explanation to its dependence on the exogenous variables. The problem is that it is often difficult to reject with any great confidence the hypothesis that the variable Y_t is generated by an ARIMA process, independent of the exogenous variables, in favour of the alternative that the exogenous variables are determinants of Y_t. As we noted earlier, the economy is a terrible design. Of course it is open to the economist to reject this hypothesis on a priori theoretical grounds, but even then it is difficult to see what quantity of any interest is being measured by R^2, particularly when it is noted that if, as is frequently the case, Y_t is non-stationary, the probability limit of $\Sigma(Y_t - \bar{Y})^2/n$ does not exist!

5. Model checking

Time series models are prescriptions for producing white noise, so that in the general multivariate case (2.4) the objective of a time series analysis is to derive a filter which will transform the given series X_t to a white noise error series a_t. It is natural then, in trying to determine whether or not the fitted model is adequate, to check the assumption of white noise errors. For the univariate model (2.1) this involves calculating the autocorrelations of the sample residuals from the fitted equation. The distribution of these quantities is studied by Box and Pierce (1970). Similarly, for the multivariate model, the autocorrelations of and cross-correlations between the residuals from individual equations are computed. If these quantities are unusually large the implication is that the wrong model has been fitted and the specification is re-examined. An additional check involves overfitting. Extra coefficients are added to the model which is then re-estimated and the statistical significance of the estimates of the additional parameters checked. Basically the time series analyst has no deep commitment to his originally chosen model, which Box and Jenkins describe as being "tentatively entertained". It has no theoretical backing, and he is quite prepared to drop it in favour of any alternative suggested by the data as being reasonable at this stage. The new model will then be re-estimated and re-checked, and the analysis continues until a satisfactory structure has been obtained.

A great variety of procedures is available for checking the adequacy of a fitted econometric model, though in practice two strategies have become dominant. The first might be described as a "plausibility check". It is often emphasised by econometricians – see, for example, Howrey, Klein and McCarthy (1974) – that building a model is not a simple one-off exercise. In addition to postulating structural equations, economic theory often specifies the signs of many of the parameters. Moreover, based on theory or his own experience or intuition, the model builder will often have strong views as to what magnitudes are or are not "reasonable" for many of the coefficients. The estimated model is examined for plausibility and results which do not accord with prior expectations are generally viewed with scepticism. The econometrician is likely, in such circumstances, to re-examine the specification of the part of the model that is giving trouble and to try to modify this specification.

The second commonly made check is on time series error structure. Typically it is assumed at the outset that the errors are white noise, this hypothesis being checked by the Durbin and Watson (1950, 1951) statistic

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

where the e_t's are the residuals from the estimated equation. Clearly

$$d \simeq 2(1 - r_1)$$

where r_1 is the first sample autocorrelation of the residuals. The test is designed to check against the alternative of first order autoregression (3.5) in the errors, and this form is generally assumed if the test rejects the hypothesis of white noise. We have already noted that, on occasions, econometricians go rather farther than this, but it should be added that frequently even less is done. The literature contains an abundance of studies in which the Durbin–Watson statistic is either not calculated at all or, when calculated, its warning is unheeded. One recent example, among many, is the model of Andersen and Carlsen (1974); although in three of the five behavioural equations R^2 is bigger than d no corrective action is taken. The consequences of ignoring autocorrelated errors are illustrated in terms of the "spurious regression" phenomenon by Granger and Newbold (1974). It should be added that, although the statistic is frequently calculated in such situations, the Durbin–Watson test is invalid when lagged dependent variables are included in the regressors, being biased against rejection of the null hypothesis (see Nerlove and Wallis (1966)). A test which is valid for large samples in this situation is given by Durbin (1970).

In time series analysis only the data is used to check the validity of a fitted model, while in econometric analysis an appeal is also made to economic theory. This provides a worthwhile added dimension, though there is always the danger of rejecting a perfectly good theory simply because data of sufficient quality is not available. We have seen that the usual econometric test for autocorrelated errors is based on only the first sample autocorrelation of the residuals. This may well not be adequate to detect particular departures from white noise in the errors and the usual time series practice of also calculating higher order autocorrelations seems preferable.

The assumption that the first order autoregressive process is the only alternative to white noise that need be contemplated should give a time series analyst cause for concern. This assumption is so prevalent, though certainly not universal, in applied econometric work that its consequences deserve further study. Hendry and Trivedi (1972) produce interesting simulation results on the consequences of assuming an autoregressive structure when the true structure is moving average, and vice versa. A more extreme form of misspecification has been considered by Newbold and Davies (1978). These authors are concerned about the possibility of spurious regressions when the variables involved exhibit time series behaviour typical of the levels of economic series. Series of 50 observations were generated from

$$X_{j,t} - X_{j,t-1} = a_{j,t} - \theta a_{j,t-1}; \quad j = 1, \ldots, k; X_{j,0} = 100; 0 \leqslant \theta < 1$$
$$Y_t - Y_{t-1} = a_t - \theta^* a_{t-1}; \quad Y_0 = 100; 0 \leqslant \theta^* < 1 \qquad (5.1)$$

where $a_{1,t}, \ldots, a_{k,t}, a_t$ were taken as independent Normally distributed white noise series, all with unit variance. Regression equations of the form (2.5) were then estimated for values $k = 1$, 2 and 4, taking 1,000 replicates for $k = 1$ and 500 for $k = 2$ and 4. The null hypothesis

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0 : \qquad (5.2)$$

was tested using the conventional F test at the 5% significance level and the Durbin–Watson d test carried out at the 5% level. It was found that adoption of the decision rule "reject the null hypothesis (5.2) only if F is significant *and* d is

Table 1
Percentage of times both F is significant and d is insignificant at the 5% level

		$\theta = 0.0$	$\theta = 0.2$	$\theta = 0.4$	$\theta = 0.6$	$\theta = 0.8$
$k = 1$	$\theta^* = 0.6$	6.9	7.7	7.4	6.2	2.4
	$\theta^* = 0.8$	21.2	21.6	22.2	19.0	9.2
$k = 2$	$\theta^* = 0.6$	14.6	15.4	14.0	10.0	3.6
	$\theta^* = 0.8$	27.6	23.6	22.8	17.4	6.8
$k = 4$	$\theta^* = 0.6$	28.2	27.8	28.0	17.4	5.2
	$\theta^* = 0.8$	39.0	35.2	37.2	26.6	14.4

insignificant" would not lead one astray too often except when θ^* in (5.1) is large. Results for this case are shown in Table 1. The difficulty here is that, as shown by Wichern (1973), the expected values of sample autocorrelations of order one for processes of the type (5.1) are not very large for moderate sized θ^*. Hence, it is not reasonable to expect the d statistic to pick up autocorrelation of this type sufficiently frequently.

However, this is only one side of the coin, for if autocorrelation in the errors is detected the equation is re-estimated, correcting for this autocorrelation. It was assumed that the modified equation was based, as is frequently the case in econometric work, on the (false) assumption that the true error structure was first order autoregressive, and an estimation procedure of Cochrane and Orcutt

Table 2
Percentage of times hypothesis of no relationship is rejected at the 5% level when "correcting" for first order autocorrelation in the regression errors

		$\theta = 0.0$	$\theta = 0.2$	$\theta = 0.4$	$\theta = 0.6$	$\theta = 0.8$
$k = 1$	$\theta^* = 0.4$	23.6	22.9	15.9	11.0	8.2
	$\theta^* = 0.6$	32.3	30.7	25.5	19.7	9.8
	$\theta^* = 0.8$	28.9	27.3	25.5	22.7	12.0
$k = 2$	$\theta^* = 0.4$	26.4	25.8	16.4	10.4	5.0
	$\theta^* = 0.6$	40.2	35.6	34.6	22.0	7.6
	$\theta^* = 0.8$	35.4	31.2	27.6	23.2	9.6
$k = 4$	$\theta^* = 0.4$	59.6	54.8	36.0	23.4	10.8
	$\theta^* = 0.6$	66.4	62.2	59.0	37.2	16.8
	$\theta^* = 0.8$	52.6	48.8	59.4	35.8	21.6

(1949) was then used. It emerged from the simulation experiments that the null hypothesis (5.2) was rejected far too frequently. For example, for $k = 4$, a significant relationship was found on about 50% of occasions for a wide range of values of θ and θ^*. Some of the simulation results obtained are summarised in Table 2. It appears then that, looked at from this particular point of view, error mis-specification can have very serious consequences in econometric work and that the usual procedures employed to correct for autocorrelated errors will not always be adequate. The conclusion must be that econometricians should follow the time series analysts' practice of considering a range of error structures wider than the standard white noise/first order autoregression alternative.

It should be added that the econometrician may react differently to the time series analyst on discovering autocorrelated errors. He may, for example, take them as an indication that the structural model is mis-specified, and then try alternative specifications.

6. Forecasting

Once a time series model has been fitted satisfactorily, forecasts are derived from it in a mechanical way. The model is simply projected forward and values of future white noise terms set equal to zero. Algorithms for deriving forecasts from (2.1) and (2.2) are fully discussed by Box and Jenkins (1970), while analagous algorithms for (2.3) and (2.4) follow in an obvious way.

For many econometric exercises the transition from achieved model to published forecast is by no means as straightforward. For convenience we illustrate by considering the linear simultaneous equations model (2.6), which can be written

$$By_t + \Gamma x_t + U_t = 0 \tag{6.1}$$

where

$$y_t' = (Y_{1t} Y_{2t} \ldots Y_{Mt}), \qquad U_t' = (U_{1t} U_{2t} \ldots U_{Mt}),$$

$$x_t' = (X_{1t} X_{2t} \ldots X_{kt}),$$

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} \ldots \beta_{1M} \\ \beta_{21} & \beta_{22} \ldots \beta_{2M} \\ \cdot & \cdot \quad \cdot \quad \cdot \\ \beta_{M1} & \beta_{M2} \ldots \beta_{MM} \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \cdots \gamma_{1k} \\ \gamma_{21} & \gamma_{22} \cdots \gamma_{2k} \\ \cdot & \cdot \quad \cdot \quad \cdot \\ \gamma_{M1} & \gamma_{M2} \cdots \gamma_{Mk} \end{pmatrix}.$$

Multiplying through (6.1) by B^{-1} gives

$$y_t = \Pi x_t + V_t \tag{6.2}$$

where

$$\Pi = -B^{-1}\,\Gamma, \qquad V_t = -B^{-1}U_t.$$

The formulation (6.2), which expresses the endogenous variables directly in terms of the exogenous variables and the error terms, is known as the "reduced form" of the model. As it stands it does not necessarily have any direct economic interpretation but is simply a consequence of the structural form (6.1) imposed by economic theory. Suppose that the present time is denoted as time n, and forecasts are required for time $n + h$. If the model is assumed to hold in the future, then

$$y_{n+h} = \Pi x_{n+h} + V_{n+h}$$

and, *provided* U_t of (6.1), and hence V_t, is not autocorrelated, the best linear forecast of y_{n+h} is Πx_{n+h}. Of course Π is unknown and must be estimated from data. This is almost invariably done by first estimating the structural form (and hence incorporating information provided by economic theory) and estimating Π as

$$\hat{\Pi} = -\hat{B}^{-1}\hat{\Gamma}$$

where $\hat{B}$ and $\hat{\Gamma}$ are the structural form estimators. However, as shown by Dhrymes (1973), these need not necessarily be more efficient than direct estimation of (6.2) by least squares, unless the most efficient structural form estimators are used. Of course, as we noted earlier, the structural form need not be linear in the variables, in which case the forecast of y_{n+h} given x_{n+h} is obtained by setting future error terms in the structural form to zero and solving numerically.

This kind of approach yields directly conditional forecasts; that is, forecasts of what the endogenous variables will be given specified future values of the exogenous variables. In order to obtain unconditional forecasts – estimates of what values the future endogenous variables actually will take – it is necessary to insert forecasts of the exogenous variables x_{n+h}. These cannot, of course, be obtained from the econometric model and, in fact, are usually obtained judgmentally rather than through any formal quantitative analysis. The ability of econometric models to produce conditional forecasts is often urged as one of the great virtues of the structural approach, since to the extent that the exogenous variables are instruments of policy, the policy maker can be presented with estimates of the consequences of alternative options. However, typically not all exogenous variables in a model will be of this type, and the problem of forecasting these non-policy variables remains.

A national economy is a very complex entity indeed. The values taken by macro-economic variables are generally held to arise as a result of a great many interacting mechanisms. The econometrician's aim is to model this behaviour as faithfully as possible, taking into account the fact that the various components of such broad aggregates as consumption and investment may not all be determined in the same way. This has led to the construction of some very large models of national economies. Models of 100–200 equations are now very common, while several models are much larger even than this. (For example, Eckstein, Green and Sinai (1974) describe a model of 698 equations.) For an account of this kind of development, from an enthusiastic insider, see Klein (1971b). The economist is naturally driven to this kind of model by economic theory, but its construction involves formidable conceptual and computational problems, and it should not be surprising to find that short-cuts are taken along the way. Unfortunately the short-cuts taken are those which would give time series analysts greatest cause for alarm. It seems that in many of the larger models treatment of time series problems, particularly those involving error structure, is cursory or non-existent. Indeed it seems inevitable to expect some kind of trade-off between the size of the model and the quality of the time series treatment, if only because time series analysts have great difficulty in handling many series simultaneously. This being the case one wonders whether it might be more fruitful to spend relatively more effort on smaller models, such as those of Hendry (1974), Fair (1970) and Wall et al (1975), where the potential for applying time series considerations has been demonstrated. The opposite view would hold that there is a need to know and understand the many intricacies of economic behaviour, and that this is impossible in a small system involving broad aggregates and excluding many important factors. However, the "knowledge" derived from the larger models could well be illusory, for if a relationship is mis-specified the estimates obtained and conclusions drawn from them must be treated with great scepticism.

In practice, econometric forecasts are rarely derived in the formal way described above. It is generally held by macro-economic model builders that no quantitative model, however large, could be trusted to produce worthwhile forecasts without human aid. An interesting discussion of the typical econometric forecasting exercise is given by Evans, Haitovsky and Treyz (1972). These authors distinguish three separate steps which are often taken in addition to mechanical solution of the fitted structural equations. First, as already noted, it is usual to forecast future values of the exogenous variables judgmentally. Second, before the model is solved, adjustments are made to individual equations. These are generally described either as the setting of non-zero values for the error terms or, equivalently, the adjustment of the intercept terms. The intention here is to incorporate into the forecast any information the economist might have which is not specifically taken account of in the formal model. The amount of modification made is generally determined on judgmental grounds. Occasionally, too, modification is made at this stage in a belated attempt to take into account autocorrelation in the errors of fitted equations (see, for example, Green, Liebenberg and Hirsch (1972)). This may be better than doing nothing, but it is surely preferable to take account of the problem of autocorrelated errors when the model is estimated. The model is now solved to obtain forecasts of the endogenous variables. However, if the forecasts so obtained differ markedly from the economist's a priori assessment of the likely range of future values, he may well decide to modify them, either directly or by modifying the adjustments made to the structural equations and re-solving the model. The heavy use of judgmental adjustments in economic forecasting can be defended on the grounds that ideally the forecaster ought to take into account all the information available at the time the forecast is made, irrespective of whether or not such information can be incorporated into a formal quantitative framework. The quality of the forecasts will then reflect both the quality of the information and the efficiency of its use.

