MODELING TIME SERIES WHEN CALENDAR EFFECTS ARE PRESENT

W. P. Cleveland
M. R. Grupe
Board of Governors of the Federal Reserve System

1. INTRODUCTION

Recent papers by W. S. Cleveland and Devlin (1980), Liu (1980) and W. P. Cleveland and Pollner (1978) have pointed to the existence of calendar effects in economic data. The background for much of this work is contained in Young (1965), which describes the relevant properties of the calendar and gives the equations for the calendar adjustment in X−11 computer programs for seasonal adjustment.

Work on seasonal adjustment based on time series models has stimulated new interest in modeling these effects. It turns out that a straightforward use of the methods contained in Box and Jenkins (1970) is not sufficient to capture calendar variations because they are not precisely periodic or not so within a useful time frame. As an example, consider an Easter effect in monthly data y_t. When Easter falls in the same month in consecutive years it is correct to compare y_t with y_{t-12}. In other cases there is no precise comparison, as Easters are 11 or 13 months apart while the remaining seasonal has the usual 12-month separation. The autocorrelation function will capture the average relationships in the series, but no standard ARIMA model can simulate acutal behavior for each year. In this paper, regression models are used for calendar effects. Using appropriate design matrices, the precise form of a deterministic effect can be estimated and subtracted from the series. An ARIMA model is applied to the residuals. This is analogous to subtracting out the intervention part of the model as in Box and Tiao (1975), or more generally, the deterministic part of the model as in Pierce (1978). If seasonal adjustment is desired, it can be accomplished by using the complete model as in Pierce (1978), by applying the ARIMA-based filters in Box, Hillmer, and Tiao (1978) to the residuals, or by applying X−11 to the residuals. Improved forecasts are obtained by extending the design matrix into the future. The current limitation of this approach is the implied assumption that the effects are the same from year to year relative to the calendar.

Three types of effects will be discussed for monthly data. These are length-of-month, composition-of-month (number of Sundays, etc.,), and Easter effects. Week-of-month effects, seasonal effects, and certain holiday effects in weekly data will also be analyzed. The effects of the composition of the month are often called trading-day effects, referring to business transactions typical for a given day of the week.

2. SIGNALS OF CALENDAR EFFECTS IN MONTHLY DATA

Calendar effects are generally second-order effects which cannot be observed until other sources of variation have been eliminated or reduced. Figures 1A and 2A show the currency and demand deposit components of M-1A as compiled by the Federal Reserve Board. Figure 3A shows total apparel retail sales as compiled by the U.S. Bureau of the Census. The three series illustrate a range of seasonal variation relative to the trend. The three series are also shown in figures 1B-3B after a straight line detrending to achieve a scale which reveals their seasonal patterns better. Any calendar effects are masked by the strong seasonal. Since all appear to have a seasonal which increases with trend, the series were logged for further analyses.

The spectra of the first difference of the logged currency and retail sales series appear in figures 4 and 5. While not the same, both are dominated by seasonal effects. The same is true for demand deposits. The autocorrelation functions have high values at multiples of 12. A periodic mean was fit to each series as an approximation to a seasonal. This was removed to reveal remaining effects. The resulting spectra and autocorrelations are shown in figures 6 to 8. All three series exhibit spectral peaks at $f = .348$, the primary calendar frequency cited in Cleveland and Devlin (1980). The frequency .432 shows up in the currency series and the additional frequency .304 in the retail sales series. All three of these frequencies relate to composition of the month. The frequency .304 corresponds to the first harmonic of the weekly frequency, while .432 corresponds to a week by seasonal fundamental interaction. The latter may be due to different within-week patterns at different times of year, or simply reflect that the composition of the months repeats approximately some years for certain day-of-week patterns. The retail sales series exhibits four other peaks. These will be shown to relate to an Easter effect.

The autocorrelation functions illustrate the difficulty of identifying calendar effects by this means. An autocorrelation function based on a simulated monthly composition effect is given in figure 9. Some evidence of this pattern is evident at lags beyond 10 in the currency autocorrelations. However, the pattern becomes averaged in with the autocorrelations of the stochastic part of the series, as mentioned in Liu (1980). It is too weak to show up in the demand deposit autocorrelations and is dominated by the Easter pattern in retail sales autocorrelations.

3. LENGTH-OF-MONTH AND COMPOSITION-OF-MONTH EFFECTS FOR FLOW DATA

The two monetary series presented represent inventories or stocks of money. They reflect the cumulative sums of transactions or flows. The sales series represents transactions directly. It is useful to characterize a day of the month in terms of the net transactions on that day. On a holiday this would be zero. Other days might have positive or negative values. Since the model will be based on transactions or flows, the appropriate model for sales is considered first.

The complete model for the observation y_t is

$$\Phi(B)(z_t - f_t(\alpha)) = \theta(B)a_t \tag{3.1}$$

where z_t is either y_t or $\log y_t$ as appropriate, B is the backshift operator such that $B^k z_t = z_{t-k}$, $\Phi(B)$ is an autoregressive polynomial in B with roots on or outside the unit circle, and $\theta(B)$ is a moving average polynomial with roots outside the unit circle. The form of $f_t(\alpha)$ is to be specified, and a_t is a sequence of independent uncorrelated noise terms with zero mean and constant variance σ_a^2. A discussion of models not containing $f_t(\alpha)$ may be found in Box and Jenkins (1970) or Nelson (1973).

The vector f $(f_t: t = 1, 2, \ldots, T)$ can be expressed as $f = X\alpha$ where α is a vector of estimated coefficients and X is the design matrix with T rows to be generated from the calendar. The matrix X given here is a modification of the one in Young (1965).

Let $m_t(k)$ be the number of days of type k in month t, where $k = 1$ represents Sunday, $k = 2$ represents Monday, etc. Then $m_t = \sum_{k=1}^{7} m_t(k)$ will denote the length in days of month t, and $m_t/7$ is the average number of days of each type in month t. Finally $\overline{m} = 365.25/12$, the average month length over a 4-year cycle, and the combined effect of composition and length of month will be denoted by $(TD)_t$. The basic equation for these effects is

$$(TD)_t = \sum_{k=1}^{7} m_t(k)\beta_k \tag{3.2}$$

where the β_k are regression coefficients. The rows of X could be taken directly from this equation as $m_t(k)$, $k = 1, 2, \ldots, 7$, but it is useful to alter them slightly. Equation (3.2) may be rewritten as

$$(TD)_t = \sum_{k=1}^{7} m_t(k)\beta_k - \frac{m_t}{7}\sum_{k=1}^{7}\beta_k + \frac{m_t - \overline{m}}{7}\sum_{k=1}^{7}\beta_k + \frac{\overline{m}}{7}\sum_{k}\beta_k \tag{3.3a}$$

$$= \sum_{k=1}^{7} m_t(k)(\beta_k - \overline{\beta}) + (m_t - \overline{m})\overline{\beta} + \overline{m}\overline{\beta} \tag{3.3b}$$

$$= \sum_{k=1}^{7} \left(m_t(k) - \frac{m_t}{7}\right)(\beta_k - \overline{\beta}) + (m_t - \overline{m})\overline{\beta} + \overline{m}\overline{\beta} \tag{3.3c}$$

The first term of (3.3c) corresponds to the composition-of-month effect normalized so the coefficients add to zero. The first seven columns of X are $m_t(k) - m_t/7$, $k = 1, 2, \ldots, 7$, and the first seven α_k are $\beta_k - \overline{\beta}$. The seven columns have rank 6. An eighth column $(m_t - \overline{m})$ is used to represent length-of-month effects with $\alpha_8 = \overline{\beta}$. Observe that for 28-day Februarys $m_t(k) = m_t/7$ for each k, so there is no composition effect. Column 8 would be spanned by a set of seasonal dummies except for leap years. We orthogonalize seasonal dummies with respect to this column to permit a specific length-of-month estimate using α_8. The seasonal dummies are now nearly rank deficient, but the effect estimated takes a much simpler form. It now represents the intensity of the net transactions in each season and not the number of days on which they occurred. This is more in keeping with the general sense of seasonal effects. If the data are differenced to achieve stationarity, the columns of X are also differenced so the transformed effects are correctly modeled. An intercept column is added to X which absorbs $\overline{m}\overline{\beta}$ in the absence of differencing.

