THE EFFECT OF VARYING MONTH-LENGTH ON THE ANALYSIS OF ECONOMIC TIME SERIES

C.W.J. Granger

Econometric Research Program
Research Memorandum No. 41
30 July 1962

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

SUMMARY

In this paper, the effects due to the number of days in a month or the number of working days in a month not being constant are considered. It is found that for instantaneously-recorded series the effect is negligible, but that accumulated series may be noticeably altered when the ratio between the mean and standard deviation is large. For such series, the high frequencies are most affected and spurious coherence between two such series may arise.

1. Introduction.

Many economic time series are recorded monthly, but the length of a calendar month is not constant and the number of working days in a month varies quite considerably. It is the object of this paper to study the effect of such variations upon the spectrum of an economic series and upon the cross-spectrum between two series.

Two very different types of economic series can be distinguished: (a) instantaneously-recorded series, and (b) accumulated series.

The former are discrete series that can be thought of as being the values of a continuous time series at certain moments of time. The temperature at noon at a particular location is a perfect example of such a series. Economic series which may be considered to fall within this classification are price-series, interest rate series, and exchange rate series. In each case the variable can be thought of as existing at all moments of time, and our series is the value of the variable at the times chosen to record this value. It is the variation in the length of the month which could be disturbing in this case.

The second class of series includes those which represent the sum or accumulation of a variable since the previous recording was taken. A good example of such a series is rainfall measurements at a certain location. Examples of such series in economics are production data, import and export data, and volume of transactions data. It is clear that the number of working days in the month can have a profound effect on the recorded series.

In sections 2 and 3 the theoretical disturbances for the two types of series are considered separately, and the "real world" is then introduced in section 4. The final section briefly discusses methods of reducing the effect.

2. Instantaneously Recorded Series.

An instantaneously recorded series x_t may usually be considered as values taken from a continuous stationary time series $x(t)$ at specified equi-distant moments of time, $t = 0, 1, \ldots$. In this case Cramér [1] has shown that $x(t)$ and x_t can be represented by

$$x(t) = \int_{-\infty}^{\infty} e^{it\omega}\, dz(\omega) + m \tag{2.1}$$

$$x_t = \int_{-\pi}^{\pi} e^{it\omega}\, dz(\omega) + m \tag{2.2}$$

where in each case $dz(\omega)$ is a random function with zero mean and with independent increments. Thus $E[x(t)] = E[x_t] = m$.

If, now, the continuous series $x(t)$ is sampled at time units $t + \epsilon_t$, $t = 0, 1, \ldots$ where ϵ_t is a stationary series with zero mean, the resulting series y_t can be (at least approximately) represented by

$$y_t = \int_{-\pi}^{\pi} e^{i\omega(t+\epsilon_t)}\, dz(\omega) + m \, , \tag{2.3}$$

where

$$E[dz(\omega_1)\ \overline{dz(\omega_2)}] = 0\ , \quad \omega_1 \neq \omega_2$$

$$E[dz(\omega)\ \overline{dz(\omega)}] = f(\omega)\ d\omega\ ,$$

$f(\omega)$ being the power spectrum of the original instantaneously recorded series x_t based on constant time intervals.

Clearly, $E[y_t] = m$, but the power spectrum of y_t will not necessarily be the same as that for x_t . We henceforth, without any loss of generality, assume $m = 0$ for ease of discussion.

In this case

$$E[y_t\ \overline{y}_{t-\tau}] = \int_{-\pi}^{\pi} e^{i\tau\omega}\ E[e^{i\omega(\epsilon_t - \epsilon_{t-\tau})}]\ f(\omega)\ d\omega\ . \tag{2.4}$$

Thus $E[y_t\ \overline{y}_t] = \int_{-\pi}^{\pi} f(\omega)\ d\omega$, i.e., the variance of the series y_t is also the same as the variance of x_t .

If we assume ϵ_t to be a Gaussian series, i.e., the k-dimensional distribution of any subset $\epsilon_{t_1}, \epsilon_{t_2}, \ldots, \epsilon_{t_k}$ is Gaussian, then the theory of characteristic functions gives

$$E[\exp\ \{i\omega(\epsilon_t - \epsilon_{t-\tau})\}] = \exp\ \{-\omega^2\sigma^2(1 - \rho_\tau)\}\ ,$$

where

$$\left.\begin{aligned} \sigma^2\rho_\tau &= E[\epsilon_t\ \overline{\epsilon}_{t-\tau}] \\ \sigma^2 &= E[\epsilon_t^{\ 2}] \end{aligned}\right\} \tag{2.5}$$

and we let

$$\rho_\tau = \int_{-\pi}^{\pi} e^{i\tau\omega}\ g(\omega)\ d\omega\ . \tag{2.6}$$

Thus

$$\begin{aligned} E[y_t\ \overline{y}_{t-\tau}] &= \int_{-\pi}^{\pi} e^{i\tau\omega}\ e^{-\omega^2\sigma^2}\ \exp(\rho_\tau\ \omega^2\sigma^2)\ f(\omega)\ d\omega \\ &= \int_{-\pi}^{\pi} e^{i\tau\omega}\ e^{-\omega^2\sigma^2}\ f(\omega)\ d\omega \\ &\quad + \int_{-\pi}^{\pi} e^{i\tau\omega}\ e^{-\omega^2\sigma^2} \sum_{k=1}^{\infty} \frac{(\omega^2\sigma^2)^k}{k!}\ \rho_\tau^{\ k}\ f(\omega)\ d\omega\ . \end{aligned} \tag{2.7}$$