The heavy use of judgment in econometric forecasting causes a problem when one attempts to formally evaluate this work. Should the end-products, the forecasts, simply be judged on their merits, or should one also try to assess the forecasting ability of the econometric model? In fact, evaluation exercises have attempted to do both. A discussion of the methodology of forecast evaluation is given by Granger and Newbold (1973), and a survey of results obtained by major evaluation exercises is contained in Granger and Newbold (1977). Many exercises have regarded econometric models and univariate time series models, such as (2.1), as competitors, and their forecast performances have been directly compared. Broadly speaking the models, unaided by judgment, have not done terribly well (see, for example, Cooper (1972) and Nelson (1972)). However, the actual judgment-aided forecasts have done very much better. This leads to a perplexing question, posed by Evans, Haitovsky and Treyz (1972): "This study has shown that econometricians have had a better forecasting record to date than an analysis of the econometric models they used would have led us to predict. Our results offer no substantive evidence that the same econometricians, forecasting without the 'benefit' of an econometric model, would have done any better or any worse in their prediction."

7. Overview

When two alternative approaches to the same problem gain wide acceptance, it is often the case that an "ideal" solution would incorporate elements of both. We feel strongly that this is the correct view with regard to the classical econometric and time series methodologies as applied to economic model building and forecasting. In contrasting the two approaches and trying to assess their strengths and weaknesses it has not been our aim to set up a contest in which a choice has to be made between one philosophy or the other. Rather, we have tried to show how deficiencies in one procedure might be remedied by incorporating the principles of the other. We believe, then, that much is to be gained from a fuller amalgamation of time series and traditional econometric ideas, and that this represents the most fruitful direction for further research effort currently open to economic model builders.

This notion is certainly not new, and we claim no credit for it. Indeed, as we have seen, econometricians have long been concerned with time series problems, and several attempts have been made to incorporate, for example, more sophisticated time series error structures into econometric models. In addition, econometricians have recently been concerned with the similarities to be found between the structures of the models used by time series analysts and by economists (see, for example, Zellner and Palm (1974) and Wallis (1975)). Much of this interest almost certainly stems from recent developments in time series model building. However, it seems fairly clear that the merger movement still has a long way to go.

It is interesting to speculate on the possible course of future developments. Here one immediately runs up against a fundamental difficulty. We have seen that time series analysts experience tremendous practical problems in trying to build models relating more than a very few series, whereas, in macro-economic model building, many practitioners regard small models as hopelessly unrealistic and incapable of answering many important questions. The instinct of the time series analyst is to work with a manageable structure so that a logical framework can be devised for the solution of the model selection problem. On the other hand, as more computing power, more data and ever more sophisticated economic theories become available, many econometricians instinctively move towards the creation of ever larger models. The temptation, in so doing, is to simply ignore the time series problems, which are virtually insoluble anyway, and in which so many econometric model builders have little interest. We have already seen that much has been done already with small models, and it is tempting to suggest that research effort should be entirely concentrated here, at least initially. However, it is impossible to deny that a market for large models exists, and that these models will continue to be built. It seems unreasonable that the time series analyst should leave these model builders out on a limb. He may not be able to construct a grand design in which a methodology can be neatly outlined, as it has been for univariate and bivariate model building. Nevertheless he may be able to provide help of a more ad hoc kind in an area where it is sorely needed. It is a reasonable guess that model builders will come to employ time series analysts in the next few years as routinely as they presently employ computer programmers.

How might these new employees proceed? First, it seems wholly sensible to take the model structure, as specified by the econometrician, as a starting point. It would be natural, then, to look at the residuals from the individual fitted equations and model these in such a way that a white noise error series is achieved for each equation. A further step would then involve examination of the cross-correlations between the errors from individual equations. This could lead, presumably after a good deal of trial and error, to a multivariate error structure (of an admittedly rather crude kind) in which at least the major inter-relationships were accounted for. The improved forecasting performance which would undoubtedly result should then induce a mood of co-operation in place of the model builder's initial scepticism. The eager time series analyst should quickly exploit this change of heart by opening a dialogue with the econometrician in an attempt to discover why particular lag structures were built into the equations. He may then find that their "theoretical basis" admits the possibility of alternatives which could be further explored. In our ideal world the happy partnership would then go from strength to strength. One useful spin-off, for example, might be the persuasion of economists that more effort should be devoted to developing truly dynamic models of economic behaviour, and that the static theories, which so often form the basis of models, are of themselves insufficient. In addition, it may be realised that a trade-off exists between size of model and quality of time series treatment, and that a workable compromise is needed.

References

Almon, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica*, **33**, 178–196.

Andersen, L. C. and Carlson, K. M. (1974). St. Louis model revisited. *Int. Ec. Rev.*, **15**, 305–327.

Ball, R. J. et al. (1975). The London Business School quarterly econometric model of the U.K. economy. *In:* G. A. Renton, ed., *Modelling the Economy* (Heinemann, London).

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *J. Roy. Statist. Soc. B*, **26**, 211–243.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control* (Holden Day, San Francisco).

Box, G. E. P. and Jenkins, G. M. (1973). Some comments on a paper by Chatfield and Prothero and on a review by Kendall. *J. Roy. Statist. Soc. A*, **136**, 337–345.

Box, G. E. P. and Newbold, P. (1971). Some comments on a paper of Coen, Gomme and Kendall. *J. Roy. Statist. Soc. A*, **134**, 229–240.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *J. Am Statist. Assoc.*, **65**, 1509–1526.

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX model. *J. Econometrics*, **1**, 133–154.

Brown, T. M. (1952). Habit persistence and lags in consumer behaviour. *Econometrica*, **20**, 355–371.

Cagan, P. (1956). The monetary dynamics of hyper-inflations. In: M. Friedman, ed., *Studies in the Quantity Theory of Money* (Chicago University Press).

Chatfield, C. and Prothero, D. L. (1973). Box–Jenkins seasonal forecasting: problems in a case study. *J. Roy. Statist. Soc. A*, **136**, 295–336.

Cochrane, D. and Orcutt, G. H. (1949). Application of least squares regressions to relationships containing autocorrelated error terms. *J. Am. Statist. Assoc.* **44**, 32–61.

Cooper, R. L. (1972). The predictive performance of quarterly econometric models of the United States. In: B. G. Hickman, ed., *Econometric Models of Cyclical Behaviour* (Columbia University Press, New York).

Deaton, A. S. (1974). The analysis of consumer demand in the United Kingdom 1900–1970, *Econometrica*, **42**, 341–367.

Desai, M. (1975). The share of wages in national income: a closed nonlinear model for the U.K. economy 1855–1970. Paper presented at World Congress of Econometric Society, Toronto.

Dhrymes, P. J. (1971). *Distributed Lags: Problems of Formulation and Estimation*. (Holden Day, San Francisco).

Dhrymes, P. J. (1973). Restricted and unrestricted reduced forms: asymptotic distribution and relative efficiency. *Econometrica*, **41**, 119–134.

Dhrymes, P. J. et al. (1972). Criteria for evaluation of econometric models. *Ann. Ec. Soc. Meas.*, **1**, 291–324.

Durbin, J. (1970). Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables. *Econometrica*, **38**, 410–421.

Durbin, J. and Watson, G. S. (1950, 1951). Testing for serial correlation in least squares regression, I, II. *Biometrika*, **37**, 409–428; **38**, 159–178.

Eckstein, O., Green, E. W. and Sinai, A. (1974). The Data Resources Model: uses, structure and analysis of the U.S. economy. *Int. Ec. Rev.*, **15**, 595–615.

Eisner, R. and Strotz, R. (1963). The determinants of business investment. In: D. B. Suits et al., *Impacts of Monetary Policy* (Prentice-Hall, New Jersey).

Evans, M. K., Haitovsky, Y. and Treyz, G. I. (1972). An analysis of the forecasting properties of U.S. econometric models. In: B. G. Hickman, ed., *Econometric Models of Cyclical Behaviour* (Columbia University Press, New York).

Fair, R. C. (1970). *A Short-Run Forecasting Model of the United States Economy.* (D. C. Heath, Lexington).

Fisher, F. M. (1966). *The Identification Problem.* (McGraw-Hill, New York).

Granger, C. W. J. (1966). The typical spectral shape of an economic variable. *Econometrica,* **34**, 150–161.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica,* **37**, 424–438.

Granger, C. W. J. and Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *App. Econ.*, **5**, 35–47.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *J. Econometrics,* **2**, 111–120.

Granger, C. W. J. and Newbold, P. (1976a). Forecasting transformed series. *J. Roy. Statist. Soc B*, **38**. 189–203.

Granger, C. W. J. and Newbold, P. (1976b). Identification of two-way causal systems. In: M. D. Intriligator, ed., *Frontiers of Quantitative Economics III* (North Holland, Amsterdam).

Granger, C. W. J. and Newbold, P. (1977). *Forecasting Economic Time Series.* (Academic Press, New York).

Green, G. R., Liebenberg, M. and Hirsch, A. A. (1972). Short- and long-term simulations with the O.B.E. econometric model. In: B. G. Hickman, ed., *Econometric Models of Cyclical Behaviour* (Columbia University Press, New York).

Griliches, Z. (1967). Distributed lags: a survey. *Econometrica,* **35**, 16–49.

Hannan, E. J. (1969). The identification of vector mixed autoregressive moving average systems. *Biometrika,* **56**, 223–225.

Hannan, E. J. (1971). The identification problem for multiple equation systems with moving average errors. *Econometrica,* **39**, 751–765.

Hatanaka, M. (1975). On the global identification of the dynamic simultaneous equations model with stationary disturbances. *Int. Ec. Rev.*, **16**, 545–554.

Hendry, D. F. (1974). Stochastic specification in an aggregate demand model of the United Kingdom. *Econometrica,* **42**, 559–578.

Hendry, D. F. (1975). On the time series approach to econometric model building. Paper presented at Conference on New Methods of Business Cycle Research, Minneapolis.

Hendry, D. F. (1976). The structure of simultaneous equations estimators. *J. Econometrics,* **4**, 51–88.

Hendry, D. F. and Trivedi, P. K. (1972). Maximum likelihood estimation of difference equations with moving average errors: a simulation study. *Rev. Econ. Studies,* **39**, 117–145.

Howrey, E. P., Klein, L. R. and McCarthy, M. D. (1974). Notes on testing the predictive performance of econometric models. *Int. Ec. Rev.*, **15**, 366–383.

Jenkins, G. M. (1974). The interaction between the muskrat and mink cycles in North Canada. Paper presented at International Biometric Conference, Constanta, Roumania.

Johnston, J. (1972). *Econometric Methods* (2nd ed.). (McGraw-Hill, New York).

Jorgenson, D. W. (1966). Rational distributed lag functions. *Econometrica,* **34**, 135–149.

Jorgenson, D. W. Hunter, J. and Nadiri, M. I. (1970). A comparison of alternative econometric models of quarterly investment behaviour. *Econometrica,* **38**, 187–212.

Jorgenson, D. W. and Siebert, C. D. (1968). A comparison of alternative theories of corporate investment behaviour. *Am. Ec. Rev.*, 58, 681–712.

Jorgenson, D. W. and Stephenson, J. A. (1967). The time structure of investment behaviour in United States manufacturing 1947–1960. *Rev. Econ. Stat.* 49, 16–27.

Klein, L. R. (1971a). *An Essay on the Theory of Economic Prediction.* (Markham, Chicago).

Klein, L. R. (1971b). Forecasting and policy evaluation using large scale econometric models: the state of the art. In: M. D. Intriligator, ed., *Frontiers of Quantitative Economics* (North Holland, Amsterdam).

Klein, L. R. and Ball, R. J. (1959). Some econometrics of the determination of absolute prices and wages. *Econ. J.*, 69, 465–482.

Koyck, L. M. (1954). *Distributed Lags and Investment Analysis.* (North Holland, Amsterdam).

Mundlak, Y. (1966). On the microeconomic theory of distributed lags. *Rev. Econ. Stat.*, 48, 51–60.

Mundlak, Y. (1967). Long-run coefficients and distributed lag analysis: a reformulation. *Econometrica*, 35, 278–293.

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *J. Am. Statist. Assoc.*, 55, 299–306.

Nelson, C. R. (1972). The prediction performance of the FRB-MIT-PENN model of the U.S. economy. *Am. Ec. Rev.*, 62, 902–917.

Nerlove, M. (1956). Estimates of the elasticities of supply of selected agricultural commodities. *J. Farm. Econ.*, 38, 496–509.

Nerlove, M. (1958). *Distributed Lags and Demand Analysis for Agricultural and other Commodities.* U.S.D.A., Agriculture Handbook 141.

Nerlove, M. and Wallis, K. F. (1966). The use of the Durbin–Watson statistic in inappropriate situations. *Econometrica*, 34, 235–238.

Newbold, P. (1973). Bayesian estimation of Box–Jenkins transfer function–noise models. *J. Roy. Statist. Soc. B*, 35, 323–336.

Newbold, P. (1974). The exact likelihood function for a mixed autoregressive-moving average process. *Biometrika*, 61, 423–426.

Newbold, P. and Davies, N. (1978). Error mis-specification and spurious regressions. *Int Econ Rev*, 19, 513–519.

Newbold, P. and Granger, C. W. J. (1974). Experience with forecasting univariate time series and the combination of forecasts. *J. Roy. Statist. Soc. A*, 137, 131–146.

Peseran, M. H. (1974). On the general problem of model selection, *Rev. Econ. Studies*, 41, 153–171.

Phillips, A. W. (1966). Estimation of systems of difference equations with moving average disturbances. Paper presented at Meeting of Econometric Society, San Francisco.

Pierce, D. A. (1977). Relationships – and the lack thereof – between economic time series, with special reference to money, reserves and interest rates. *J. Am. Statist. Assoc.*, 72, 11–21.

Prothero, D. L. and Wallis, K. F. (1976). Modelling macroeconomic time series (with Discussion). *J. Roy. Statist. Soc. A*, 139, 468–500.