4. STOCK DATA

Consider now a series which represents stock at the end of the month. Let $F_t(i)$ be the transactions or flow on day i of month t, S_t be the stock at the end of month t.

$$S_t = S_{t-1} + \sum_i F_t(i) \tag{4.1}$$

Equation (4.1) states that the new stock or inventory is the old one plus the inflows and outflows which took place during the month. Since (3.3) is intended to describe flows, it is related to the first difference of stock data. The same matrix X can be used, but it is not differenced when the data is for estimation of α. To transfrom the effect back to the undifferenced data, the columns of X must be integrated to form W. The first row of W is set equal to zero. The rest are formed recursively from

$$w_t' = w_{t-1}' + x_{t-1}' \tag{4.2}$$

where w_t' is row t of W. The entire matrix W can then be centered so $W\alpha$ has mean zero. Thus for end-of-month stocks we have

$$S_t - S_{t-1} = x_t'\alpha \tag{4.3}$$

$$S = W\alpha \tag{4.4}$$

5. DAILY AVERAGED STOCKS

The demand deposit and currency series compiled by the Federal Reserve Board do not represent end-of-month stocks, but the average of daily stocks. This causes an uneven weighting of the flows. Let $S_t(i)$ be the stock on day i of month t, and suppose a 3-day month.

$$S_t(1) = S_{t-1} + F_t(1)$$

$$S_t(2) = S_{t-1} + F_t(1) + F_t(2) \tag{5.1}$$

$$S_t(3) = S_{t-1} + F_t(1) + F_t(2) + F_t(3)$$

$$\sum_i S_t(i) = 3S_{t-1} + 3F_t(1) + 2f_t(2) + F_t(3) \tag{5.2}$$

$$\bar{S}_t = S_{t-1} + F_t(1) + \frac{2}{3}F_t(2) + \frac{1}{3}F_t(3) \tag{5.3}$$

Since day 1 flow appears in all subsequent averages, it gets the most weight. For a calendar month, one must know on which days of the month each day of the week fell and compute its appropriate weight. If day k occurs on the first day of the month, then the expression corresponding to (5.2) for the month has for the coefficients of the flows for that day

$$N_t(k) = m_t + m_t - 7 + m_t - 14 + m_t - 21 + m_t - 28 \tag{5.4}$$

where $N_t(k)$ is the cumulative times day k of the week occurs in the daily stock sum for the month. Let $J_t(k)$ be the sum of the days of the month less one on which day k occurs, i.e. $0 + 7 + 14 + 21 + 28$ in (5.4). Then a general expression for $N_t(k)$ is

$$N_t(k) = m_t \times m_t(k) - J_t(k) \tag{5.5}$$

Corresponding to (5.3) is

$$\bar{S}_t = S_{t-1} + \sum_{k=1}^{7} [m_t(k) - \frac{1}{m_t}J_t(k)]F(k) + \bar{p}_t \tag{5.6}$$

Here $F(k)$ represents the average flow attributed to day k, and p_t the flow due to other sources. This expression assumes any patterns in flows other than day-of-the-week patterns apply equally to each day of the week. The logic of (5.6) gives for the next month

$$\bar{S}_{t+1} = S_{t-1} + \sum_{k=1}^{7} [m_t(k) + m_{t+1}(k) - \frac{1}{m_{t+1}}J_{t+1}(k)]F(k) + \bar{p}_t + \bar{p}_{t+1} \tag{5.7}$$

Thus the first difference of the data is related to

$$\bar{S}_{t+1} - \bar{S}_t = \sum_{k=1}^{7} [m_{t+1}(k) + \frac{1}{m_t}J_t(k) - \frac{1}{m_{t+1}}J_{t+1}(k)]F(k) + \bar{p}_{t+1} - \bar{p}_t \tag{5.8}$$

Now $F(k)$ plays the role of α in (4.3), and $m_{t+1}(k)$ represents the same rows of X as before. The only modification to (4.3) is to change X by the $J_t(k)$ terms of (5.8). Then W and X have the same relation as in (4.2).

Length-of-month effects for stock at the end of the month are completely analogous to the effects for flows, except they are related to $(1 - B)y_t$. The situation for daily averages is somewhat more complicated. Corresponding to (5.4) is

$$L_t = m_t + m_{t-1} + m_{t-2} + \ldots + 1 = \frac{m_t(m_t + 1)}{2} \tag{5.9}$$

Letting $\bar{S}$ be the average daily flow

$$\bar{S}_t \propto S_{t-1} + \frac{m_t + 1}{2} \tag{5.10}$$

and

$$\bar{S}_{t+1} \propto S_{t-1} + [m_t + \frac{m_{t+1} + 1}{2}] \tag{5.11}$$

where α indicates proportionality. These imply

$$\bar{S}_{t+1} - \bar{S}_t \propto \frac{m_t + m_{t+1}}{2} \tag{5.12}$$

One should observe that calendar effects will produce the same spectral peaks and autocorrelations in flow, stock at the end of the month, or daily averaged stock data. These signals are more related to frequencies of repetition in the calendar than to the specific form of the effect. Further, the three design matrices relating to $y_t - y_{t-1}$ are correlated, so that use of the wrong one will still reduce composition-of-month effects considerably.

6. RESULTS FOR DAY-OF-THE-WEEK MODELS

The spectrum of the retail sales series after removal of the length-of-month and composition-of-month effects is shown in figure 10. The frequencies corresponding to these effects are gone. An F test based on the reduction in residual sums of squares (RRSS) divided by the residual sum of squares (RSS) for a complete model is $F_{7,121} = 25$. The corresponding tests for currency and demand deposits are $F_{7,101} = 38$ and $F_{7,101} = 20$. Figure 11 shows the resulting spectrum for currency. Use of the flow model on the stock series also produced significant results, but the RRSS figures were smaller and the spectra somewhat less free of calendar effects. Estimation was accomplished by minimizing the conditional sum $\sum_t a_t^2$ over the parameters of Φ, θ, and $f_t(\alpha)$ in equation (3.1).

7. EASTER EFFECT IN SALES DATA

Models for holiday effects are necessarily more specific to a particular series. While the timing of the effect is general, the response to it is not. Composition-of-month effects only consider variations in flow intensity by day of the week. A holiday response might be a decrease or increase before or after the date. The changes might be linear over several days or might take some other form. Thus a parameterization should be partly informed by subject-matter knowledge and partly by the data itself, as in intervention analysis.

A plot of April minus March differences in z_t vs. the date of Easter is given in figure 12 for the retail sales series with composition-of-month effects removed. The date of Easter is coded as the distance from April 1. There is a clear difference between March dates and late April dates. The piecewise linear form drawn in represents the simplest form which adequately represents the pattern. Easter falls in the critical April 1 to April 10 period so few times that a more complex shape cannot be justified. The theoretical perspective is an increase in sales prior to Easter. This sets the left corner in April 1, as any increase must fall entirely in March. The April 10 corner is more arbitrary, but includes the Friday a week before Easter and is consistent with the data.

The model corresponding to this linear change is a uniform increase in daily sales over the period 10 days prior to Easter. The amount of this which falls in March is the sum of the days in this period which fall in March, the integral of a constant.

Figure 13 shows a simulation of this model using the constraint that March plus April sales should not be affected. Note how few times the swing amplitudes are reduced, and the irregular pattern of the Easter dates. The spectrum and autocorrelations of the first difference of this simulation are given in figure 14. Both patterns can be identified in figures 8 and the spectral peaks in figure 10. These were completely removed by the model with an $F_{1,121} = 171$.

8. WEEKLY DATA

Weekly data poses a different set of problems from monthly data with respect to seasonality and to calendar variations. Whereas January always covers the same time span relative to the seasons, the first Sunday-through-Saturday period of the year does not. Thus, the phasing of the weeks within the year must be considered in estimating seasonality, and a $1 - B^{52}$ operator is not as precise a seasonal filter as a $1 - B^{12}$ operator is for monthly data. Also, many more holidays vary from one week of the year to another than vary from one month of the year to another. Some specific holiday effects will be modeled in section 10. The next section concentrates on a more severe problem in monetary series.

9. WEEK-OF-THE-MONTH EFFECTS

Weekly data on demand deposits of member banks are compiled by the Federal Reserve Board. These figures are an average of daily figures for a week spanning Thursday of one week through Wednesday of the next. They are dated by the Wednesday date. A weekly series was grouped into sets of four or five observations by month in which the Wednesdays occurred. The groups were then sorted by the day of the month on which the first Wednesday of the month fell, 1-7. The averages of the corresponding Wednesdays contained in the same day-of-the-month pattern are shown in the first seven lines of figure 15. The bottom is an average of all Wednesdays by day of the month on which they occurred. In general, there are high values in the first part of the month and a fall toward the end of the month. As the first Wednesday moves from the first to the third, it rises relative to later values. The third pattern down shows months on which the first Wednesday fell on the 3rd of the month and the fifth Wednesday (if there was a fifth) fell on the 31st. Note how the last value appears to rise toward a first-of-the-month level.

We do not have a complete understanding of the reasons for this pattern at this time. Social Security payments and other monthly payroll checks which go out near the end of the month are probably major contributors to the pattern.

This within-month variation may be as important for weekly data as the usual seasonal is for monthly data. A plot of the demand deposit series along with its spectrum is shown in figure 16. Note the peaks at $f = .23$ and $f = .46$. These are due to the sampling of a monthly frequency on a weekly basis. A month is $365.25/12 = 30.4375$ days or 4.348 weeks. The reciprocal of 4.348 is .23. Thus $f = .23$ is the fundamental and $f = .46$, the first harmonic of a monthly pattern in weekly data.

Within-month variations such as the one illustrated in figure 15 can be modeled using a Fourier expansion. Let m_t be the number of days in the month containing week t. Let i be the day of the month of the data value y_t. Then values x_{tj} for a design matrix X may be generated according to

$$x_{tj} = \sin(\frac{2\pi ji}{m_t}) \text{ or } \cos(\frac{2\pi ji}{m_t}) \qquad (9.1)$$

Use of three to five frequences (j values) is usually sufficient, implying 6-10 parameters. This model assumes the same within-month patterns for all months. Estimation and removal of an $X\alpha$ having the described form completely eliminated the peaks at $f = .23$ and $f = .46$.

10. HOLIDAY EFFECTS IN WEEKLY DATA

The total retail sales series was published weekly through 1977, as well as monthly. Though the data are not strictly comparable due to sampling differences, the series provides an interesting comparison and confirmation of holiday models in weekly and monthly data. Both types of models are based on assumed daily patterns, but the timing and the levels of aggregations are different.