Quenouille, M. H. (1957). *The Analysis of Multiple Time Series* (Griffin, London).

Ramsey, J. B. (1969). Tests for specification errors in classical least squares regression analysis. *J. Roy. Statist. Soc. B*, 31, 350–371.

Ramsey, J. B. (1970). Models specification error and inference: a discussion of some problems in econometric methodology. *Bull. Oxford Inst. Econ.*, 32, 301–318.

Ramsey, J. B. and Zarembka, P. (1971). Specification error tests and alternative functional forms of the aggregate production function. *J. Am. Statist. Assoc.*, **66**, 471–477.

Reid, D. J. (1966). A comparative study of time series prediction techniques on economic data. Ph.D. thesis, Department of Mathematics, University of Nottingham.

Sargan, J. D. (1964). Wages and prices in the United Kingdom: a study in econometric methodology. In: P. E. Hart, G. Mills and J. K. Whitaker, eds., *Econometric Analysis for National Economic Planning* (Butterworth, London).

Solow, R. (1957). Technical change and the aggregate production function. *Rev. Econ. Stat.*, **39**, 312–320.

Trivedi, P. K. (1970). Inventory behaviour in U.K. manufacturing 1956–67. *Rev. Econ. Studies*, **37**, 517–536.

Trivedi, P. K. (1975). Time series versus structural models: a case study of Canadian manufacturing inventory behaviour. *Int. Ec. Rev.*, **16**, 587–608.

Wall, K. D., Preston, A. J., Bray, J. W. and Peston, M. H. (1975). Estimates of a simple control model of the U.K. economy. In: G. A. Renton, ed., *Modelling the Economy* (Heinemann, London).

Wallis, K. F. (1972). Testing for fourth order autocorrelation in quarterly regression equations. *Econometrica*, **40**, 617–636.

Wallis, K. F. (1975). Testing dynamic specification from the final form. Paper presented at World Congress of Econometric Society, Toronto.

Wichern, D. W. (1973). The behaviour of the sample autocorrelation function for an integrated moving average process. *Biometrika*, **60**, 235–239.

Wilson, G. T. (1973). The estimation of parameters in multivariate time series models. *J. Roy. Statist. Soc. B*, **35**, 76–85.

Zellner, A. and Geisel, M. S. (1970). Analysis of distributed lag models with applications to consumption function estimation. *Econometrica*, **38**, 865–888.

Zellner, A. and Palm, F. (1974). Time series analysis and simultaneous equation econometric models. *J. Econometrics*, **2**, 17–54.

O.D. Anderson & M.R. Perryman, eds., TIME SERIES ANALYSIS

THE TREATMENT OF EXTREME VALUES IN THE X-11-ARIMA PROGRAM

John Kuiper

Department of Economics
University of Ottawa
Ottawa, Canada

In this study, the identification and treatment of extremes is analyzed in the context of the X-11-Arima seasonal adjustment method. It is shown that for some series over-adjustment for extremes takes place in the Box-Jenkins step. For other series, the Box-Jenkins step may still be used to identify outliers, even if extrapolation of the time series is inappropriate. Finally, it is shown that the seasonal factors obtained for the last three years of the time series using the X-11-Arima method are more reliable than those calculated using the regular X-11 method.

INTRODUCTION

The X-11-Arima method of seasonal adjustment consists of the application of the regular X-11 program to a time series which has been extended by one year using a Box-Jenkins model. The X-11 method is described in Shiskin, Young and Musgrave (1967) and the X-11-Arima method in Dagum (1980). For the Box-Jenkins analysis, see Box and Jenkins (1970).

X-11-Arima also has an option to replace outliers by the fitted values obtained as part of the Box-Jenkins step.

TIME SERIES USED

The following time series are analyzed in this paper:

Identifier	Description
USE	Total employed - United States
USU	Total unemployed - United States
11134	Money - United States
11234	Money - United Kingdom
13234	Money - France
13270	Exports FOB - France
13271	Imports CIF - France
13834	Money - Netherlands
15632	Domestic Credit - Canada

All time series are monthly series for the period January 1957 to December 1978. The two labour force series were obtained from the Survey of Current Business and the other time series from the International Financial Statistics published by the International Monetary Fund.

THE X-11-ARIMA METHOD OF SEASONAL ADJUSTMENT

The X-11-Arima program automatically selects the model which performs best from among three models which have been found to fit a wide variety of time series. The selection criteria are the average forecast error (which must be less than 12%), the average backcast error (which must be less than 18%) and the χ^2 probability (which must be greater than 10%).

The Box-Jenkins models selected for the nine time series and their performance on the selection criteria were as follows:

Series	Model Selected	χ^2	Avg. % Error
USE	(2,1,2) (0,1,1)	11.4	1.07
USU	Log (0,2,2) (0,1,1)	47.6	9.72
11134	Log (0,1,1) (0,1,1)	20.3	1.06
11234	Log (0,1,1) (0,1,1)	43.9	1.90
13234	Log (0,1,1) (0,1,1)	23.1	1.04
13270	Log (0,1,1) (0,1,1)	7.5	5.31
13271	(2,1,2) (0,1,1)	13.3	8.97
13834	Log (0,1,1) (0,1,1)	52.1	4.33
15632	Log (0,1,1) (0,1,1)	78.1	0.68

The time series analyzed all passed the selection criteria, except for series 13270 which failed the χ^2 test.

CALCULATIONS PERFORMED

The Box-Jenkins models selected were used to extrapolate the time series by 12 months. Because the time series cover a sufficiently long period, the addition of one year's data at the beginning of the series through a backcast was not considered necessary.

The X-11 program has an option called 'prior adjustment' which allows the user to replace extreme values before the seasonal adjustment is performed. The prior adjustment option should be used if, for example, monthly production, shipments or foreign trade statistics are affected by major industrial disputes. Ideally, these prior adjustments should be made by subject matter specialists. However, due to the large number of series which are currently adjusted for seasonality, it is useful to consider mechanical procedures.

Such a procedure is readily available in X-11-Arima because, in the X-11-Arima method, a Box-Jenkins model is fitted to the time series, and it is therefore possible to identify outliers. The program includes an option to modify the time series by replacing those observations with the corresponding fitted value. Observations are considered to be an outlier if the error exceeds 2.5 σ. (See Dagum [1980], p.29)

To determine the effects of the extension of the time series using the fitted model and of 'prior adjustment' separately as compared to the combined effect of both these modifications to the regular X-11 program, the following alternative sets of calculations were performed.

Method	Description
X-11-Arima AR1	The time series is extrapolated 12 months, based on the Box-Jenkins model selected, without modifying extreme values. The X-11 program is applied to the extrapolated original series.
X-11-Arima AR2	The appropriate Box-Jenkins model is fitted to the original time series. The model is then re-fitted to the time series modified for outliers. The time series is then extrapolated 12 months. The X-11 program is applied to the extrapolated original series.
X-11-Arima AR3	The appropriate Box-Jenkins model is fitted to the original time series. The model is then re-fitted to the time series modified for outliers. The time series is then extrapolated 12 months. The X-11 program is applied to the extrapolated modified series.
X-11-Arima AR4	The appropriate Box-Jenkins model is fitted to the original time series. The model is then re-fitted to the time series modified for outliers. The X-11 program is applied to the modified series, but the series is not extrapolated.
X-11 Regular	The original series is adjusted using the regular X-11 program.

As an overall measure of the extent to which the series have been smoothed by seasonal adjustment, the average absolute month-to-month percentage change has been used. These percentage changes are presented in Table I for the full period analyzed, i.e., 1957-1978; for the sub-period 1960-1975, i.e., excluding the first and last three years; for the three year period 1976-1978 and for the year 1978.

For the period 1960-1975, i.e., the time span for which two-sided filters could be applied, options AR1 and AR2 do not result in significant reductions in this summary measure when compared with the regular X-11 results. However, options AR3 and AR4 resulted in small increases in the mean absolute percentage change, except for series 11134, 13271 and 15632 which improved marginally.

For the period 1976-1978, the X-11-Arima program performed better than the regular X-11 program. This is to be expected because the time series has been extrapolated. It should be noted that options AR1 and AR2 performed best for most series. The exceptions are series 11234 and 13270 for which option AR3 resulted in the lowest mean absolute percentage change. Note, however, that option AR3 performed worse when applied to the other series. This implies that for these series, it would have been preferable not to use the 'prior adjustment' option of X-11 or, in other words, to use the original time series together with the extrapolated values obtained from the Box-Jenkins step.

Option AR4 would be appropriate if the average forecast error were above the selection criteria (i.e., 12% for forecasts and 18% for backcasts), but the Box-Jenkins model fits the overall time series well.

In this case, the selection criteria should be extended to include, for example, a 6% average error of estimate, which would trigger the use of the modified series, even in cases where extrapolation is not indicated.

For other studies in which comparisons between various methods of seasonal adjustment were undertaken, see Fase, Koning and Volgenant (1973) and Kuiper (1978).

TABLE I

MEAN ABSOLUTE PERCENTAGE CHANGES
SEASONALLY ADJUSTED SERIES

	1957-78	1960-75	1976-78	1978
Series USE:				
X-11 Regular	0.316	0.296	0.342	0.373
X-11-Arima (AR1)	0.313	0.295	0.322	0.330
X-11-Arima (AR2)	0.313	0.295	0.322	0.331
X-11-Arima (AR3)	0.312	0.300	0.332	0.339
X-11-Arima (AR4)	0.314	0.300	0.345	0.379
Unadjusted Series	0.918	0.915	0.838	0.729
Series USU:				
X-11 Regular	3.194	3.110	2.075	2.477
X-11-Arima (AR1)	3.184	3.106	2.006	2.142
X-11-Arima (AR2)	3.191	3.108	2.044	2.398
X-11-Arima (AR3)	3.384	3.335	2.223	2.648
X-11-Arima (AR4)	3.399	3.338	2.333	2.850
Unadjusted Series	8.649	9.012	6.574	5.533
Series 11134:				
X-11 Regular	0.619	0.590	1.028	0.966
X-11-Arima (AR1)	0.618	0.593	1.005	0.816
X-11-Arima (AR2)	0.619	0.593	1.014	0.915
X-11-Arima (AR3)	0.622	0.591	1.044	1.024
X-11-Arima (AR4)	0.623	0.591	1.057	1.057
Unadjusted Series	1.943	1.986	2.373	1.660
Series 11234:				
X-11 Regular	1.028	0.979	1.638	1.341
X-11-Arima (AR1)	1.026	0.978	1.623	1.229
X-11-Arima (AR2)	1.028	0.978	1.634	1.271
X-11-Arima (AR3)	1.061	1.022	1.609	1.285
X-11-Arima (AR4)	1.059	1.024	1.593	1.308
Unadjusted Series	1.264	1.142	1.942	1.652
Series 13234:				
X-11 Regular	1.045	1.116	0.976	1.171
X-11-Arima (AR1)	1.046	1.117	0.975	1.103
X-11-Arima (AR2)	1.049	1.118	0.997	1.186
X-11-Arima (AR3)	1.072	1.152	1.010	1.166
X-11-Arima (AR4)	1.068	1.150	0.994	1.195
Unadjusted Series	1.994	2.026	2.602	2.793
Series 13270:				
X-11 Regular	4.690	4.833	3.111	3.535
X-11-Arima (AR1)	4.683	4.836	3.037	3.262
X-11-Arima (AR2)	4.680	4.836	3.016	3.242
X-11-Arima (AR3)	4.756	4.879	3.014	3.242
X-11-Arima (AR4)	4.763	4.872	3.101	3.580
Unadjusted Series	10.962	11.039	11.908	12.818

TABLE I (cont.)

	1957-78	1960-75	1976-78	1978
Series 13271:				
X-11 Regular	5.381	5.904	3.517	2.424
X-11-Arima (AR1)	5.349	5.904	3.272	1.862
X-11-Arima (AR2)	5.381	5.902	3.527	2.456
X-11-Arima (AR3)	5.500	5.845	4.700	4.136
X-11-Arima (AR4)	5.556	5.834	5.187	4.664
Unadjusted Series	9.947	10.604	8.082	8.025
Series 13834:				
X-11 Regular	1.054	1.032	1.447	1.305
X-11-Arima (AR1)	1.052	1.033	1.416	1.120
X-11-Arima (AR2)	1.051	1.033	1.406	1.148
X-11-Arima (AR3)	1.074	1.052	1.477	1.059
X-11-Arima (AR4)	1.093	1.063	1.567	1.192
Unadjusted Series	1.689	1.614	2.617	2.619
Series 15632:				
X-11 Regular	1.107	1.088	1.440	1.637
X-11-Arima (AR1)	1.107	1.091	1.430	1.616
X-11-Arima (AR2)	1.106	1.088	1.436	1.748
X-11-Arima (AR3)	1.091	1.061	1.500	1.748
X-11-Arima (AR4)	1.086	1.062	1.459	1.693
Unadjusted Series	1.555	1.584	1.638	1.916

THE TREATMENT OF EXTREMES

The X-11 program includes an elaborate procedure to identify and modify extremes. Extremes which fall above 2.5 σ are replaced, and those that fall between 1.5 and 2.5 σ are partially replaced.

The adjustments for extremes made by the X-11 program were analyzed for the nine time series. The results are presented in Table II. For series 11134, 11234, 13270 and 13834, option AR3 significantly reduced the mean absolute values of the correction for extremes made by the X-11 program. This indicates that the 'prior adjustments' for this series were consistent with those that resulted from applying the X-11 routine to the unmodified series.

However, for series USU and series 15632, the X-11 adjustments for extremes are larger for the X-11-Arima options. This indicates that the 'prior adjustments' resulted in an over-adjustment for extremes. In these cases, X-11-Arima option AR1 should be selected, i.e., the series should not be modified.

STABILITY OF THE SEASONAL COMPONENT

One of the important properties which may be used to rank seasonal adjustment procedures, is the extent to which the preliminary estimates of the seasonal factors diverge from the final estimate, i.e., the extent to which the seasonal component changes when additional observations become available.

To measure this property, the seasonal factors for the last three years of the time series were compared with the seasonal factors calculated when three additional years of data have become available. At that time, the seasonal factors may be considered final in a statistical sense.