10.1 Easter

The Easter pattern made an interesting study in the monthly series because of its movement from March to April and its various positionings within those months. The weekly series is aggregated from Sunday through the following Saturday. The figures just before Easter always represent the 7 days before Easter, i.e., the same 7 days relative to Easter. The week of the year will vary from year to year. This is fundamentally simpler than the monthly situation where both the month and phasing within the months change. Figure 17 shows a plot for each year of the week of Easter along with the week after and 3 weeks before. The series was logged and detrended with a straight line before plotting to remove most of the trend variation. The week before Easter appears lightly elevated, while Easter week is much higher than the first 2 and last weeks. This is generally consistent with the pattern of 10 days of increased sales prior to Easter postulated for the monthly series, as three would fall to the first elevated week and seven to the second. The estimated weights for these weeks were .28 and .72, respectively.

10.2 Christmas

The effect of Christmas on the monthly retail sales series was large December values. As neither the month nor the phase within the month changes, it becomes part of the regular seasonal. In relation to the weekly series Christmas behaves much like Easter in the monthly series. Both the week of the year and the day of the week change from one year to the next. Figure 18 shows the data for each year grouped by week relative to Christmas. The set labeled X is for the Saturday just prior to Christmas up to and including Christmas falling on Saturday. When Christmas falls on Friday, all of the sales days just prior to the holiday will be credited to the Saturday after Christmas, or $X + 1$. This situation is represented in the leftmost two lines of each group. The Saturdays before Christmas tend to have their minimum values, as they are at their maximum distance from the holiday. Moving right in each group, the next two lines represent Christmas on Thursday. The Saturday after Christmas shows a decline, as 1 less day prior to Christmas falls in that week. The Saturdays prior to Christmas show an increase as they move closer to Christmas. The next three lines are for Christmas on Wednesday and the next line is for Christmas on Tuesday. By the time Christmas reaches Monday, the relevant sales days for the week after Christmas are only Sunday, when the stores would likely be closed, so that this week shows no further decline. The earlier weeks continue to increase. Now assume a linearly daily increase in sales is running from December 1 through December 24, omitting December 24 if it is a Sunday and rescaling the rest to the same total. This would imply a linear increase in a value for a Saturday before Christmas as the Saturday moved toward Christmas. The values over $X - 1$, $X - 2$, and $X - 3$ should be a maximum for Christmas on Saturday, while the values over X would be a maximum for Christmas on Sunday. The last three lines for Saturday show a decrease in the X pattern because a day is removed from the week.

A more general model for the daily pattern would be

$$W_t = \frac{W_0}{1 - \delta B} S_t$$

where $S_t = 1$ from December 1 through December 24. This is one of the intervention models described in Box and Tiao (1976). For $\delta = 1$ a linear increase is obtained. In general $W_t = W_0(1 + \delta + \delta^2 + \ldots + \delta^{t-1})$ which approaches an asymptotic value $W_0(1 - \delta^{24}/1 - \delta$ for $\delta < 1$ and increases exponentially for $\delta > 1$. For $t > 24$ $W_t = 0$. The estimated value of δ was .98, not significantly different from 1. December 1 was used as a starting point because the $X - 4$ values in figure 18 did not show an increase as they approached Christmas. Figure 19 shows a plot of 3 years of the series with dots indicating values after removing estimated holiday effects for Easter and Christmas. Other calendar effects may remain, but this alone reduces the sum of squares of the first differences to 30 percent of its original value. Application of an ARIMA model containing terms in B^{52} obtained only a 46-percent reduction.

11. FIXED SEASONALS IN WEEKLY DATA

The causes of seasonality which do not change timing from year to year relative to the seasons are the subject of this section. When their effects are the same year to year, some fixed functional form is appropriate. In additive models for monthly data, 12 seasonal dummies constrained to add to zero are usually specified. An equivalent specification is

$$s_t = \sum_{j=1}^{5} [\alpha_j \sin(\frac{2\pi jt}{12}) + \beta_j \cos(\frac{2\pi jt}{12})] + \beta_6 \cos(\pi t)$$

The extension of this equation to weekly data is not straightforward for two reasons. The first Sunday-through-Saturday period of the year is generally not the first 7 days of the year. Hence, the first Saturday (or any other day) data for a year does not represent exactly the same season. In addition, the number of terms required would be 52 or 53. This seems excessive.

We have tentatively adopted a model which mimics the monthly model for weekly data. Let n_y be the number of days in year y. Let u_t be the first day of the year in week t and v_t the last. Then the definition of s_t is

$$s_t = \sum_{i=u_t}^{v_t} \sum_{j=1}^{6} \alpha_j \sin(\frac{2\pi ji}{n_y}) + \beta_j \cos(\frac{2\pi ji}{n_y})$$

This formula includes frequencies up to six times a year. The effect for a specific week is obtained by summing over the appropriate days. If all frequencies up to 26 per year were used, Christmas and other holidays which are tied to a given day of the year could be modeled this way. It seems more parsimonious to use the trigonometric formulation for low frequencies and holiday models for special effects.

The variance reduction accomplished by this formulation depends on the rest of the model. In stationary series or a series detrended by a very low frequency filter such as a straight line, the variance reduction is substantial. If the series is detrended by a first difference operator, the effect is much less because the frequencies involved are mostly removed by the differencing.