TABLE II

EXTREME VALUES
1976-1978

	Number	Mean Absolute Value	Mean Squared Value
Series USE:			
X-11 Regular	3	0.204	0.316
X-11-Arima (AR1)	3	0.168	0.237
X-11-Arima (AR2)	3	0.166	0.237
X-11-Arima (AR3)	2	0.326	0.335
X-11-Arima (AR4)	4	0.176	0.288
Series USU:			
X-11 Regular	4	0.089	0.120
X-11-Arima (AR1)	5	0.103	0.125
X-11-Arima (AR2)	5	0.118	0.139
X-11-Arima (AR3)	6	0.162	0.181
X-11-Arima (AR4)	3	0.265	0.290
Series 11134:			
X-11 Regular	8	2.817	3.686
X-11-Arima (AR1)	9	2.617	3.371
X-11-Arima (AR2)	11	2.185	3.167
X-11-Arima (AR3)	8	1.823	2.077
X-11-Arima (AR4)	7	1.833	2.365
Series 11234:			
X-11 Regular	6	285.9	322.1
X-11-Arima (AR1)	9	229.6	268.3
X-11-Arima (AR2)	9	224.0	267.9
X-11-Arima (AR3)	10	215.2	287.4
X-11-Arima (AR4)	6	321.9	350.2
Series 13234:			
X-11 Regular	3	1.414	1.951
X-11-Arima (AR1)	6	3.022	3.644
X-11-Arima (AR2)	6	3.087	3.763
X-11-Arima (AR3)	5	2.914	3.786
X-11-Arima (AR4)	3	0.800	0.858
Series 13270:			
X-11 Regular	3	2075.7	2103.2
X-11-Arima (AR1)	5	1261.1	1556.8
X-11-Arima (AR2)	5	1244.6	1533.9
X-11-Arima (AR3)	5	1244.2	1533.6
X-11-Arima (AR4)	3	2084.9	2110.6
Series 13271:			
X-11 Regular	1	730.0	730.0
X-11-Arima (AR1)	2	1085.7	1352.5
X-11-Arima (AR2)	2	1169.0	1613.4
X-11-Arima (AR3)	4	1273.4	1332.8
X-11-Arima (AR4)	5	761.8	1049.4

TABLE II (cont.)

	Number	Mean Absolute Value	Mean Squared Value
Series 13834:			
X-11 Regular	7	0.917	1.042
X-11 Arima (AR1)	8	0.896	1.002
X-11 Arima (AR2)	10	0.756	0.945
X-11 Arima (AR3)	9	0.446	0.559
X-11 Arima (AR4)	5	0.630	0.757
Series 15632:			
X-11 Regular	6	0.563	0.689
X-11-Arima (AR1)	7	0.619	0.727
X-11-Arima (AR2)	6	0.693	0.842
X-11-Arima (AR3)	7	0.591	0.721
X-11-Arima (AR4)	7	0.682	0.824

Tables III and IV show respectively the mean algebraic percentage differences and the mean absolute percentage differences of the seasonal factors for the nine time series studied. These results are averages for the period January 1967 to December 1975, which is a sufficiently large period to exclude random fluctuations.

These two measures are presented for the seasonal factor forecasts, the concurrent seasonal factors, the first revised and the second revised seasonal factors. The concurrent seasonal factors are generated from a series that includes the last available observations, whereas the seasonal factor forecasts are obtained from a series that ended one year before. The first revised, i.e., the seasonal factors calculated when one additional year of observations has become available, and the second revised, calculated when two additional years of observations have become available, are included in these Tables to indicate the extent to which the seasonal factors approach the 'final' factors.

The mean algebraic percentage differences of the seasonal factors are a measure of the bias in the preliminary estimates. The X-11-Arima method shows a significant reduction in bias for most of the series. The magnitude is approximately 30%. An exception is series 15632, while series 13270 only showed a marginal improvement. There is no apparent difference between the various X-11-Arima options studied. Note that for this analysis, option AR4 was not available.

The mean absolute percentage differences of the seasonal factors are a measure of the dispersion of the preliminary estimates. The X-11-Arima method results in a significant reduction in this measure. An exception is series 13271, while series 13234 and 13270 only show small improvements. Overall, the order of magnitude of the reduction is approximately 30%. Similar results were reported by Dagum, [1978], p. 50.

The seasonal factor forecasts obtained using the X-11-Arima method, i.e., seasonal factors for the last year of the extrapolated time series, are thus more reliable than the seasonal factors calculated using the regular X-11 method. The results for the first revised and second revised seasonal factors show, as expected, that the differences with the regular X-11 method gradually diminished.

TABLE III

MEAN ALGEBRAIC PERCENTAGE DIFFERENCES
SEASONAL FACTORS

	Forecast	Concurrent	1st revised	2nd revised
Series USE:				
X-11 Regular	0.086	0.056	0.050	0.037
X-11-Arima (AR1)	0.061	0.042	0.029	0.018
X-11-Arima (AR2)	0.054	0.035	0.025	0.017
X-11-Arima (AR3)	0.058	0.042	0.026	0.017
Series USU:				
X-11 Regular	0.821	0.587	0.344	0.107
X-11-Arima (AR1)	0.674	0.471	0.258	0.100
X-11-Arima (AR2)	0.691	0.449	0.215	0.089
X-11-Arima (AR3)	0.703	0.450	0.227	0.111
Series 11134:				
X-11 Regular	0.233	0.141	0.060	0.046
X-11-Arima (AR1)	0.176	0.089	0.050	0.033
X-11-Arima (AR2)	0.189	0.107	0.048	0.030
X-11-Arima (AR3)	0.196	0.107	0.056	0.028
Series 11234:				
X-11 Regular	0.286	0.191	0.102	0.054
X-11-Arima (AR1)	0.255	0.160	0.094	0.037
X-11-Arima (AR2)	0.252	0.165	0.103	0.037
X-11-Arima (AR3)	0.211	0.137	0.077	0.025
Series 13234:				
X-11 Regular	0.367	0.220	0.122	0.051
X-11-Arima (AR1)	0.330	0.185	0.103	0.047
X-11-Arima (AR2)	0.312	0.178	0.103	0.053
X-11-Arima (AR3)	0.261	0.163	0.093	0.050
Series 13270:				
X-11 Regular	0.741	0.506	0.317	0.164
X-11-Arima (AR1)	0.844	0.653	0.466	0.173
X-11-Arima (AR2)	0.737	0.499	0.326	0.122
X-11-Arima (AR3)	0.764	0.518	0.345	0.139
Series 13271:				
X-11 Regular	2.228	1.722	1.030	0.427
X-11-Arima (AR1)	2.313	1.859	1.210	0.531
X-11-Arima (AR2)	2.219	1.716	0.998	0.480
X-11-Arima (AR3)	2.090	1.597	0.919	0.419
Series 13834:				
X-11 Regular	0.523	0.370	0.222	0.100
X-11-Arima (AR1)	0.417	0.291	0.168	0.052
X-11-Arima (AR2)	0.412	0.293	0.173	0.057
X-11-Arima (AR3)	0.376	0.282	0.166	0.060
Series 15632:				
X-11 Regular	0.096	0.077	0.079	0.052
X-11-Arima (AR1)	0.117	0.090	0.061	0.018
X-11-Arima (AR2)	0.125	0.095	0.058	0.018
X-11-Arima (AR3)	0.172	0.131	0.075	0.015

TABLE IV

MEAN ABSOLUTE PERCENTAGE DIFFERENCES
SEASONAL FACTORS

	Forecast	Concurrent	1st revised	2nd revised
Series USE:				
X-11 Regular	0.188	0.140	0.099	0.061
X-11-Arima (AR1)	0.140	0.111	0.078	0.041
X-11-Arima (AR2)	0.137	0.107	0.077	0.041
X-11-Arima (AR3)	0.137	0.109	0.078	0.042
Series USU:				
X-11 Regular	2.042	1.498	0.911	0.417
X-11-Arima (AR1)	1.418	1.164	0.806	0.416
X-11-Arima (AR2)	1.470	1.111	0.717	0.373
X-11-Arima (AR3)	1.430	1.090	0.680	0.364
Series 11134:				
X-11 Regular	0.320	0.234	0.137	0.077
X-11-Arima (AR1)	0.276	0.188	0.121	0.087
X-11-Arima (AR2)	0.292	0.198	0.128	0.085
X-11-Arima (AR3)	0.305	0.220	0.135	0.095
Series 11234:				
X-11 Regular	0.619	0.462	0.283	0.146
X-11-Arima (AR1)	0.537	0.392	0.256	0.150
X-11-Arima (AR2)	0.555	0.417	0.266	0.147
X-11-Arima (AR3)	0.501	0.394	0.257	0.160
Series 13234:				
X-11 Regular	0.562	0.438	0.287	0.151
X-11-Arima (AR1)	0.529	0.414	0.268	0.145
X-11-Arima (AR2)	0.576	0.438	0.303	0.182
X-11-Arima (AR3)	0.581	0.444	0.285	0.180
Series 13270:				
X-11 Regular	1.415	1.119	0.752	0.373
X-11-Arima (AR1)	1.502	1.171	0.747	0.378
X-11-Arima (AR2)	1.401	1.060	0.675	0.336
X-11-Arima (AR3)	1.363	1.080	0.704	0.385
Series 13271:				
X-11 Regular	3.015	2.662	1.991	1.117
X-11-Arima (AR1)	3.061	2.626	1.917	1.216
X-11-Arima (AR2)	3.110	2.684	1.956	1.100
X-11-Arima (AR3)	3.190	2.821	2.451	1.573
Series 13834:				
X-11 Regular	0.653	0.506	0.333	0.170
X-11-Arima (AR1)	0.627	0.469	0.330	0.197
X-11-Arima (AR2)	0.634	0.474	0.324	0.181
X-11-Arima (AR3)	0.580	0.480	0.344	0.209
Series 15632:				
X-11 Regular	0.281	0.191	0.120	0.077
X-11-Arima (AR1)	0.208	0.163	0.121	0.077
X-11-Arima (AR2)	0.206	0.173	0.140	0.084
X-11-Arima (AR3)	0.246	0.193	0.130	0.075

CONCLUSION

This study has analyzed the application of the X-11-Arima method to nine time series from various countries.

It was found that the application of the X-11-Arima method results in more reliable seasonal factors for the last three years of the time series. In addition, forecasted factors are also superior.

Secondly, the Box-Jenkins model which has been identified for the time series, should only be used to modify the original time series if it has been found that for the particular time series no over-adjustment for extremes results.

Finally, it is suggested that even in cases where it is inappropriate to extrapolate the time series, the Box-Jenkins step may still be useful to identify outliers to be excluded as 'prior adjustments' from the X-11 seasonal adjustment run.

REFERENCES

[1] Box, G.E.P. and Jenkins, G.M., Time Series Analysis, Forecasting and Control (Holden Day, San Francisco, 1970).

[2] Dagum, E.B., A Comparison and Assessment of Seasonal Adjustment Methods for Employment and Unemployment Statistics, Background Paper No. 5, (National Commission on Employment and Unemployment Statistics, Washington, D.C., February 1978).

[3] Dagum, E.B., The X-11-Arima Seasonal Adjustment Method, Catalogue 12-564E Occasional (Statistics Canada, 1980).

[4] Fase, M.M.G., Koning, J., and Volgenant, F., An Experimental Look at Seasonal Adjustment, De Economist, 121 (1973) 441-480.

[5] Kuiper, J., A Survey and Comparative Analysis of Various Methods of Seasonal Adjustment, Proceedings of the NBER/Bureau of the Census Conference on Seasonal Analysis of Economic Time Series, Arnold Zellner (editor) (Washington, D.C., 1978).

[6] Shiskin, J., Young, H. and Musgrave, J.C., The X-11 Variant of Census Method II Seasonal Adjustment, Technical Paper No. 15 (Bureau of the Census, U.S. Department of Commerce, Washington, D.C., 1967).

Detecting Outliers in Time Series Data

MICHAEL R. CHERNICK, DARRYL J. DOWNING, and DAVID H. PIKE*

Occasional large errors in data can have drastic effects on estimates for such quantities as correlation coefficients, regression coefficients, and spectral density estimates. In this article we investigate the effect of outliers on time series data by considering the influence function for the autocorrelations $p(k)$ of a stationary time series. This influence function matrix is applied to simulated data, to power plant data, and to inventory data on nuclear materials.

KEY WORDS: Autocorrelation; Influence function; Stationary time series; Outliers.

1. INTRODUCTION

Occasional large errors in data can have drastic effects on estimates for such quantities as correlation coefficients, regression coefficients, and spectral density estimates. Devlin, Gnanadesikan, and Kettenring (1975) show that the influence function for bivariate correlation can be a useful tool for detecting outliers. Gnanadesikan (1977) gives a discussion of this method and other useful multivariate methods. A formula for the influence function for bivariate correlation is given in Gnanadesikan (1977) and is attributed to Mallows.

In Section 2, we define an influence function matrix for the autocorrelations and in Section 3 show how it can be used as a graphical tool for detecting outliers as is illustrated on the monthly power plant series, the nuclear materials inventory data, and the simulation of an example of Miller.

One widely used technique of model fitting for time series is the identification and estimation of parameters for integrated autoregressive moving average models. This approach was made popular by Box and Jenkins (1970), who proposed a systematic method for model identification, parameter estimation, and diagnostic checking. The model identification part relies heavily on the behavior of the sample autocorrelation function. Since individual outliers can have dramatic effects on several correlations, particularly for relatively short series, this may have implications on the identification phase of the Box and Jenkins method when errors are suspected in the series. In such instances the user may want to compute the influence function matrix for the data set before starting the Box and Jenkins method.

The detection of outliers in a multivariate or time series setting is more complex than in the case of uncorrelated univariate data, since an outlier may not be simply an extremely large or small observation. With large data bases it may be fruitful to focus attention on the types of outliers that have large influence on estimators of interest to users of the data.

In this article we study the use of the influence function for the estimator of the autocorrelation function of a time series, one of the most important functions studied in time series analysis.

2. THE INFLUENCE FUNCTION MATRIX

The influence function for an estimate depends on the parameters being estimated, the observation vector whose influence is being measured, and the distribution function of that observation vector. The parameter can be considered as a functional of the distribution function F and is commonly written $T(F)$. Often the estimator under consideration can be expressed as $T(F_m)$, where F_m is the empiric distribution function (i.e., $F_m(\mathbf{y})$ is the fraction of the m sample vectors $\mathbf{x}$ whose coordinates are all less than or equal to the coordinates of $\mathbf{y}$). The influence function as defined by Hampel (1974) is given by the following equation when the limit on the right side exists:

$$I(F, T(F), \mathbf{x}) = \lim_{\epsilon \to 0} \{T((1 - \epsilon) F + \epsilon \delta \mathbf{x}) - T(F)\}/\epsilon. \tag{2.1}$$

In (2.1), $\mathbf{x}$ is the point of interest in the observation space, ϵ is a positive real number, and $\delta \mathbf{x}$ is the distribution function that has all its probability mass concentrated at the point $\mathbf{x}$. For many practical applications the influence function is well defined.