Figure 1A. CURRENCY COMPONENT OF M-1A

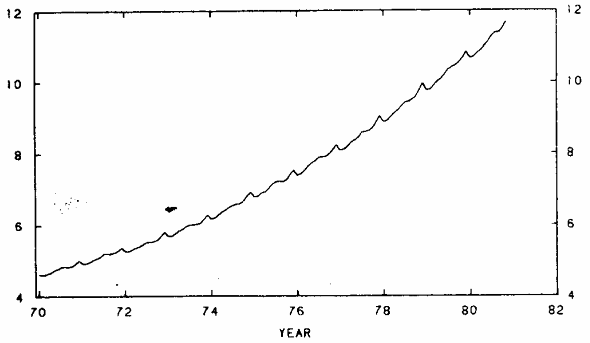


Figure 1B. CURRENCY COMPONENT OF M-1A DETRENDED

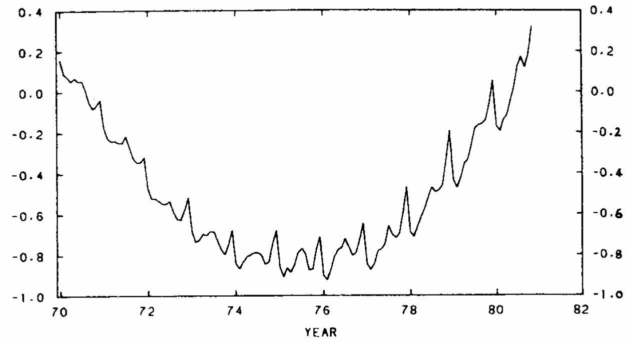


Figure 2A. DEMAND DEPOSIT COMPONENT OF M-1A

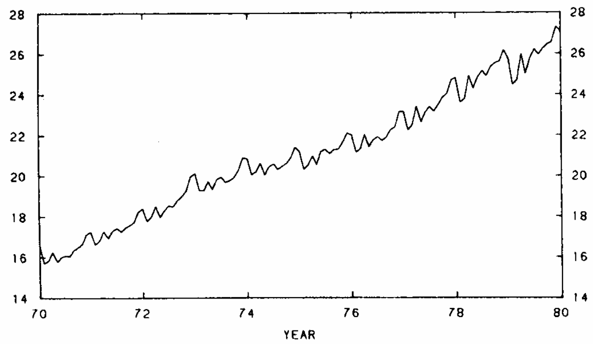


Figure 2B. DEMAND DEPOSIT COMPONENT OF M-1A DETRENDED

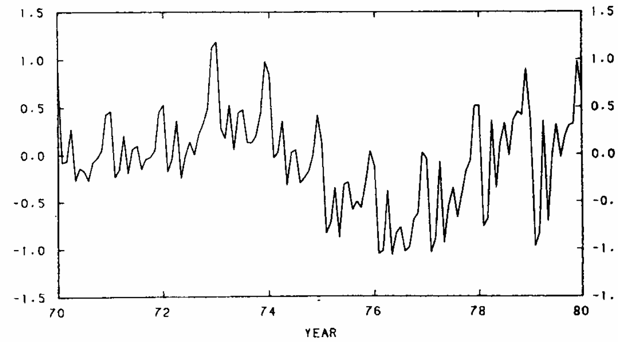


Figure 3A. RETAIL SALES, APPAREL AND ACCESSORY STORES

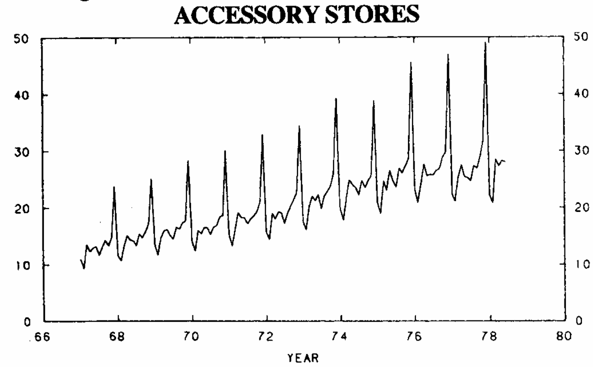


Figure 3B. RETAIL SALES, APPAREL AND ACCESSORY STORES DETRENDED

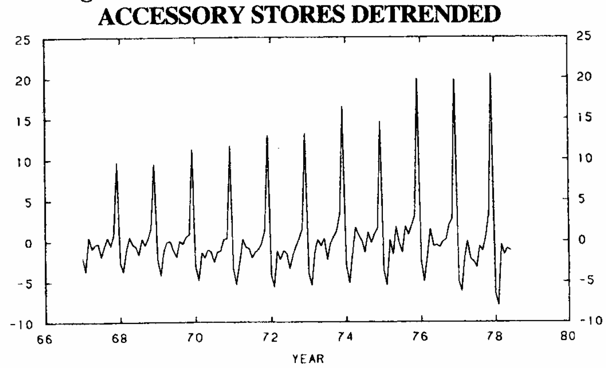


Figure 4. LOGGED CURRENCY, SERIES MEAN

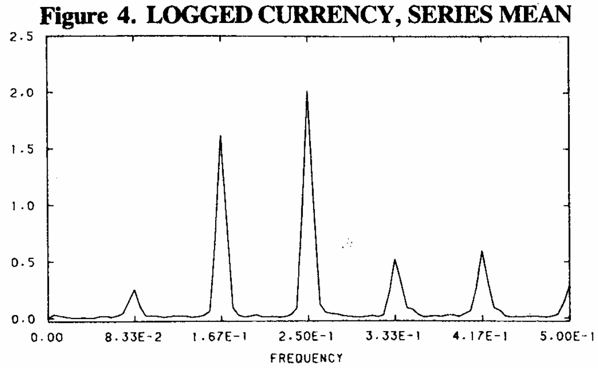


Figure 5. LOGGED RETAIL SALES, SERIES MEAN

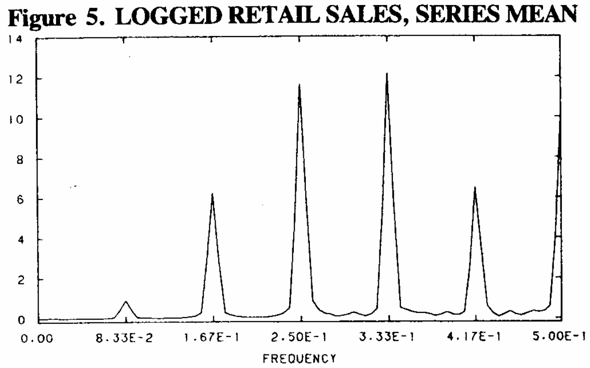


Figure 6A. LOGGED CURRENCY, PERIODIC MEAN

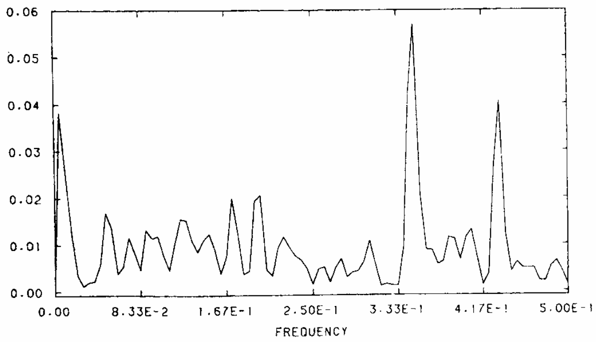


Figure 6B. LOGGED CURRENCY, AUTOCORRELATIONS

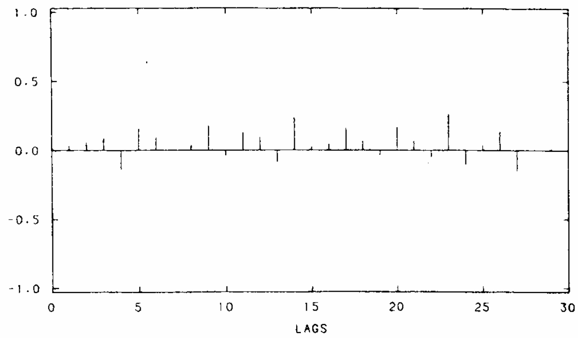


Figure 7A. LOGGED DEMAND DEPOSITS, PERIODIC MEAN

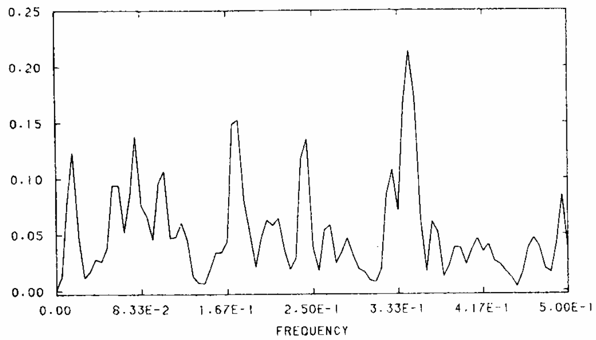


Figure 7B. DEMAND DEPOSITS, AUTOCORRELATIONS

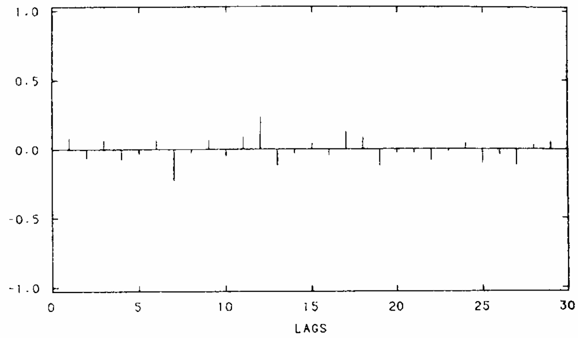


Figure 8A. LOGGED RETAIL SALES, PERIODIC MEAN

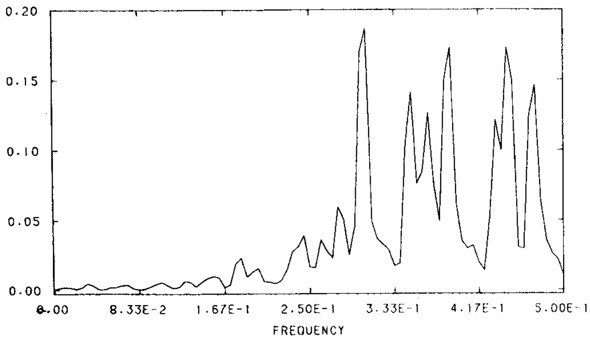


Figure 8B. RETAIL SALES, AUTOCORRELATIONS

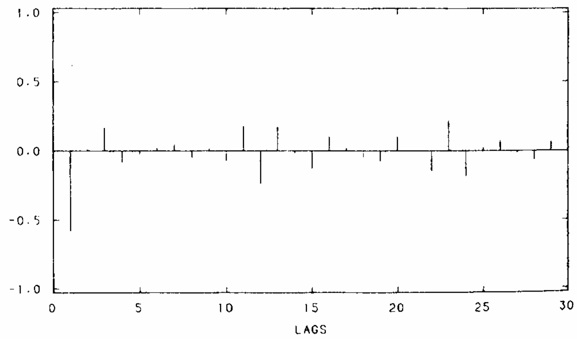


Figure 9A. RETAIL SALES, SIMULATED DAY-OF-WEEK EFFECT, SERIES MEAN

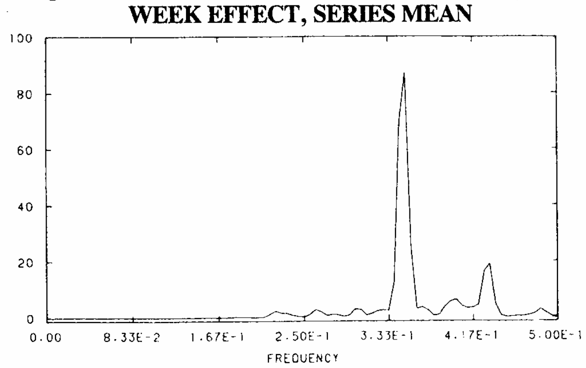


Figure 9B. RETAIL SALES, SIMULATED DAY-OF-WEEK EFFECT, AUTOCORRELATIONS

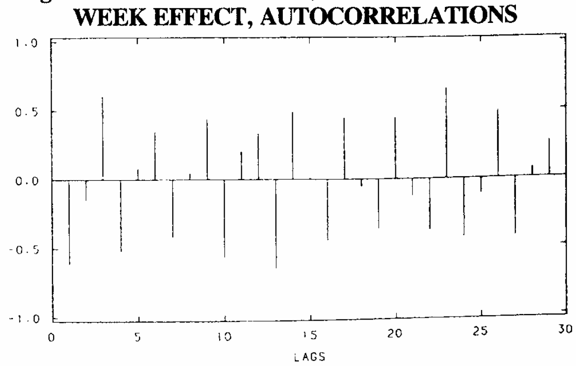


Figure 10. RETAIL SALES WITH DAY-OF-WEEK EFFECTS REMOVED, PERIODIC MEAN

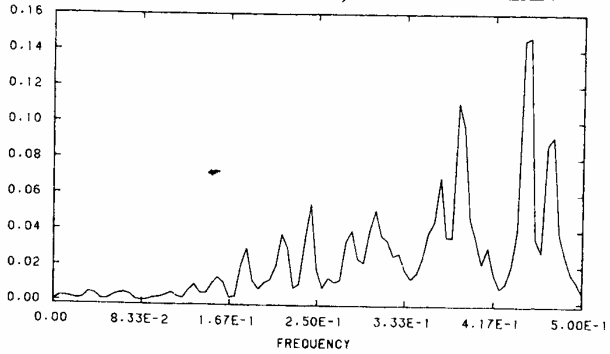


Figure 11. CURRENCY WITH DAY-OF-WEEK EFFECTS REMOVED, PERIODIC MEAN

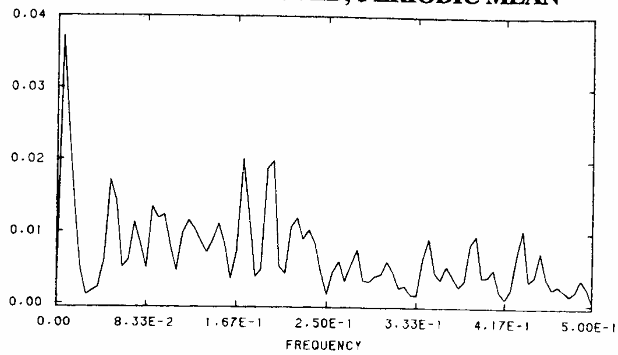


Figure 12. RETAIL SALES, EASTER EFFECT IN APRIL MINUS MARCH VALUES

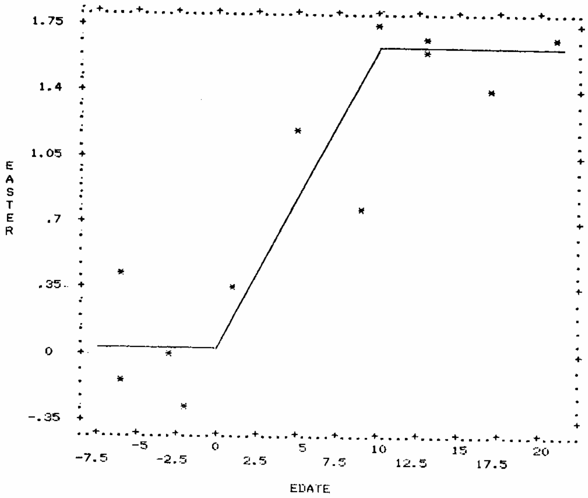


Figure 13. RETAIL SALES, SIMULATED EASTER EFFECT

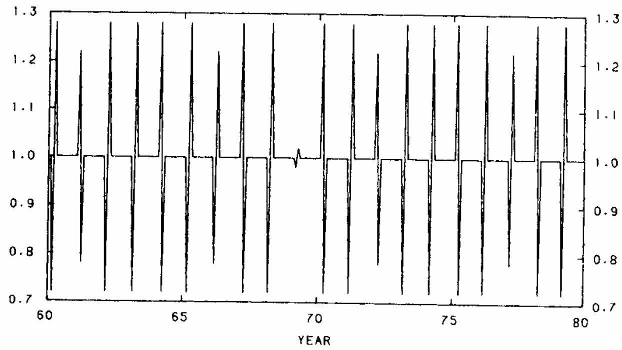


Figure 14A. RETAIL SALES, FIRST DIFFERENCE OF SIMULATED EASTER EFFECT, SERIES MEAN

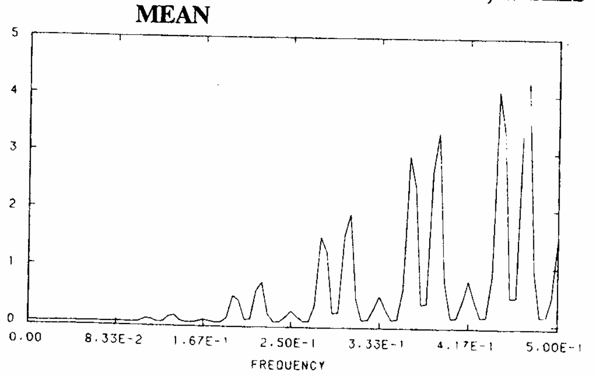


Figure 14B. RETAIL SALES, FIRST DIFFERENCE OF SIMULATED EASTER EFFECT, AUTOCORRELATIONS