Let $p(k)$ denote the autocorrelation at lag k for a stationary time series $\{X_t\}_{t=1}^{\infty}$. Let $\mu = E(X_t)$ and $\sigma^2 = \text{var}(X_t)$. It is convenient to let $Y_t = (X_t - \mu)/\sigma$ for each t. This transformation from X_t to Y_t does not affect $p(k)$ and therefore $I(F, p(k), \mathbf{x}) = I(H, p(k), \mathbf{y})$, where H is the distribution for (Y_t, Y_{t+k}).

Lemma 2.1.

$$I(H, p(k), (z_t, z_{t+k})) = z_t z_{t+k} - p(k) (z_t^2 + z_{t+k}^2)/2,$$

* Michael Chernick is a member of the technical staff, The Aerospace Corporation, El Segundo, CA 90245. His research was conducted while employed at the Oak Ridge National Laboratory. Darryl Downing and David Pike are staff members of the Union Carbide Corporation Nuclear Division, Oak Ridge, TN 37830. The research was sponsored by the U.S. Department of Energy under contract W-7405-eng-26. The authors wish to thank Victor Lowe for his help in providing the nuclear materials inventory difference data.

December 1982, Volume 77, Number 380
Applications Section

where H is any bivariate distribution function with mean zero, unit variance, and covariance $p(k)$.

This result is based on the result for bivariate correlation due to Mallows and applied by Devlin, Gnanadesikan and Kettenring (1975). We see that we can compute the influence of any pair of observations k units apart on the estimate of $p(k)$ based on the formula given in the lemma. When $p(k)$, σ, and μ are not known, estimates can be used in the formula in order to estimate the influence function.

An $n \times m$ matrix, where n is the number of observations and m is a fixed number of lags (m should be considerably less than n) has its (j, k) entry given by $I(H, p(k), (y_j, y_{j+k})$, where y_j is the jth standardized observation. The observation y_t influences several lagged autocorrelation estimates. It appears in the computation of every element in the tth row and also in the diagonal elements of the preceding rows beginning in column 1 of row $t - 1$ and proceeding up and to the right. An outlier will often have a very large positive or negative influence on each estimate of correlation. Hence, if all the elements in the tth row and the above diagonal are large in absolute value, this will indicate that the tth observation is probably an outlier. This "clothes-pin" pattern in the matrix directs one's attention to the suspect observation. Applications of the influence function matrix to detect outliers are given in the next section.

Define

$$U_{i,k,1} = \left\{\frac{(y_i + y_{i+k})}{\sqrt{1 + p(k)}} + \frac{(y_i - y_{i+k})}{\sqrt{1 - p(k)}}\right\} \Big/ 2$$

and

$$U_{i,k,2} = \left\{\frac{(y_i + y_{i+k})}{\sqrt{1 + p(k)}} - \frac{(y_i - y_{i+k})}{\sqrt{1 - p(k)}}\right\} \Big/ 2.$$

It is easy to see that

$$(1 - p^2(k))U_{i,k,1}U_{i,k,2} = y_i y_{i+k} - p(k)\,(y_i^2 + y_{i+k}^2)/2$$

and so

$$I\,(H, p(k), (y_i, y_{i+k})) = (1 - p^2(k))U_{i,k,1}\ U_{i,k,2}.$$

For a stationary Gaussian process with μ, σ, and $p(k)$ all known, $U_{i,k,1}$ and $U_{i,k,2}$ are observations from independent standard normal distributions. The quantity $I(H, p(k), (y_i, y_{i-k}))$ then has the distribution of a constant times a product of standard normal random variables. This distribution can be used to determine what values for the influence function would be unusually large for a realization from a stationary Gaussian process.

To clearly see the patterns in the influence function matrix, we propose choosing a critical value based on the product standard normal distribution. Influence function estimates exceeding this critical value in absolute value are designated + or − depending on the sign of the estimate. Other observations are left blank. The matrix will then appear with patterns of +s and −s and the clothes-pin effect should be evident to the eye. While we do not expect that the stationary and Gaussian assumptions hold in all applications, we do believe that this approach provides a method for revealing outliers in a wide variety of situations. This was apparently the case for the power plant data. A critical value of 1.0 was used in the power plant and nuclear materials data examples.

3. APPLICATIONS OF THE INFLUENCE FUNCTION MATRIX

3.1 The Monthly Power Plant Data

The influence function matrix, as described in Section 2, was applied to the monthly power plant data. For the power plant data base 36 months of electricity generation and energy consumption data were available for the 25 plants that were selected. Table 1 shows the influence function matrix for this example. Large values in row 23 and the above diagonal indicate that observation number 23 is an outlier. That observation was in fact an incorrect consumption figure.

The respondents at the power plant sent a corrected value for observation 23. Table 2 shows the influence function matrix after observation 23 was replaced with the revised value supplied by the power plant. Note that observation number 23 no longer appears suspicious. The first and seventh observations now appear to be outliers. Previously this effect was masked by observation 23. Since these outliers were not detected by the methods of the previous study or by the data editing routines of the Department of Energy, the authors do not know whether the observations are errors.

3.2 Miller's Example

In discussing Martin's paper (see Martin 1980) Miller generates a series of length 100 from a stationary Gaussian AR(1) process with $p(1) = .5$. The series is contaminated by adding normal independent random variables,

Table 1. The Influence Function Matrix for the Uncorrected Consumption Data

Time/Lag	1	2	3	4	5	6	7	8	9	10
1						+				
2										
⋮										
13										
14									−	
15								−		
16										
17						−				
18										
19										
20										
21										
22										
23				−		−	−	+	−	
24										
⋮										
35										

Table 2. The Influence Function Matrix for the Corrected Consumption Data

Time/ Lag	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	–	–		–	–	+			–	–
2										
3				–						
4										
5										
6	–									
7					+	–	–		+	
8										
9										
⋮										
35										

with variance 100, to five observations. The variance for the uncontaminated series is one and hence the innovations have variance .75. The outliers were introduced at observations 10, 37, 49, 75, and 98. Miller notes that the effect of the outliers is to make the sample autocorrelation estimates small (i.e., not significantly different from zero). The uncontaminated estimates for $p(1)$ and $p(2)$ were .527 and .194, respectively, as compared with .066 and $-.065$ for the contaminated series.

Miller's example illustrates the dramatic effect a few outliers can have on estimates of autocorrelation and autoregression coefficients. It is difficult to distinguish the contaminated series from white noise. It seems clear that the influence function matrix would be an effective tool for detecting outliers in Miller's example.

Miller's experiment was repeated using various points in the time series for the introduction of contamination. This was done for 100 different examples. Of the 500 contaminated points only 374 of the disturbances were greater than 3 in absolute value. This is to be expected since a normal distribution with mean zero and variance 100 has 23.6 percent of its probability mass in the interval $[-3, 3]$. Only those 374 observations were considered to be outliers. A procedure based on the product standard normal distribution was used to identify the observations with large influence. This decision rule identified 309 out of the 374 outliers, a 82.6 percentage of identifications. For more details see Chernick and Downing (1980).

3.3 Nuclear Safeguards Problem

In the safeguarding of nuclear materials, inventory differences are computed at periodic intervals to detect loss of materials if they should occur. Various methods have been applied to accomplish this including CUSUM charts and Kalman filtering methods. A major drawback of previous methods is their inability to detect the theft until several periods after the loss has occurred. Since the inventory differences can be modeled as a 1-dependent sequence and a theft at time t corresponds to an outlying observation at t, the influence function matrix provides an alternative method for detecting loss. Since the influence function matrix could have large values for the correlations between the observations prior to time t with the observation at time t, thefts could possibly be detected just after their occurrence.

Data on inventory differences were taken from the Energy Research and Development Administration's 1977 "Report on Strategic Special Nuclear Material Inventory Differences." Inventory differences for plutonium, enriched uranium, U-233, and Pu-238 from 1949 to 1976 are available for various sites. In this analysis, data from Los Alamos Scientific Laboratory (LASL), Oak Ridge National Laboratory (ORNL), Richland Hanford, and Savannah River are used. The data are given in Table 3. In 1977 the Department of Energy began to conduct inventory checks on a semi-annual basis. The data for 1977 in Table 3 are based on the inventory for the first half of 1977, reported in the Department of Energy's "Semi-Annual Report on Strategic Special Nuclear Material Inventory Differences," January 1978.

The Nuclear Regulatory Commission (NRC) carefully checks on the nuclear materials stored in its facilities. Whenever large negative differences in inventory data are found, reasons for the discrepancies are sought and reported. The two reports referenced here provide a description of the inventory difference data and the types of error that can lead to discrepancies.

In some cases the material is not available in a measurable form. At some of the sites data are recorded as "none" for no inventory difference when the material can be checked without measurements. Also, when measured differences are less than 50 grams, they are reported as "<0.1." We shall treat "none" as zero and "<.1" as $-.05$ in this analysis. The data are reported in kilograms and represent the inventory in year t minus the inventory in year $t - 1$. Negative values correspond to possible losses or thefts and are of particular concern. Large positive values are unusual and occur less frequently than the large negative values. These positive values ("gains") are of less concern to the NRC since they do not represent a potential theft.

The columns in Table 3 are numbered from 1 to 7 to designate the seven data sets that were studied. The influence function matrix was computed for five lags using standard estimates of the mean, standard deviation, and correlation coefficients. Note that data sets 1–5 have 29 observations, data set 6 has 26 observations, and data set 7 has 19 observations.

The analysis revealed significant entries in the influence function matrix, based on the cut-off value of 1.0, for the data sets 1–7, respectively.

For data set 1, we note that observations 6, 10, and 24 stand out as outliers. These correspond to the observed values -8.1, -9.3, and 6.1, respectively. For data set 2, observations 26 and 28 are outliers. These correspond to the values -79.6 and 22.7. For data set 3, observations 12, 13, 16, and 17 are outliers. It is interesting here that although observations 12 and 13 are the two largest negative values and 17 is the third largest, 16 is not the fourth largest. For data set 4, we again see consecutive outliers

Table 3. Inventory Differences

	Facility						
	LASL		Richland	ORNL			Savannah
Year	Plutonium 1	Enriched Uranium 2	Plutonium 3	Plutonium 4	Enriched Uranium 5	U-233 6	Pu-238 7
49	.1	−1.1	1.0	−.05	−.05		
50	.1	−2.1	.3	−.05	−.1		
51	−.2	−7.8	−2.9	0	−.05		
52	0	−3.7	1.0	−.1	−.05	−.05	
53	−5.7	−1.7	.3	−.1	−.3	−.05	
54	−8.1	−1.8	−24.6	−.4	−.6	−.05	
55	−.4	−1.1	−23.7	.3	−.1	.1	
56	.9	−1.5	−11.6	−.05	.5	−.1	
57	−3.6	−1.0	−55.7	−.1	−.1	.1	
58	−9.3	−1.4	−86.4	−.1	−.4	−.8	
59	−2.1	.4	−90.1	0	−.3	−1.4	−.05
60	−4.9	−2.6	−143.7	−.05	.1	−.05	−.05
61	−3.1	−1.9	−169.2	−.05	−.1	−.1	−.05
62	−6.8	−1.2	−106.8	0	−.7	−.4	−.1
63	−1.5	−.4	−66.9	−.05	−.9	−.1	−.5
64	−3.5	−2.1	−94.9	−.05	−.05	−.9	−1.9
65	−3.4	.9	−118.8	−.05	.9	−.6	.8
66	−3.4	1.0	17.2	−.05	−.3	.3	1.3
67	−1.0	−.7	−1.5	−.3	−.2	−.4	−1.0
68	3.2	−6.1	−32.0	−.6	.3	.1	−2.5
69	.2	−14.3	62.9	−.4	−.7	−.5	−1.7
70	1.9	1.3	−9.0	−.5	−.6	−.5	−19.6
71	−.1	5.7	.9	−.4	−.3	−.05	−2.9
72	6.1	−15.7	−40.6	−.2	−2.0	−.05	1.6
73	−1.3	5.2	−49.1	−.2	−.1	−.1	2.1
74	−4.8	−79.6	21.2	−.2	−.3	−.05	−2.1
75	−5.2	4.8	−3.4	−.1	−.1	−.05	−1.3
76	−.05	22.7	6.7	−.05	.2	−.05	−2.0
77	.2	−2.6	−3.0	−.05	−.05	−3.3	−2.5

at 6 and 7. Observation 20 is clearly an outlier and observations 22 and 23 may also be considered outliers.

For data set 5, the outliers at observations 17 and 24 are apparent. Since observation 24 is larger in absolute value by a factor of two, it is not surprising to see it identified. However, observation 17 is not any larger in absolute value than observation 15 and yet it stands out. Apparently this is due to the fact that most of the observations are negative. In fact, for this case, the sample mean is −.222. Table 4 shows outliers at observation year 59 and 77 for data set 6. Year 77 is worth noting because it illustrates the outlier being detected at the time of occurrence. Since 77 is the last observation in data set

Table 4. The Influence Function Matrix for Selected Inventory Difference Data Sets

	Data Set 6				
Time/Lag	1	2	3	4	5
52					
53					
54					
55				−	
56					
57		−			
58	+				
59					+
60					
⋮					
70					
71					
72					−
73				−	
74			−		
75		−			
76	−				
77					

Table 5. The Influence Function Matrix for Selected Inventory Difference Data Sets

	Data Set 7				
Time/Lag	1	2	3	4	5
59					
⋮					
64					
65					−
66				−	
67					
68		+			
69					
70		−	−	+	
71					
⋮					
77					

6, it is visible as an outlier only through the upper diagonal entries in the matrix. Table 5 for data set 7 shows only year 70 to be an outlier. Here year 70 is very extreme since it is nearly 7 times larger in absolute value than the next most extreme observation.

The influence function matrix for correlation provides another interesting way to analyze the inventory difference data for outliers. One could also estimate the influence function using robust estimates of location and scale. That might aid some in the detection of outliers, especially for the cases when multiple outliers are present. Also, if inventory differences were only due to measurement error one would expect the mean to be zero. Clearly the data indicate that this is not the case. However, if the influence function matrix were computed with zero replacing the sample mean it might increase the number of negative differences detected while reducing the number of positive differences detected. This may be desirable for inventory difference data since detection of losses is critical.

[Received June 1980. Revised May 1982.]

REFERENCES

BOX, G.E.P., and JENKINS, G.M. (1970), *Time Series Analysis Forecasting and Control*, San Francisco, Holden-Day.

CHERNICK, M.R., and DOWNING, D.J. (1980), "The Influence Function Method Applied to Energy Time Series Data," *Proceedings of the 1980 DOE Statistical Symposium*, 102–111.

DEVLIN, S.J., GNANADESIKAN, R., and KETTENRING, J.R. (1975), "Robust Estimation and Outlier Detection With Correlation Coefficients," *Biometrika*, 62, 531–545.