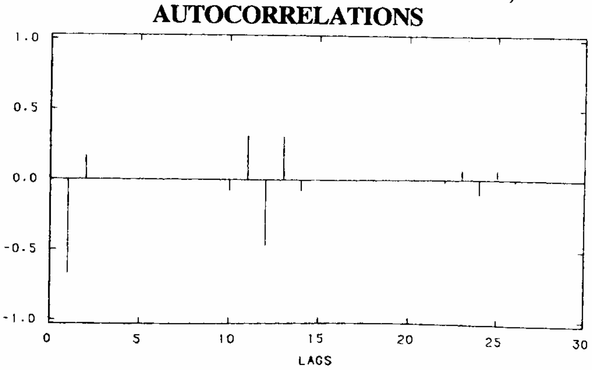


Figure 15. DEMAND DEPOSITS, WITHIN-MONTH PATTERNS

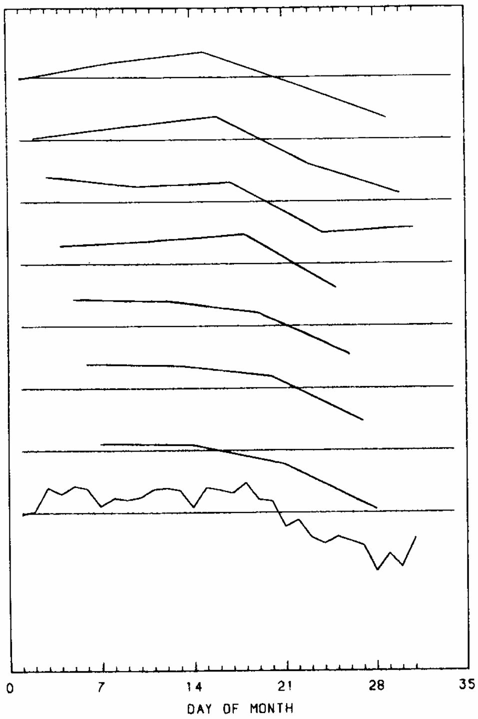


Figure 16A. WEEKLY DEMAND DEPOSITS

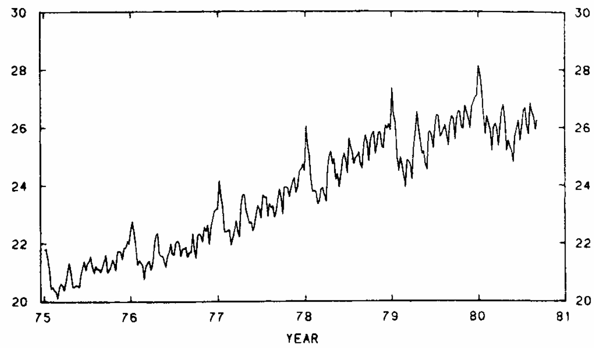


Figure 16B. LOGGED WEEKLY DEMAND DEPOSITS

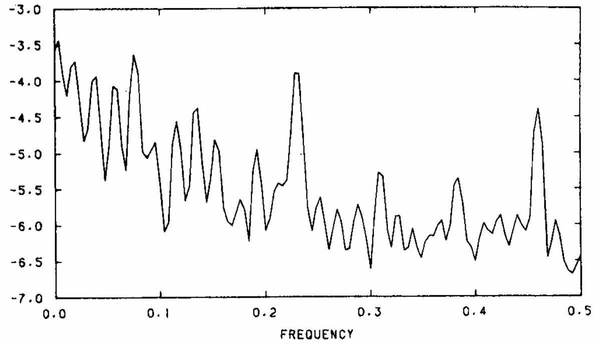


Figure 17. ADJUSTED WEEKLY RETAIL SALES

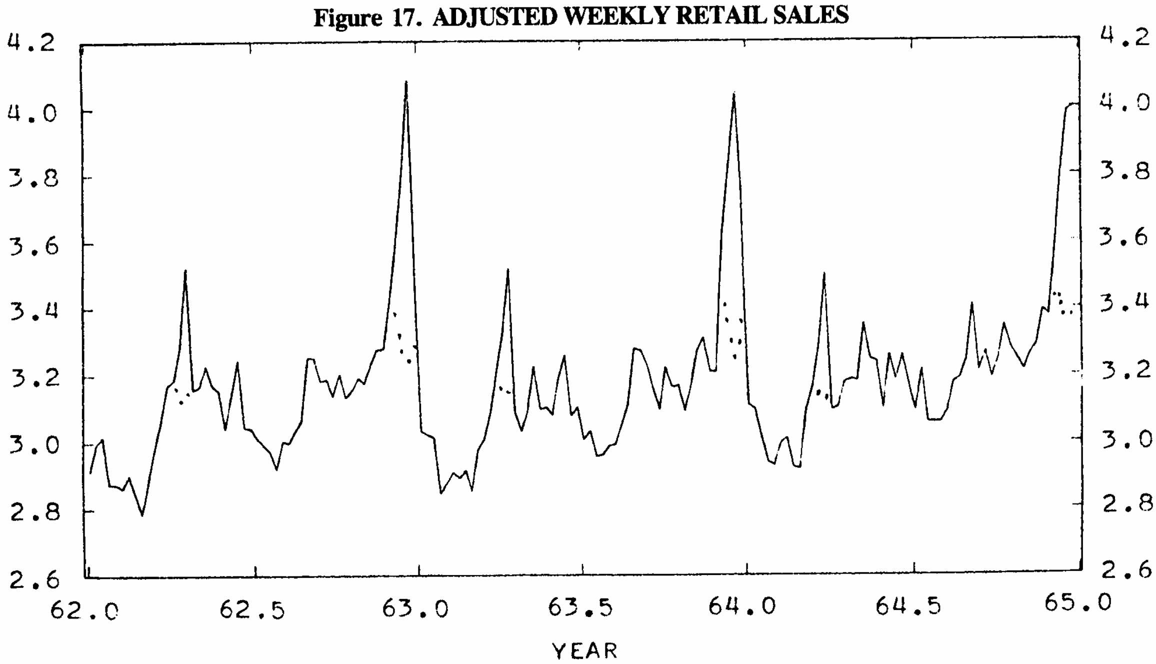


Figure 18. EASTER EFFECT ON RETAIL SALES

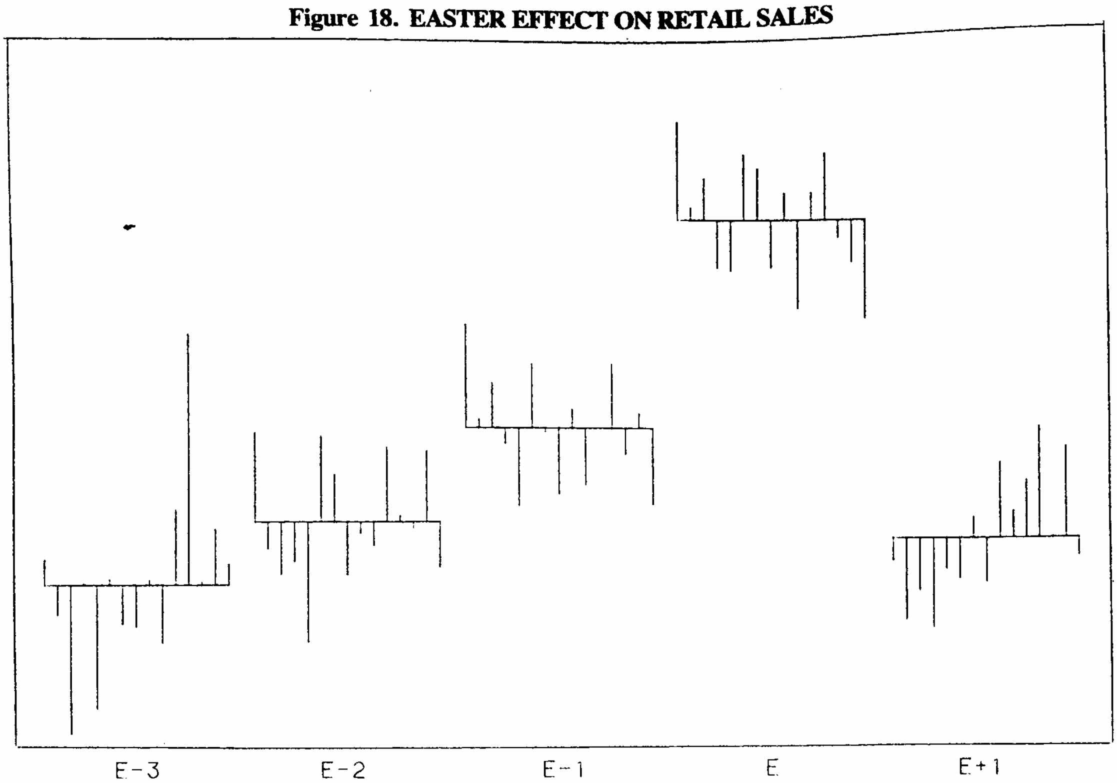
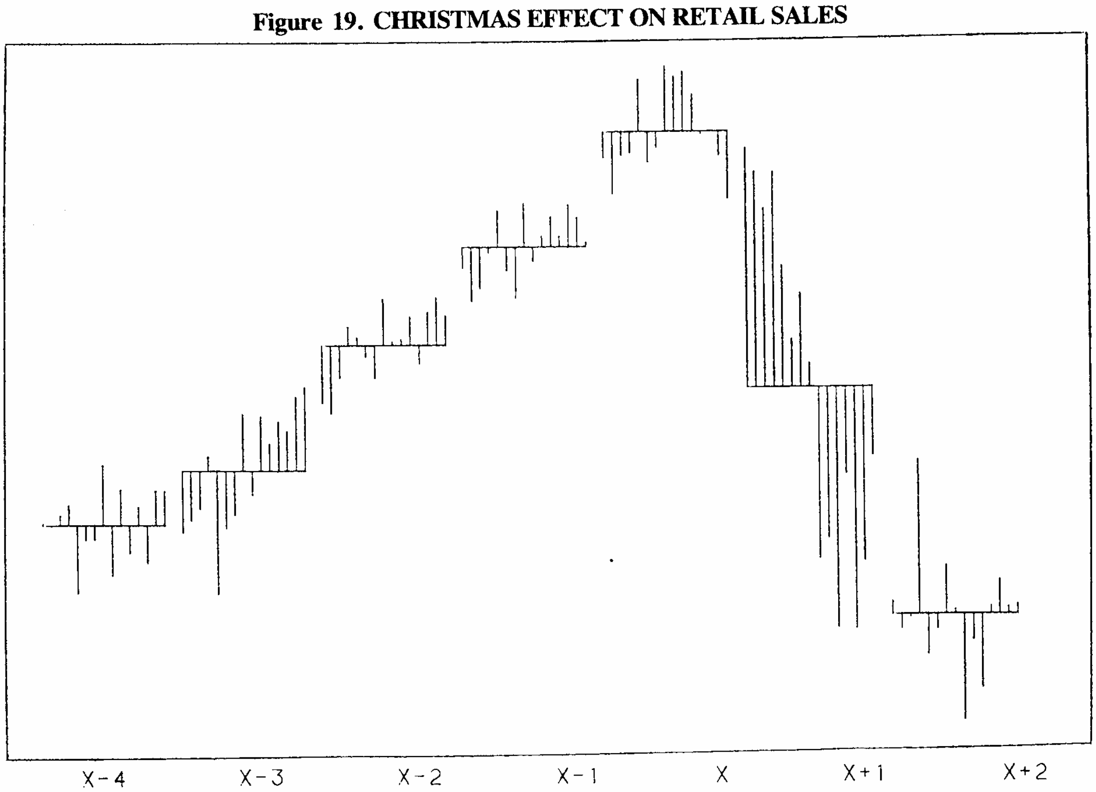


Figure 19. CHRISTMAS EFFECT ON RETAIL SALES



12. DISCUSSION

Some techniques for detecting and modeling calendar effects in economic data have been presented. They are necessary because ARIMA models cannot capture the irregular timing of the events. The problems of weekly data are more severe because of the number of annual events that can move and the possible complexities of the seasonal pattern. The fundamental period of the weekly seasonal is much longer with respect to the sampling rate and therefore closer to trend-cycle phenomena. A component model for the trend involving a $(1 - B)$ operator may not be appropriate. Long memory models being explored by Dempster and Granger may be needed.

REFERENCES

BOX, G. E. P.; HILLMER, S.; and TIAO, G. C. (1978), "Analysis and Modeling of Seasonal Time Series," *Proceedings of the Conference on Seasonal Analysis of Econometric Time Series*, Bureau of the Census, Washington, D.C.

BOX, G. E. P., and JENKINS, G. M. (1970), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

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.

CLEVELAND, W. P., and POLLNER, S. (1978), "Irregular Reporting in Economic Time Series," presented at the NBER conference on time series at Madison, Wisc., in May.

CLEVELAND, W. S., and DEVLIN, S. (1980), "Calendar Effects in Monthly Time Series: Detection by Spectral Analysis and Graphical Methods," *Journal of the American Statistical Association* 75, 487-96.

LIU, LON-MU (1980), "Analysis of Time Series With Calendar Effects," *Management Science* 26, 106-12.

NELSON, C. R. (1973), *Applied Time Series Analysis for Managerial Forecasting*, Holden-Day, San Francisco.