GNANADESIKAN, R. (1977), *Methods for Statistical Data Analysis of Multivariate Observations*, New York: John Wiley.

HAMPEL, F.R. (1974), "The Influence Curve and its Role in Robust Estimation," *Journal of the American Statistical Association*, 69, 383–393.

MARTIN, R.D. (1980), "Robust Estimation of Autoregressive Models," *Reports on Directions in Time Series*, *Institute of Mathematical Statistics*, 228–262.

Performing the X11-ARIMA Seasonal Adjustment

SUGI 11 Tutorials

John C. Brocklebank
SAS Institute Inc.

David A. Dickey
North Carolina State University

Most seasonal adjustment methods are based on the assumption that seasonal fluctuations can be measured and separated from the underlying trend and irregular components.

The U.S. Bureau of Census Method II-X11 variant of a seasonal adjustment program developed by Shiskin, Young, and Musgrave (1967) has often been criticized. Critics state that the estimates for observations of the most recent years do not have the same degree of reliability as those of the central observations.

The Statistics Canada X11-ARIMA developed by Dagum (1975) extrapolates the unadjusted data one year ahead at both ends using ARIMA models of the Box and Jenkins type. The extended original series is then seasonally adjusted using various moving averages of the Method II-X11 variant.

This paper discusses how X11-ARIMA methodology can be applied using several procedures provided in SAS/ETS® software. In particular, many of the "advantages" of the X11-ARIMA over the Method II-X11 variant are discussed.

Session Objectives

- Understand the principal features of the SAS® seasonal adjustment procedure X11, its outputs, and its uses.
- Obtain projected deseasonalizing weights for the next year.
- Obtain an output data set containing needed variables for plotting and further analysis.
- Perform an analysis using the X11-ARIMA method.

Session Overview

- Seasonal adjustment procedures can signal economic turning points.
- The Bureau of Census program (SAS X11 procedure) can be run in the additive or multiplicative mode.
- The X11 procedure is robust with respect to outlying observations.
- The X11 procedure, together with PROC AUTOREG, can be used to produce year-ahead forecast values.
- Realistic confidence limits on the X11 forecast are somewhat difficult to produce.
- X11 is an inexpensive and powerful tool that you can use even if you are a nonstatistician, but you are limited to a "canned" model applied to all series.
- The X11 procedure can be combined with PROC ARIMA to perform the X11-ARIMA method.

SEASONALITY

Seasonality refers to the regular periodic fluctuations that recur each year with the same timing and intensity.

The majority of procedures for seasonal analysis involve smoothing to eliminate unwanted irregular variation from patterns that are meaningful to the analyst.

Seasonal Adjustments

Many economic series show seasonal variation. Income from an orange grove farm may rise each year from the late fall until early spring and then drop very sharply.

The main use of a seasonal adjustment is to remove such fluctuations and to expose the underlying trend-cycle.

Usually, a wide variety of factors influence economic data, so it can be useful to remove the seasonal component and then observe which influences dominate changes in a time series.

Seasonal Adjustments

Most seasonal adjustment methods are based on the assumption that seasonal fluctuations can be measured and separated from the underlying trend and irregular fluctuations.

The main assumption underlying the X11 program is that a time series is composed of seasonal, trend-cycle, trading-day, and irregular components.

The seasonal adjustment is categorized as either multiplicative or additive.

FORMS OF SEASONAL ADJUSTMENTS

In the multiplicative version, the time series is a product of the following factors.

1. The seasonal component (S) is defined as the intrayear pattern of variation that is repeated constantly from year to year.

2. The trend-cycle (C) includes the long-term trend and business cycle. This can be increasing, decreasing, or unchanged.

3. The trading-day component (TD) consists of variations attributed to the composition of the calendar.

4. The irregular component (I) consists of residual variations such as unseasonable weather conditions, report and sampling errors, and so on.

5. The seasonally adjusted series (CI) consists of a trend-cycle and irregular components.

Thus, the multiplicative version is assumed to have the following structure:

$$Y_t = S_t * C_t * TD_t * I_t$$

An alternate additive formation occurs when the original time series is a summation of these components:

$$Y_t = S_t + C_t + TD_t + I_t \quad .$$

The multiplicative model usually produces the best seasonal factors for most series, but it will not work for series that have negative values or for series that are highly volatile. We suggest the additive model for these.

SAS STATEMENTS

The following program seasonally adjusts the number of personal checks cashed in Canadian clearing centers for the period January 1968 to December 1976:

```
OPTIONS LS=72 NODATE;
DATA CHECKS;
   TITLE
   'CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.'
   INPUT CHECKS @@;
   RETAIN DATE '01DEC 67'D;
   DATE=INTNX('MONTH',DATE,1);
   FORMAT DATE MONYY.;
   CARDS;
456 447 447 530 619 555 607 576 630 666 744 782
697 734 734 789 901 884 854 777 838 881 911 928
881 792 871 911 961 972 1016 923 1043 1005 1092 1088
962 968 1115 1044 1137 1195 1164 1108 1180 1136 1301 1290
1279 1287 1430 1437 1654 1626 1562 1674 1572 1763 2187 1625
1735 1576 1753 1767 2223 2189 2072 2116 1980 2238 2259 2090
2204 2107 2285 2714 3029 2691 2746 2626 2654 2720 2927 2771
2695 2597 2592 2931 2931 3120 3108 2894 3162 3160 3017 3288
3077 3100 3360 3262 3559 3777 3574 3499 3782 3720 3891 3671
;
PROC X11 YRAHEADOUT;
   VAR CHECKS;
   ID DATE;
   MONTHLY DATE=DATE TDREGR=TEST CHARTS=NONE PRINTOUT=STANDARD;
   OUTPUT OUT=X11 B1=B1 D10=D10 D11=D11 D12=D12 D13=D13 C16=C16;
```

SAS STATEMENTS

```
DATA X11;SET X11;
   TIME=_N_;
   TIME2=TIME*TIME;
   D10=D10/100;
   D13=D13/100;
   C16=C16/100;
   LABEL
   B1='B1: ORIGINAL SERIES'
   D10='D10: SEASONAL FACTORS %'
   D11='D11: FINAL SEASONALLY ADJ.'
   D12='D12: TREND CYCLE'
   D13='D13: IRREGULAR SERIES %'
   C16='C16: TRADING DAY FACTORS %';
PROC AUTOREG;
   MODEL D11=TIME TIME2/NLAG=13 BACKSTEP METHOD=ML;
   OUTPUT OUT=AUTOREG P=P;
DATA AUTOREG;
   SET AUTOREG;
   IF D11=. THEN
      DO;
         FOREX11=P*D10*C16;
         OUTPUT;
      END;
DATA FORECAST;
   SET X11(KEEP=B1 DATE );
   _TYPE_='ACTUAL   ';
   IF B1=. THEN
      DO;
         _TYPE_='FORECAST';
         SET AUTOREG(KEEP=FOREX11 RENAME=(FOREX11=B1));
      END;
```

AUTOREG Output

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

A U T O R E G   P R O C E D U R E

DEPENDENT VARIABLE = D11   D11: FINAL SEASONALLY ADJ.

ORDINARY LEAST SQUARES ESTIMATES

SSE	1102777	DFE	105
MSE	10502.64	ROOT MSE	102.4824
SBC	1317.508	AIC	1309.462
REG RSQ	0.9895	TOTAL RSQ	0.9895
DURBIN-WATSON	0.6618		

VARIABLE	DF	B VALUE	STD ERROR	T RATIO	APPROX PROB
INTERCPT	1	518.897562	30.1405546	17.216	0.0001
TIME	1	8.731484	1.2764608	6.840	0.0001
TIME2	1	0.203591	0.0113454	17.945	0.0001

AUTOREG Output

ESTIMATES OF AUTOCORRELATIONS

LAG	COVARIANCE	CORRELATION
0	10210.9	1.000000
1	6655.68	0.651821
2	5118.56	0.501284
3	4629.56	0.453394
4	3785.8	0.370761
5	3301.77	0.323357
6	3512.07	0.343953
7	3848.07	0.376859
8	2716.32	0.266021
9	1855.02	0.181670
10	1262.47	0.123639
11	455.998	0.044658
12	-413.83	-0.040528
13	-473.035	-0.046327

BACKWARD ELIMINATION OF AUTOREGRESSIVE TERMS

LAG	ESTIMATE	T-RATIO	PROB
5	-0.010737	-0.0916	0.9272
11	0.014544	0.1257	0.9002
13	0.021906	0.2132	0.8316
9	0.028365	0.2464	0.8059
4	0.030810	0.2973	0.7669
10	0.041154	0.4327	0.6662
2	-0.054984	-0.5015	0.6172
6	-0.076182	-0.7551	0.4520
8	0.089787	0.9013	0.3696
3	-0.150695	-1.7577	0.0818
12	0.138100	1.8049	0.0740

AUTOREG Output

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

A U T O R E G   P R O C E D U R E

PRELIMINARY MSE= 5602.677

ESTIMATES OF THE AUTOREGRESSIVE PARAMETERS

LAG	COEFFICIENT	STD ERROR	T RATIO
1	-0.59226712	0.07772981	-7.619562
7	-0.17314685	0.07772981	-2.227548

EXPECTED AUTOCORRELATIONS

LAG	AUTOCORR
0	1.0000
1	0.6314
2	0.4081
3	0.2786
4	0.2132
5	0.1970
6	0.2260
7	0.3070

AUTOREG Output

MAXIMUM LIKELIHOOD ESTIMATES

SSE	578560.2	DFE	103
MSE	5617.089	ROOT MSE	74.94724
SBC	1258.162	AIC	1244.751
REG RSQ	0.9229	TOTAL RSQ	0.9945
DURBIN-WATSON	2.0573		

VARIABLE	DF	B VALUE	STD ERROR	T RATIO	APPROX PROB
INTERCPT	1	478.671830	75.8409074	6.312	0.0001
TIME	1	11.034113	3.1667214	3.484	0.0007
TIME2	1	0.180677	0.0280623	6.438	0.0001
A(1)	1	-0.603778	0.0769188	-7.850	0.0001
A(7)	1	-0.192075	0.0784486	-2.448	0.0160

EXPECTED AUTOCORRELATIONS

LAG	AUTOCORR
0	1.0000
1	0.6553
2	0.4409
3	0.3143
4	0.2501
5	0.2357
6	0.2682
7	0.3540

AUTOREG Output

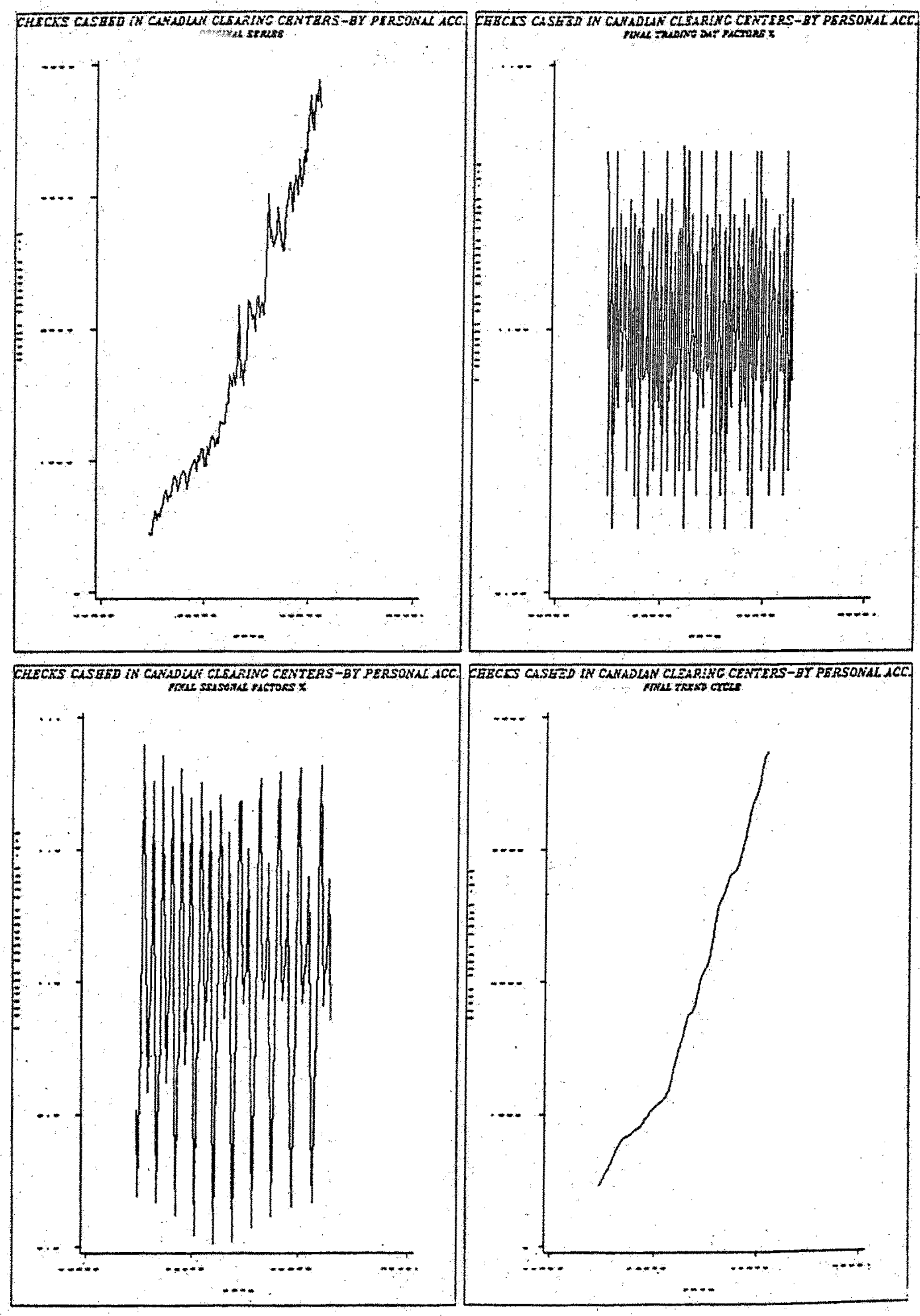
CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