PIERCE, D. A. (1978), "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality Are Present," *Proceedings of the Conference on Seasonal Analysis of Economic Time Series*, Bureau of the Census, Washington, D.C.

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

COMMENTS ON "MODELING TIME SERIES WHEN CALENDAR EFFECTS ARE PRESENT" BY W. P. CLEVELAND AND M. R. GRUPE

Allan H. Young
Bureau of Economic Analysis

The authors examine the calendar effects in economic data within the context of ARIMA modeling. They consider three causes of variation in monthly data: (1) The length of the month and its effect on the amount of activity in the month relative to other months within the year; (2) the composition of the month, that is, whether the month contains four or five Sundays, four or five Mondays, and so on, and its effect on the amount of activity in the month relative to the same month in other years; (3) the date of Easter and its effect on the amount of Easter-associated activity that takes place in March relative to that in April. The authors point out that calendar effects must be specifically dealt with in seasonal adjustment and that ARIMA models, by themselves, are not appropriate tools.

In the broader context of economic data, as opposed to the narrower context of ARIMA modeling, much of the subject of this paper is well-traveled territory. For example, consider the work of the Bureau of the Census. By the midsixties, Census was adjusting retail sales series for three types of calendar effects: (1) The composition of the month, using the same type of technique described in Cleveland and Grupe's paper, (2) Easter variation, using the same technique described in the paper, which had been adopted from work at the Organization for European Cooperation and Development (OECD), and (3) Labor Day variation, a holiday not mentioned by the authors. Census was also examining the retail sales series for the effect of the length of the Christmas shopping season. Also, in this period, Census, as well as other agencies, was adjusting many other series for calendar effects.