A U T O R E G   P R O C E D U R E

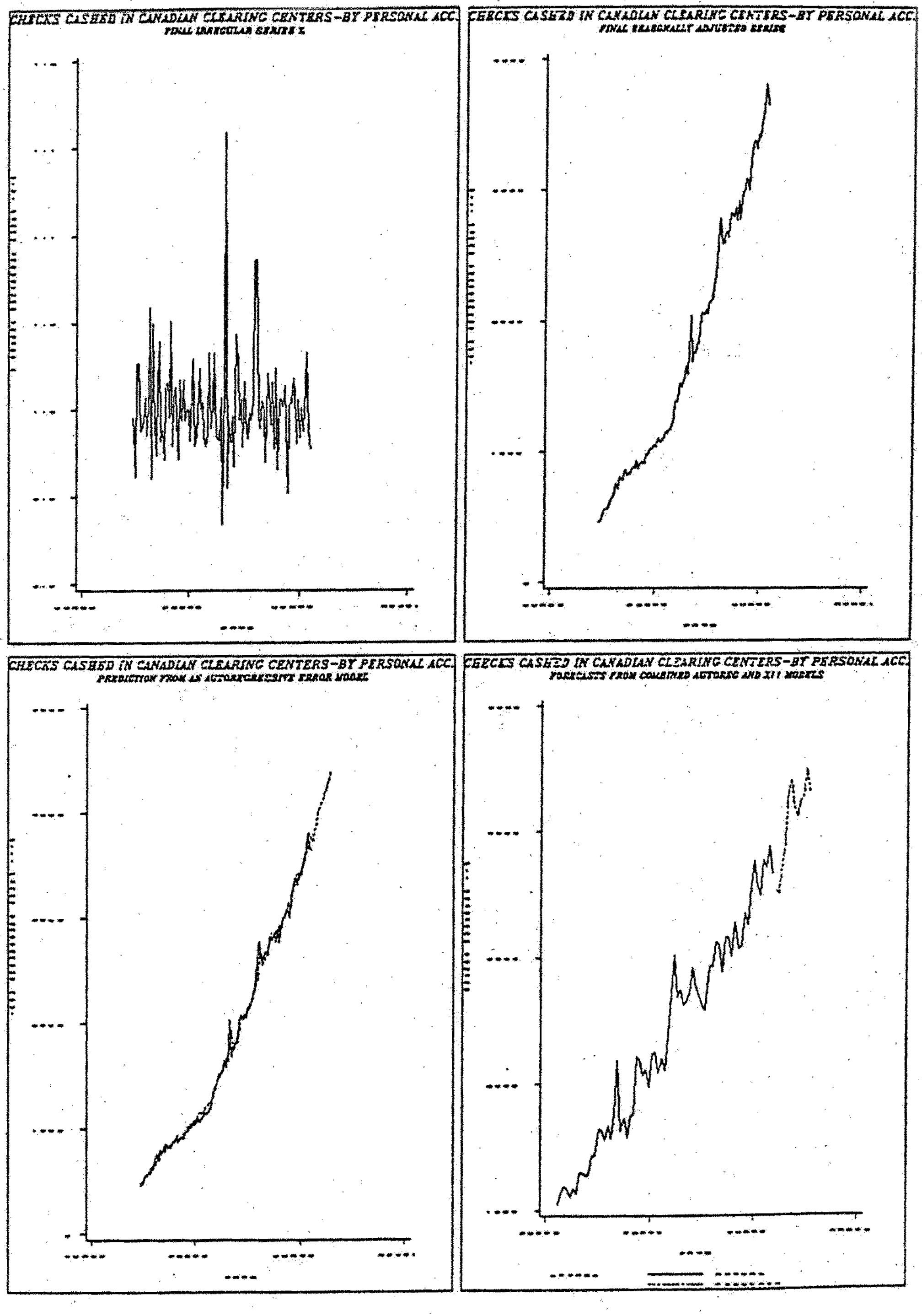
AUTOREGRESSIVE PARAMETERS ASSUMED GIVEN.

VARIABLE	DF	B VALUE	STD ERROR	T RATIO	APPROX PROB
INTERCPT	1	478.671830	75.6502274	6.327	0.0001
TIME	1	11.034113	3.1267666	3.529	0.0006
TIME2	1	0.180677	0.0274974	6.571	0.0001

Output for Selected Graphs



Output for Selected Graphs



able C15: Final Trading-Day Regression

This table examines the results of the final trading-day regression computation.

Actual regression coefficients are given in column 3. Only coefficients for Monday through Saturday are derived, and Sunday's is the negative of the sum of the other six.

The remaining statistics help to judge the significance and reliability of the combined weights.

The F-ratio is the ratio of the variance in the irregular series, explained by trading-day regression, to the variance not explained. It measures the significance of the trading-day weights.

F-values greater than 2.5 indicate there is a less than 1% probability that the trading-day weights would be this extreme due only to chance.

Standard errors and t values provide some measures of the precision of the estimated combined daily weights.

Table C15: Final Trading-Day Regression

Two hypotheses are tested:

- combined weights differ from 1 (*)
- the combined weights differ from the prior daily weight supplied (* *).

C15 FINAL TRADING DAY REGRESSION

	COMBINED WEIGHT	PRIOR WEIGHT	REG. COEFF.	ST.ERR. COMB.WT.	T (1)	T PR.WT.
MONDAY	1.301	1.000	0.301	0.172	1.746	1.746
TUESDAY	1.231	1.000	0.231	0.176	1.311	1.311
WEDNESDAY	1.515	1.000	0.515	0.176	2.926*	2.926**
THURSDAY	0.932	1.000	-0.068	0.180	-0.378	-0.378
FRIDAY	1.155	1.000	0.155	0.186	0.831	0.831
SATURDAY	0.609	1.000	-0.391	0.179	-2.183	-2.183
SUNDAY	0.258	1.000	-0.742	0.172	-4.315*	-4.315**

* COMB. WT. DIFFERS FROM 1 AT 1 PER CENT LEVEL

** COMB. WT. DIFFERS FROM PRIOR WT. AT 1 PER CENT LEVEL

SOURCE OF VARIANCE	SUM OF SQUARES	DGRS.OF FREEDOM	MEAN SQUARE	F
REGRESSION	47.403	6	7.901	18.673
ERROR	41.041	97	0.423	
TOTAL	88.444	103		

PROBABILITY OF A LARGER F IS 1.0E-04

STD. ERRORS OF TRADING DAY ADJUSTMENT FACTORS

31-DAY-MONTH	0.486
30-DAY-MONTH	0.538
29-DAY-MONTH	0.594
28-DAY-MONTH	0.000

Selected X11 Tables

Table D8: Final Unmodified SI Ratios

- The final iteration of the program starts with Table D1. This table is the original series, adjusted by any prior weights and adjusted for any trading-days, with the extreme values modified.

- The original series or original series adjusted for trading-days is divided term-by-term by a Henderson moving average, yielding a series of SI ratios in which extremes have not been modified. This series is shown in Table D8.

- Most economic time series have a seasonal pattern that changes very gradually from year to year. A stable seasonal pattern is the name given to a series without change.

- A test for the existence of seasonal variation in the original series is given at the bottom of the table. The F-test used compares the variance in the SI series, which is due to differences between months, with the variance not explained by these differences.

- This is the ANOVA F test using months as "treatments." An F ratio of 2.34 or greater signifies a less than 1% probability that the differences between monthly means are due to chance.

- The F-test can give a false indication of seasonal variation when the two variances being compared are very small. Series with rapidly changing seasonal patterns can yield a small F-ratio.

- Series showing little or no evidence of seasonal variation should not be seasonally adjusted. There is no provision in the procedure to halt the computation when the F ratio is below the level of significance since the test is fallible.

Table D8: Final Unmodified SI Ratios

D 8 FINAL UNMODIFIED SI RATIOS

YEAR	JAN	FEB	MAR	APR	MAY	JUN	
1968	97.092	91.760	92.585	99.590	111.210	101.572	
1969	91.819	96.753	95.841	95.047	109.647	106.870	
1970	100.847	89.839	97.519	98.589	107.418	101.368	
1971	93.978	91.173	99.814	95.582	106.280	104.667	
1972	97.951	88.644	96.663	100.611	105.847	105.136	
1973	94.221	88.869	97.084	95.224	111.292	109.593	
1974	94.675	89.989	96.331	105.949	115.990	107.808	
1975	93.581	92.586	94.125	100.591	103.225	108.940	
1976	95.275	93.353	96.732	95.915	106.883	106.629	
AVG	95.493	91.441	96.299	98.566	108.644	105.843	
YEAR	JUL	AUG	SEP	OCT	NOV	DEC	AVG
1968	99.779	95.278	100.133	97.886	108.277	109.328	100.374
1969	99.312	95.393	97.337	100.645	108.634	102.243	99.962
1970	103.384	95.966	100.783	99.223	107.527	102.131	100.383
1971	103.878	97.377	100.624	99.606	104.758	99.071	99.734
1972	101.390	99.962	93.270	101.738	121.749	94.260	100.602

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

D 8 (CONTINUED.)

YEAR	JUL	AUG	SEP	OCT	NOV	DEC	AVG
1973	101.708	100.095	98.380	102.127	104.390	97.221	100.017
1974	101.535	100.402	100.530	97.690	106.453	99.500	101.404
1975	102.611	98.925	101.300	99.998	99.817	99.417	99.593
1976	102.479	97.482	100.991	103.033	102.633	97.372	99.898
AVG	101.786	97.876	99.261	100.216	107.138	100.061	

TOTAL - 10823.603

STABLE SEASONALITY TEST

	SUM OF SQUARES	DGRS.OF FREEDOM	MEAN SQUARE	F
BETWEEN MONTHS	2491.597	11	226.509	21.370
ERROR	1017.539	96	10.599	
TOTAL	3509.136	107		

PROBABILITY OF A LARGER F IS 0.0001

X11-ARIMA

Introduction

The X11-ARIMA seasonal adjustment method developed by Dagum (1983) consists of:

1. Modeling the original series by autoregressive integrated moving averages (ARIMA models) of Box and Jenkins (1970).
2. Extrapolating one year of the unadjusted data at each end of the series from ARIMA models used in (1). This operation, called forecasting and backcasting, is designed to extend the observations at both ends.
3. Seasonally adjusting the extended series with the various linear filters provided in X11.

Because of this forecasting and backcasting method, the central observations now receive symmetric filters.

The use of the ARIMA option by default automatically tests three built-in ARIMA models against built-in criteria of fitting and extrapolation.

X11-ARIMA

The three built-in models using the Box and Jenkins symboli notation $(p,d,q)(P,D,Q)_s$ and equivalent SAS PROC ARIMA statements are given below.

$(2,1,2)(0,1,1)_s$

```
PROC ARIMA;
   IDENTIFY VAR=X(1,s) NOPRINT;
   ESTIMATE P=(1,2) Q=(1,2)(s) NOCONSTANT;
```

$\log(0,1,1)(0,1,1)_s$

```
PROC ARIMA;
   IDENTIFY VAR=LX(1,s) NOPRINT;
   ESTIMATE     Q=(1)(s) NOCONSTANT;
```

$\log(0,2,2)(0,1,1)_s$

```
PROC ARIMA;
   IDENTIFY VAR=LX(1,1,s) NOPRINT;
   ESTIMATE Q=(1,2)(s) NOCONSTANT;
```

X11-ARIMA

The adequacy of any one of the three automatically fitted ARIMA models by the X11-ARIMA method is verified by:

1. Testing the randomness of the residuals with a portmanteau test of fit developed by Box-Pierce.

2. A method that checks for overdifferencing based on checking the estimated values of the parameters (for a better approach see Dickey and Brocklebank (1984)).

Suggested Advantages of X11-ARIMA over X11

1. The availability of a statistical model provides relevant information on the quality of the raw data.

2. The existence of a model that fits the original series fulfills the fundamental principle of a seasonal adjustment, that is the series is decomposable.

3. One-step extrapolation from ARIMA models is a minimum-mean-square-error and can be used as a projected value or benchmark for preliminary analysis.

4. A reduction of some 30% is typically observed in the bias of the total error in the seasonal factor forecasts.

5. Trend-cycle estimates for the last observations are made with symmetric weights of the Henderson averages combined with the weights of the ARIMA model used for the extrapolated data. The weights reflect most recent movements and a better trend-cycle estimate is obtained from the combined weights. This is particularly true for years with turning points because X11 applies asymmetric weights of the Henderson filters.

6. A better estimate of the variance of the irregulars is obtained.

The next few pages perform the X11-ARIMA seasonal adjustment using PROC ARIMA and PROC X11 and compare the results with the Statistics Canada X11-ARIMA seasonal adjustment program.

```
PROC ARIMA DATA=CHECKS;
  IDENTIFY VAR=CHECKS(1,12) NOPRINT;
  ESTIMATE P=(1,2) Q=(1,2)(12) MAXIT=100 NOCONSTANT METHOD=CLS;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

ARIMA: CONDITIONAL LEAST SQUARES ESTIMATION

PARAMETER	ESTIMATE	APPROX. STD ERROR	T RATIO	LAG
MA1,1	-0.409219	0.234237	-1.75	1
MA1,2	-0.21357	0.194814	-1.10	2
MA2,1	0.725609	0.0843692	8.60	12
AR1,1	-0.955193	0.200665	-4.76	1
AR1,2	-0.627679	0.135678	-4.63	2

```
VARIANCE  ESTIMATE =   16929.8
STD ERROR ESTIMATE =   130.115
AIC                =   1199.46*
SBC                =   1212.23*
NUMBER OF RESIDUALS=        95
* DOES NOT INCLUDE LOG DETERMINANT
```

CORRELATIONS OF THE ESTIMATES

	MA1,1	MA1,2	MA2,1	AR1,1	AR1,2
MA1,1	1.000	-0.113	-0.020	0.888	0.464
MA1,2	-0.113	1.000	-0.029	-0.290	0.584
MA2,1	-0.020	-0.029	1.000	0.065	0.021
AR1,1	0.888	-0.290	0.065	1.000	0.439
AR1,2	0.464	0.584	0.021	0.439	1.000

AUTOCORRELATION CHECK OF RESIDUALS

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	4.75	1	0.029	-0.020	-0.058	-0.149	-0.125	-0.064	0.037
12	11.09	7	0.135	0.138	-0.003	0.008	-0.126	0.147	0.048
18	16.05	13	0.247	-0.066	0.056	-0.108	-0.033	0.109	0.102
24	17.52	19	0.555	-0.090	-0.005	-0.046	-0.006	0.042	-0.000

```
MODEL FOR VARIABLE CHECKS
NO MEAN TERM IN THIS MODEL.
PERIODS OF DIFFERENCING= 1,12.

 AUTOREGRESSIVE FACTORS
                 FACTOR 1
1+0.955193B**(1)+0.627679B**(2)

 MOVING AVERAGE FACTORS
                 FACTOR 1
1+0.409219B**(1)+ 0.21357B**(2)
                 FACTOR 2
1-.725609B**(12)
```

Performing the X11-ARIMA Adjustment

The parameter estimates obtained by the Statistics Canada X11-ARIMA seasonal adjustment program.

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) EXTRAPOLATION PROGRAM

A 5. ARIMA EXTRAPOLATION MODEL (FORECAST)

THIS PROGRAM WAS DEVELOPED FOLLOWING THE PROCEDURES OUTLINED IN 'TIME SERIES ANALYSIS' BY G. E. P. BOX AND G. M. JENKINS.

MODEL	TRAN.	ADDITIVE CONSTANT	AVERAGE PERCENTAGE STANDARD ERROR IN FORECASTS: LAST 3 YEARS	LAST YEAR	LAST-1 YEAR	LAST-2 YEAR	CHI-SQ. PROB.	R-SQUARED VALUE	ESTIMATED PARAMETERS			
(2,1,2)(0,1,1)	NONE	0.0	6.59	4.08	4.66	11.02	45.08%	0.9807	-0.960 0.714	-0.642	-0.419	-0.235

THE MODEL CHOSEN IS (2,1,2)(0,1,1)0 WITH TRANSFORMATION - NONE

Obtain a slightly better fitting model.

```
ESTIMATE P=(1,2) Q=(4,12) MAXIT=100 NOCONSTANT METHOD=CLS;
FORECAST OUT=FORE1 LEAD=12 ID=DATE INTERVAL=MONTH NOOUTALL;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

ARIMA: CONDITIONAL LEAST SQUARES ESTIMATION

PARAMETER	ESTIMATE	APPROX. STD ERROR	T RATIO	LAG
MA1,1	0.169134	0.0870385	1.94	4
MA1,2	0.656672	0.0914314	7.18	12
AR1,1	-0.638625	0.102297	-6.24	1
AR1,2	-0.32729	0.102072	-3.21	2

```
VARIANCE  ESTIMATE =   16524.1
STD ERROR ESTIMATE =   128.546
AIC                =   1196.21*
SBC                =   1206.42*
NUMBER OF RESIDUALS=        95
* DOES NOT INCLUDE LOG DETERMINANT
```

CORRELATIONS OF THE ESTIMATES

	MA1,1	MA1,2	AR1,1	AR1,2
MA1,1	1.000	-0.208	0.085	0.181
MA1,2	-0.208	1.000	0.166	0.006
AR1,1	0.085	0.166	1.000	0.481
AR1,2	0.181	0.006	0.481	1.000

AUTOCORRELATION CHECK OF RESIDUALS

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	2.03	2	0.362	0.016	0.007	-0.026	-0.106	-0.037	0.079
12	7.53	8	0.481	0.079	0.046	0.082	-0.115	0.150	0.007
18	13.61	14	0.479	-0.116	0.070	-0.069	0.041	0.138	0.093
24	16.01	20	0.716	-0.097	0.037	-0.005	0.018	0.090	-0.015

```
MODEL FOR VARIABLE CHECKS
NO MEAN TERM IN THIS MODEL.
PERIODS OF DIFFERENCING= 1,12.

 AUTOREGRESSIVE FACTORS
                 FACTOR 1
1+0.638625B**(1)+ 0.32729B**(2)

 MOVING AVERAGE FACTORS
                 FACTOR 1
1-.169134B**(4)-.656672B**(12)
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

MODEL FOR VARIABLE CHECKS
NO MEAN TERM IN THIS MODEL.
PERIODS OF DIFFERENCING= 1,12.

AUTOREGRESSIVE FACTORS
FACTOR 1
1+0.638625B**(1)+ 0.32729B**(2)

MOVING AVERAGE FACTORS
FACTOR 1
1-.169134B**(4)-.656672B**(12)

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

FORECASTS FOR VARIABLE CHECKS

OBS	FORECAST	STD ERROR	LOWER 95%	UPPER 95%	ACTUAL	RESIDUAL
--------FORECAST BEGINS--------						
109	3636.5186	128.5460	3384.5737	3888.4636		
110	3610.4081	136.6820	3342.5168	3878.2994		
111	3734.9652	148.0173	3444.8572	4025.0731		
112	3875.2407	166.8768	3548.1688	4202.3126		
113	4085.8749	171.3728	3749.9910	4421.7589		
114	[illegible]35.8178	180.5021	3782.0408	4489.5947		
11[illegible]	[illegible]2.0501	189.5444	3680.5506	4423.5496		
[illegible]	[illegible].3899	196.3673	3583.5179	4353.2620		
[illegible]	[illegible]8226	203.9520	3708.0850	4507.5603		
[illegible]	[illegible]441	211.1911	3686.5180	4514.3701		
[illegible]	[illegible]2.3159	217.9262	3785.1892	4639.4426		
[illegible]	4134.9850	224.6449	3694.6900	4575.2799		

Performing the X11-ARIMA Adjustment

```
DATA FORE1;
  SET FORE1(KEEP=FORECAST DATE RENAME=(FORECAST=CHECKS));
PROC PRINT;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

OBS	DATE	CHECKS
1	JAN77	3636.52
2	FEB77	3610.41
3	MAR77	3734.97
4	APR77	3875.24
5	MAY77	4085.87
6	JUN77	4135.82
7	JUL77	4052.05
8	AUG77	3968.39
9	SEP77	4107.82
10	OCT77	4100.44
11	NOV77	4212.32
12	DEC77	4134.98

Performing the X11-ARIMA Adjustment

Use an ARIMA model to "backcast" the data.

```
PROC SORT DATA=CHECKS OUT=SCHECKS;
   BY DESCENDING DATE;
DATA OUT2;
   SET SCHECKS;
   T=_N_;
PROC ARIMA;
   IDENTIFY VAR=CHECKS(1,12) NOPRINT;
   ESTIMATE P=(1,2) Q=(1,2)(12) MAXIT=100 NOCONSTANT METHOD=CLS;
   FORECAST OUT=FORE2(KEEP=FORECAST) LEAD=12 ID=T NOOUTALL;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

ARIMA: CONDITIONAL LEAST SQUARES ESTIMATION

PARAMETER	ESTIMATE	APPROX. STD ERROR	T RATIO	LAG
MA1,1	-0.363664	0.219136	-1.66	1
MA1,2	-0.193093	0.188596	-1.02	2
MA2,1	0.602937	0.0918914	6.56	12
AR1,1	-0.912178	0.187128	-4.87	1
AR1,2	-0.628106	0.131217	-4.79	2

```
VARIANCE  ESTIMATE =   19042.4
STD ERROR ESTIMATE =   137.994
AIC                =   1210.63*
SBC                =   1223.4*
NUMBER OF RESIDUALS=      95
* DOES NOT INCLUDE LOG DETERMINANT
```

CORRELATIONS OF THE ESTIMATES

	MA1,1	MA1,2	MA2,1	AR1,1	AR1,2
MA1,1	1.000	-0.119	0.034	0.877	0.399
MA1,2	-0.119	1.000	-0.033	-0.285	0.616
MA2,1	0.034	-0.033	1.000	0.109	-0.051
AR1,1	0.877	-0.285	0.109	1.000	0.379
AR1,2	0.399	0.616	-0.051	0.379	1.000

AUTOCORRELATION CHECK OF RESIDUALS

TO LAG	CHI SQUARE	DF	PROB	AUTOCORRELATIONS					
6	4.56	1	0.033	-0.022	-0.054	-0.154	-0.106	-0.084	-0.004
12	10.23	7	0.176	0.095	0.033	0.072	-0.158	0.096	0.054
18	13.29	13	0.426	-0.087	0.011	-0.056	-0.020	0.086	0.087
24	17.90	19	0.529	-0.162	0.055	0.033	-0.067	0.053	-0.001

```
MODEL FOR VARIABLE CHECKS
NO MEAN TERM IN THIS MODEL.
PERIODS OF DIFFERENCING= 1,12.

 AUTOREGRESSIVE FACTORS
                 FACTOR 1
1+0.912178B**(1)+0.628106B**(2)

 MOVING AVERAGE FACTORS
                 FACTOR 1
1+0.363664B**(1)+0.193093B**(2)
                 FACTOR 2
1-.602937B**(12)
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

FORECASTS FOR VARIABLE CHECKS

OBS	FORECAST	STD ERROR	LOWER 95%	UPPER 95%	ACTUAL	RESIDUAL
--------FORECAST BEGINS--------						
109	493.2349	137.9942	222.7719	763.6980		
110	545.1888	151.4067	248.4377	841.9398		
111	439.3630	167.3624	111.3395	767.3866		
112	397.7030	200.6256	4.4849	790.9211		
113	354.3898	212.1946	-61.5031	770.2828		
114	393.3899	227.5932	-52.6838	839.4635		
115	384.4602	247.1479	-99.9397	868.8601		
116	417.2389	258.3487	-89.1142	923.5919		
117	297.3982	272.1818	-236.0673	830.8636		
118	241.0476	286.4931	-320.4675	802.5627		
119	198.7385	297.3059	-383.9691	781.4461		
120	210.0020	309.5463	-396.6964	816.7003		

Performing the X11-ARIMA Adjustment

The parameter estimates obtained by the Statistics Canada X11-ARIMA seasonal adjustment program.

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) EXTRAPOLATION PROGRAM

A 6. ARIMA EXTRAPOLATION MODEL (BACKCAST)

THIS PROGRAM WAS DEVELOPED FOLLOWING THE PROCEDURES OUTLINED IN 'TIME SERIES ANALYSIS' BY G. E. P. BOX AND G. M. JENKINS.

MODEL	TRAN.	ADDITIVE CONSTANT	AVERAGE PERCENTAGE STANDARD ERROR IN BACKCASTS: LAST 3 YEARS	LAST YEAR	LAST-1 YEAR	LAST-2 YEAR	CHI-SQ. PROB.	R-SQUARED VALUE	ESTIMATED PARAMETERS			
(2,1,2)(0,1,1)	NONE	0.0	12.32	12.10	10.46	14.40	3.78%	0.9710	-0.989 0.516	-0.463	-0.354	0.148

THE MODEL CHOSEN IS (2,1,2)(0,1,1)0 WITH TRANSFORMATION - NONE

Fix the date values.

```
DATA FORE2;
   RETAIN DATE '01JAN68'D;
   SET FORE2(RENAME=(FORECAST=CHECKS));
   DATE=INTNX('MONTH',DATE,-1);
   FORMAT DATE MONYY.;
PROC SORT;
   BY DATE;
PROC PRINT;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

OBS	DATE	CHECKS
1	JAN67	210.002
2	FEB67	198.738
3	MAR67	241.048
4	APR67	297.398
5	MAY67	417.239
6	JUN67	384.460
7	JUL67	393.390
8	AUG67	354.390
9	SEP67	397.703
10	OCT67	439.363
11	NOV67	545.189
12	DEC67	493.235

Performing the X11-ARIMA Adjustment

Concatenate the data sets prior to performing X11.

```
DATA CHECKS2;
   SET FORE2 CHECKS FORE1;
PROC X11 DATA=CHECKS2;
   VAR CHECKS;
   MONTHLY DATE=DATE . . .
```

Performing the X11-ARIMA Adjustment

Compare the forecasts of a combined estimation using PROC X11 and PROC AUTOREG with the forecasts fit with PROC ARIMA.

```
PROC AUTOREG DATA=X11;
   MODEL D11=TIME TIME2/NLAG=13 BACKSTEP NOPRINT;
   OUTPUT OUT=AUTOREG P=P;
DATA AUTOREG;
   SET AUTOREG;
   IF D11=. THEN
      DO;
         FOREX11=P*D10*C16;
         OUTPUT;
      END;
DATA FORECAST;
   SET X11(KEEP=B1 DATE );
   _TYPE_='ACTUAL   ';
   IF B1=. THEN
      DO;
         _TYPE_='FORECAST';
         SET AUTOREG(KEEP=FOREX11 RENAME=(FOREX11=B1));
      END;
DATA COMPARE;
   INPUT TRUE @@;
   SET AUTOREG(KEEP= DATE FOREX11);
   RESX11=(TRUE-FOREX11);
   SRESX11+ABS(RESX11);
   SSEX11+(RESX11)**2;
   SET FORE1(RENAME=(CHECKS=FOREARMA) KEEP=CHECKS);
   RESARMA=(TRUE-FOREARMA);
   SEARMA+ABS(RESARMA);
   SSEARMA+(RESARMA)**2;
   CARDS;
3315 3414 3797 3449 4253 4342 3909 3991 4144 3809 4444 4128
PROC PRINT;
   ID DATE;
```

CHECKS CASHED IN CANADIAN CLEARING CENTERS-BY PERSONAL ACC.

DATE	TRUE	FOREX11	RESX11	SRESX11	SSEX11
JAN77	3315	3527.36	-212.36	212.36	45096
FEB77	3414	3506.43	-92.43	304.79	53639
MAR77	3797	3731.98	65.02	369.80	57867
APR77	3449	3891.78	-442.78	812.59	253924
MAY77	4253	4316.73	-63.73	876.32	257985
JUN77	4342	4423.00	-81.00	957.32	264547
JUL77	3909	4206.80	-297.80	1255.12	353229
AUG77	3991	4122.45	-131.45	1386.57	370509
SEP77	4144	4243.49	-99.49	1486.06	380407
OCT77	3809	4292.90	-483.90	1969.96	614569
NOV77	4444	4503.19	-59.19	2029.15	618073
DEC77	4128	4336.76	-208.76	2237.91	661653

DATE	FOREARMA	RESARMA	SEARMA	SSEARMA
JAN77	3636.52	-321.52	321.52	103377
FEB77	3610.41	-196.41	517.93	141954
MAR77	3734.97	62.03	579.97	145802
APR77	3875.24	-426.24	1006.21	327484
MAY77	4085.88	167.12	1173.33	355414
JUN77	4135.82	206.18	1379.51	397926
JUL77	4052.05	-143.05	1522.57	418389
AUG77	3968.39	22.61	1545.17	418900
SEP77	4107.82	36.18	1581.35	420209
OCT77	4100.44	-291.44	1872.80	505149
NOV77	4212.32	231.68	2104.48	558826
DEC77	4134.99	-6.99	2111.47	558875

REFERENCES

Bureau of the Census. (1967), The X-11 Variant of the Census Method II Seasonal Adjustment Program, U.S. Department of Commerce, Technical Paper No. 15 (1967 revision), Washington, DC: Government Printing Office.

Bureau of the Census. (1969), X-11 Information for the User, U.S. Department of Commerce, Washington, DC: Government Printing Office.

Dagum, E.B. (1983), The X-11-ARIMA seasonal adjustment method, Statistics Canada, Catalogue 12-563E, Ottawa.

Dickey, D.A. and Brocklebank, J.C. (1984), Preliminary Transformations in Time Series Modeling, Proceedings of the Ninth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, 58-63.