Before considering the paper, let me interject some perspective on the features of the calendar that affect economic data. As the authors note, these features, and their effects, are not necessarily fixed over time. They also vary from culture to culture.

Our modern-day calendar contains 12 months of either 28, 29, 30, or 31 days in length. This calendar was established in the 16th century by Pope Gregory, who straightened out the earlier 12-month calendar developed by the Romans. The composition of a month affects the amount of monthly activity because economic activity generally does not take place at a constant rate over the entire week. This variation in activity goes back a long way. One finds in Genesis, ". . . and he rested on the seventh day from all his work which he had done."

Superimposed on the Gregorian calendar are holidays. Some of the holidays, such as Easter, are tied to the lunar calendar and, therefore, are not tied to a specific date and can even move between months; other holidays, such as Labor Day, are not tied to a specific date of the month. Also, the scheduling of the dates of some holidays, such as Washington's Birthday, has been changed in recent years, in this case, from a fixed date, February 22, to a fixed day of the week, the third Monday of the month.

The calendars of other cultures also have features that affect economic data. For example, in Saudi Arabia, the effects of the Feast of Ramadan are reflected in the economic data. Also note the article referenced by the authors, in which Liu examines the variation in the volume of Taiwan's highway traffic arising from the Chinese New Year. This holiday, the first day of the year in the Chinese calendar, flops back and forth between January and February in the Gregorian calendar, which is used in compiling the data.

Some retail merchants have chosen to define away the problem of the composition of the month by redesigning the calendar. They keep their books on the 4–4–5 plan for each quarter. Two months of 4 weeks each, followed by a month of 5 weeks, accounts nicely for the quarter. Under the 4–4–5 plan, the composition of a month does not vary from year to year. Does this clever scheme help us? No. Since it is not used by all merchants, the result is that either some 4–4–5 sales reports pollute the Census monthly data, or Census personnel, on the basis of incomplete information, must attempt to undo the retailer's data and place them back on the calendar of Pope Gregory. Additionally, Eastman Kodak and a few other firms keep their records on the basis of a 13-month year with 4 weeks in each month. In fact, Eastman Kodak rewrote X–11 to include 13 months per year.

The authors set forth three models for estimating the composition-of-month effect. The first, which they state is appropriate for flows, is essentially the same specification as that in X–11. In presenting the other two, one for monthend stocks and the other for daily average stocks, the authors move into less-explored territory. Their work implies that adjustments for the composition-of-month effect are needed for stock series; however, they do not demonstrate this need for monthend stocks. In addition, they present the results for the daily average money supply only in terms of spectra, which leave much to be desired in practical time series work,

Allan H. Young is Deputy Director, Bureau of Economic Analysis, U.S. Department of Commerce, Washington, D.C. 20230.

and F-ratios, which must be interpreted cautiously when computed from smooth series.

The general wisdom has been that monthend stock series such as inventories contain little, if any, composition-of-month effects. This has been backed up by tests, using the flow specification in X−11 to test for composition-of-month effects, that found little, if any, significant variation. I am not aware of the equivalent general wisdom, if any, for the money supply or what sort of testing may have been done previously. The published monthly seasonal adjustment of the money supply does not contain any adjustment for the composition of the month.

Why might one conclude that monthend stock series, such as inventories, would not be affected by the composition of the month? Essentially, for two reasons. First, consider the gross flow into and out of the stock rather than the net flow referred to by the authors. An example is a manufacturing plant that works a 5-day week. It receives shipments of raw materials 5 days a week; it builds up work-in-process and finished goods inventories by steadily applying labor and other inputs 5 days a week; and it ships its product 5 days a week. Thus, the stock will be steady, balanced by equal inflows and outflows. The stock will not vary with the composition of the month. Granted, the real world is not this simple: Raw materials may pile up on the rail siding on the weekend and be logged in on Monday, and there may be a big push to ship everything before Friday. Nevertheless, my example captures the essence of many stock situations.

Second, consider what happens if stocks build up in a recession and then are depleted in a boom. If the gross inflows and outflows have the same composition-of-month effect as they do in my example, the effect on the stock in the recession period will be of opposite sign to that in the boom. When estimated over both periods, the composition-of-month effect will tend to cancel out. Thus, for both reasons, one would expect any composition-of-month effect to be small in a series such as inventories. It would be limited to a net effect reflecting the extent to which gross additions outweighed gross deductions, or vice versa. Possibly, one might find more effect in disaggregated data for detailed industries or individual plants.

To me—a neophyte when it comes to money and banking—it appears that the above considerations should also apply to the daily average of the money supply; however, there may be some subtle differences. For example, the provision of currency before the weekend could, by operating through reserve requirements, shrink demand deposits by more than the increase in currency. Also, perhaps the effect on the money supply of the flow of Government securities to and from nonbanks varies with the composition of the month.

The practical question, which the authors do not examine, is, how important is the composition-of-month effect in monthly stock series? In what follows, I examine this question in terms of the two components of the money supply—currency and demand deposits—used by the authors, and in terms of monthend manufacturers' inventories of durable goods and of nondurable goods, as compiled by the Census Bureau. My investigation was limited by the time available and was hampered because it became apparent that the X−11 trading-day adjustment procedure, which I had intended to use, is less appropriate for stock series than I had initially thought—or than the authors perhaps imply—because of the specification for February. In assessing my results, one should also note the procedure developed by William S. Cleveland and Susan Devlin, at Bell Laboratories, for estimating trading-day variation. It provides a somewhat better estimate of the combined trading-day irregular component, from which one can estimate trading-day variation, than does X−11. The procedure allows for the preconditioning of the irregulars that result from the use of moving averages in prior steps. Nevertheless, I doubt that more refined analysis would lead to conclusions different from those that I derived. My basic finding is that, of the four series, only the currency component of the money supply has trading-day variation of any consequence.

I first examined the series for the composition-of-month effect using the X−11 trading-day adjustment procedure. The F-ratios shown in table 1 were obtained. The critical limit in the X−11 program for the F-ratio at the 1-percent level is about 3. Experience suggests that a somewhat higher value for the limit would be better. Thus, only the currency component appeared to have significant variation. The daily weights were very close to 1.0, compared with those estimated for flow series. Those for currency are shown in table 2.

Table 1. TEST FOR COMPOSITION-OF-MONTH EFFECT

Series	F-ratio
Currency, 1959−68	29.8
1969−79	10.6
Demand deposits, 1959−68	2.1
1969−79	3.5
Durable manufacturing inventories, 1959−79	1.8
Nondurable manufacturing inventories, 1959−79	.8

Table 2. DAILY WEIGHTS FOR CURRENCY

Day of the week	1959−68	1969−79
Monday	1.005	1.004
Tuesday	1.004	.971
Wednesday	.993	1.005
Thursday	.995	.981
Friday	1.036	1.017
Saturday	.991	1.007
Sunday	1.016	1.015

I wanted to carry the analysis one step further and ascertain the effect on the series of removing the estimated trading-day variation. The X−11 procedure, however, assumes that more activity occurs in a leap-year February than in a nonleap year. The adjustment factors, therefore, are not appropriate for a stock series. (Februaries were largely excluded as extremes in the regression estimation, so the daily weights and F-ratios were little affected.)

Given the problem with X−11, monthly trading-day adjustment factors were estimated for each series by classifying, by type of month, the monthly irregulars from a run with no trading-day adjustment. For example, all 31-day months beginning on Monday were placed in one group, all 31-day months beginning on Tuesday in another, and so on. Then, for each of the seven groups of 31-day months and the seven groups of 30-day months, the median was taken as the trading-day adjustment factor. The median for all nonleap-year Februaries was also used. For each leap year, the actual February irregular was used as the factor. The resulting monthly factors for currency for the period 1959-79 are shown in table 3.

It should be noted that 31-day months preceded by 31-day months might have been classified separately from those preceded by 30-day months or by Februaries. This type of refinement would have been closer to the authors' specifications for stocks; however, inspection of the currency series revealed that little would have been gained. The differences in the medians were virtually nil.

Table 3. TRADING-DAY ADJUSTMENT FACTORS BASED ON MEDIANS: CURRENCY, 1959–79

First day of month	31-day months	30-day months	Leap-year Februaries
Monday	1.0000	1.0003	1.0000
Tuesday	.9991	.9994	1.0004
Wednesday	.9997	.9987	(X)
Thursday	.9995	.9998	.9982
Friday	1.0010	1.0007	(X)
Saturday	1.0003	1.0000	.9990
Sunday	1.0003	1.0008	.9994
Nonleap-year Februaries: 1.0002			

X Not applicable.

Summary measures of the month-to-month variation in the trading-day component (obtained from the medians), irregular component, and seasonally adjusted component are shown in table 4. For the last two components, the summary measures are shown with and without the trading-day adjustment. Only the average month-to-month variation in the irregular component of the currency series was reduced more than 0.01 percent by the trading-day adjustment. Relative to the variation in the seasonally adjusted series, the trading-day variation was very small in each series. In contrast, the trading-day variation in retail sales was about twice as large as the variation in the seasonally adjusted series.

In reviewing the subject of the paper, it became apparent that, in some instances, the practices followed by the statistical agencies in making trading-day adjustments in flow series fall short of what is possible. For example, the authors cite a recent paper by W. S. Cleveland and S. Devlin that presented evidence of residual trading-day variation in manufacturers' shipments of durable goods as compiled by the Census Bureau. Further, Cleveland and Devlin indicated that Census used X−11 to make the trading-day adjustment. As the developer of the X−11 trading-day adjustment, I became curious. I found that the trading-day adjustment was not made by X−11 or by any other procedure that estimates the composition-of-month effect from the monthly data. Instead, Census adjusted data for detailed industries using external evidence concerning the activity on each day of the week.

Table 4. SUMMARY MEASURES OF STOCK SERIES, 1959–79

Series	$\overline{TD}$	$\overline{I}$ With TD	$\overline{I}$ W/o TD	$\overline{CI}$ With TD	$\overline{CI}$ W/o TD
Currency	0.08	0.10	0.14	0.54	0.54
Demand deposits	.07	.28	.28	.48	.48
Durable manufacturing inventories	.06	.25	.26	.79	.79
Nondurable manufacturing inventories	.05	.24	.24	.63	.64

Note: $\overline{TD}$, $\overline{I}$, and $\overline{CI}$ are averages, without regard to sign, of month-to-month percent changes in the trading-day, irregular, and seasonally adjusted components, respectively.

This type of procedure usually comes down to assuming a simple pattern of activity, such as a 5-day or 7-day week, with equal weights for each workday. Sometimes it closely approximates an adjustment for the composition of the month that is based on evidence contained in the monthly time series. In my experience, however, the evidence in the monthly data usually must be used to obtain an adequate adjustment. Simple assumptions or casual observation will not work.

I also found that the Federal Reserve Board (FRB) uses external evidence to adjust the production indexes for composition-of-month effects. The FRB seems to have originated this approach before World War II. Apparently one reason Census uses it is to make the shipment series comparable to the production series. At this point, it seemed worthwhile to examine both the Census and FRB series using the X−11 procedure. The results are summarized in table 5.

Table 5. SUMMARY MEASURES OF FLOW SERIES PREVIOUSLY ADJUSTED FOR TRADING-DAY VARIATION

Series	F-ratio	$\overline{TD}$	$\overline{I}$ With TD	$\overline{I}$ W/o TD	$\overline{CI}$ With TD	$\overline{CI}$ W/o TD
Durable manufacturing shipments, 1960–74	11.3	0.78	1.20	1.39	1.63	1.74
Nondurable manufacturing shipments, 1960–74	2.8	.21	.68	.71	.94	.98
Durable manufacturing production, 1960–74	1.2	.12	.77	.80	1.16	1.18
Nondurable manufacturing production, 1960–74	5.6	.23	.66	.69	.92	.96

Note: $\overline{TD}$, $\overline{I}$, and $\overline{CI}$ are averages, without regard to sign, of month-to-month percent changes in the trading-day, irregular, and seasonally adjusted components, respectively. The F-ratio and $\overline{TD}$ measure residual trading-day variation in the series previously adjusted for trading-day variation by the source agency. The residual trading-day variation is included in $\overline{I}$ and $\overline{CI}$ in the columns headed "W/o TD." It is removed in I and CI in the columns headed "With TD." The estimate of the residual trading-day variation does not reflect the difference between leap-year and nonleap-year Februaries because this variation was removed by the source agency.

In this table, the measures of trading-day variation are not of the total of this variation but of the residual that remains in the published series. The findings for the Census series on durable shipments agreed with those of Cleveland and Devlin, and they suggested that the series could be improved somewhat by basing the trading-day adjustment on the evidence contained in the monthly data. The problem seems to be that the external weights given the weekend are too heavy. The daily weights representing the residual variation are shown in table 6. Possibly the detailed nondurable components and the FRB detail could also be improved, although this is not apparent from the aggregates. Perhaps the FRB series fares better than the Census series in the test because, for some industries, production is measured by proxies that are not subject to composition-of-month effects.

Perhaps the deliberations of this conference will encourage Census and FRB to review their procedures for these particular series.

Table 6. DAILY WEIGHTS REPRESENTING RESIDUAL VARIATIONS

Day of the week	Weight
Monday	0.910
Tuesday	1.074
Wednesday	1.148
Thursday	.978
Friday	1.063
Saturday	.961
Sunday	.867

COMMENTS ON "MODELING TIME SERIES WHEN CALENDAR EFFECTS ARE PRESENT" BY W. P. CLEVELAND AND M. R. GRUPE

Charles Plosser
University of Rochester

Cleveland and Grupe discuss a number of phenomena that arise in both monthly and weekly time series that are related to the calendar. Broadly speaking, these phenomena can be categorized into composition effects and holiday effects. Composition effects in monthly data are primarily related to the number of days in the month and composition of those days in terms of the frequencies of Sundays, Mondays, etc. In weekly data, composition effects are related to the position of the week in the month. The effect of Easter is the holiday effect of most concern in the paper because it is a holiday that changes from year to year. The modeling strategy of Cleveland and Grupe is to construct a design matrix that can be used to estimate via regression the deterministic calendar effects.

In the first part of the paper, Cleveland and Grupe discuss composition effects on monthly data. The design matrix used to account for such phenomena is similar to that suggested by Young (1965) and more recently by Cleveland and Devlin (1980) and Hillmer (1981). In this setup, the monthly series is regressed on seven variables that indicate the number of each day of the week occurring in the month. Cleveland and Grupe are more careful than most, however, in making a distinction between flow and stock data and data that are constructed as averages of daily stocks. Nevertheless, as pointed out by the authors, even though the design matrices used to model calendar composition effects are different, they are highly correlated and one is likely to be as effective as another.

The modeling strategy for the Easter effect on sales assumes that there is a uniform increase in daily sales for 10 days before Easter. The total Easter effect is then allocated to either March or April in proportion to the number of days for this 10-day period occurring in the respective month. Unfortunately, the authors are not very clear as to how this total effect is estimated. It appears that it is the historical mean of the difference between April and March when Easter (and the 10 days before) occurs entirely in one month or the other. If so, then the procedure has the same difficulties that any adjustment procedure has that uses fixed weights computed from past data. Nonetheless, the technique is interesting and warrants further scrutiny.

My criticisms of the paper are of a more general variety. First, Cleveland and Grupe provide very little in the way of motivation as to why the problem of calendar effects is of interest and what are their ultimate objectives. Second, as might be expected when the objective of a procedure is ill-defined, the evaluation procedures are inadequate to arrive at an informed judgment as to the usefulness of the modeling strategies proposed.

These problems are certainly not unique to Cleveland and Grupe, but are common weaknesses in many papers that address seasonality issues. It is noteworthy that a major focus of the conference held here in 1976 concerned questions such as "what is seasonality?" and "why do we seasonally adjust economic data?" Unfortunately, the lessons learned from these discussions seem to be lost in the morass of techniques present in many recent papers on seasonal adjustment. One might consider such a comment as unfair. However, I believe it is important to reemphasize the importance of these questions in our study and evaluation of seasonal adjustment procedures.

In the context of the calendar effects discussed by Cleveland and Grupe, the question of why calendar effects are of concern must be answered before a modeling strategy can be pursued and evaluated. For example, are we interested in calendar effects because, by modeling them directly, forecasts of future values can be improved, or because we are interested in better modeling the relationships among economic time series, or because we want a better estimate of some "trend-cycle" component of the series? The modeling strategy and evaluation procedures clearly depend on which of these questions is being asked.

If improved forecasts is the objective, then Cleveland and Grupe may be on the right track by choosing design matrices that seem to mimic certain characteristics they apparently observed in the data. This is most likely to be the case in the context of forecasting a series based on its own history. Unfortunately, the authors never provide us with an example of the potential for forecast improvement. My guess is that for the series investigated by Cleveland and Grupe the potential is small simply because of the relative small amount of variance accounted for by calendar effects. Notice that the peaks in the spectra associated with the calendar effects are not even detectable in the raw data.

The second rationale suggested above is that by considering calendar effects, we might get a better understanding about the relationships among economic time series. This is a much more difficult question. It is not at all obvious that modeling deterministic calendar effects will be productive in this context. Particularly since these effects are likely to be common to many economic time series. Unfortunately, to adequately address these issues requires much more structure to the problem and generally would be done on a case by case basis incorporating some economic structure to the problem.

Finally, if the objective is adjustment to obtain an improved estimate of some trend-cycle component, Cleveland and Grupe should provide a more complete model of the unobserved components model being assumed and the consequences of their procedures. Although it is unclear from their discussion, it appears that Cleveland and Grupe carry out their procedures conditional on the removal of a trend and deterministic seasonal. One might ask how the transformations affect the decomposition strategy. It appears, for example, that Cleveland and Grupe assume that the decomposition is linear in the logs of the data. From an evaluation point of view, if the procedures proposed by the authors are to improve adjustment procedures, a comparison of their techniques with other methods is warranted. Do their methods improve (in some appropriately defined sense) on the methods in the X−11 or other smoothing techniques? Once again, the authors provide us with very little that would enable us to answer such questions.

In closing, let me add that I found the procedures advocated by Cleveland and Grupe potentially appealing at an intuitive level. Nonetheless, as the above comments suggest, I feel it would be preferable to define more precisely the ultimate objectives of the adjustment procedures so that appropriate evaluation techniques can be implemented.

REFERENCES

CLEVELAND, W. S., and DEVLIN, S. J. (1980), "Calendar Effects in Monthly Time Series: Modeling and Adjustment," Working Paper, Bell Laboratories.

HILLMER, S. C. (1981), "Modeling and Forecasting Time Series With Trading-Day Variation," Working Paper, University of Kansas, School of Business.

YOUNG, A. (1965), *Estimating Trading-Day Variation in Monthly Economic Time Series*, Technical Paper No. 12, Bureau of the Census.