If we note that

$$\rho_\tau^{\,k} = \int_{-\pi}^{\pi} e^{i\tau\omega} \, g_k{*}(\omega) \, d\omega$$

where $g_k{*}(\omega)$ is the $k^{\underline{th}}$ convolution of the spectral function of ϵ_t , i.e.,

$$g_k{*}(\omega) = \int_{-\pi}^{\pi} g(\omega-\lambda) \, g_{k-1}^{*}(\lambda) \, d\lambda$$

$$g_1{*}(\omega) = g(\omega) \; ,$$

then using the convolution theorem for products of functions having Fourier integral representations, (2.7) becomes

$$E[y_t \, \overline{y}_{t-\tau}] = \int_{-\pi}^{\pi} e^{i\tau\omega} \, [f(\omega) + \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{k!} \int_{-\pi}^{\pi} \lambda^{2k} \, f(\lambda) \, g_k{*}(\omega-\lambda) d\lambda] \, e^{-\sigma^2\omega^2} \, d\omega \; .$$

It thus follows that the power spectrum of the series y_t is

$$\varphi(\omega) = e^{-\sigma^2\omega^2} \, [f(\omega) + \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{k!} \int_{-\pi}^{\pi} \lambda^{2k} \, f(\lambda) \, g_k{*} \, (\omega-\lambda) \, d\lambda] \; . \qquad (2.8)$$

As σ^2 will usually be extremely small, the sum of the first two or three expressions in the series provides a very good approximation to $\varphi(\omega)$.

If ϵ_t is white noise, (2.8) becomes

$$\varphi(\omega) = e^{-\sigma^2\omega^2} \, f(\omega) + A \qquad (2.9)$$

where A is a constant such that

$$\int_{-\pi}^{\pi} [A + e^{-\sigma^2\omega^2} \, f(\omega)] \, d\omega = \int_{-\pi}^{\pi} f(\omega) \, d\omega \; .$$

With small σ^2 , we see immediately that for this case the effect on the shape of the spectrum of having variable month-length will be very slight. The actual effect will be fully discussed in section 4.

The steps for studying the effect on the cross-spectrum are similar. If the undisturbed series are $x_t^{(1)}$, $x_t^{(2)}$ with power spectra $f_1(\omega)$, $f_2(\omega)$ respectively, and with (complex) power cross-spectrum $c(\omega)$, then the power cross-spectrum for the two series $y_t^{(1)}$, $y_t^{(2)}$ observed at the time points $t + \epsilon_t$, $t = 0, 1, \ldots$ is given by

$$c_y(\omega) = e^{-\sigma^2\omega^2} [c(\omega) + \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{k!} \int_{-\pi}^{\pi} \lambda^{2k} c(\lambda) g_k^*(\omega-\lambda) d\lambda] . \qquad (2.10)$$

In general, the difference between $c_y(\omega)$ and $c(\omega)$ will be small, and, in particular, we note that if $c(\omega) \equiv 0$ then $c_y(\omega) \equiv 0$ also, and so no spurious coherence is generated.

3. Accumulated Series.

For accumulated series a complete investigation of the effect becomes less manageable. For instance, if the daily production series is z_t and if the number of working days in a month is given by the series N_t, to be accurate we would need to consider the new series formed by

$$y_t = \sum_{j=N_{t-1}+1}^{N_t} z_j . \qquad (3.1)$$

However, a useful approximation, which is likely to be sufficiently close to indicate the more important features of the effect, can be formed by considering the "real" series y_t to be a product of two stationary series. Suppose that it were possible to divide the year into twelve units, all containing an equal number of working days. If the accumulated series were based on these units, the resulting series x_t may be called the "undisturbed" series. The effect of the number of working days in a month not being a constant can be discussed in terms of the difference between the power spectrum of x_t and that of y_t. The approximation we make is to assume that

$$y_t = \frac{N_t x_t}{N} , \qquad (3.2)$$

where N_t is the series of the number of working days in a month. We further assume $E[N_t] = N$ to be the number of days in the "perfect" month over which x_t is measured. (In practice, of course, N will not necessarily be an integer.) We also need to assume that $\{x_t\}$ and $\{N_t\}$ are independent series.

To study y_t defined by (3.2) we require the following theorem:

Theorem.*

If $\{x_t\}$, $\{y_t\}$ are stationary, Gaussian stochastic processes having zero means, power spectra $f_x(\omega)$, $f_y(\omega)$, and power cross-spectrum $c(\omega) = c_o(\omega) + iq(\omega)$, then the power spectrum of the process $\{z_t\}$ defined by

$$z_t = x_t\, y_t$$

is given by

$$f_z(\omega) = \int_{-\pi}^{\pi} f_x(\omega-\lambda)\, f_y(\lambda)\, d\lambda + \int_{-\pi}^{\pi} c_o(\omega-\lambda)\, c_o(\lambda)\, d\lambda + \int_{-\pi}^{\pi} q(\omega-\lambda)\, q(\lambda)\, d\lambda \;. \tag{3.3}$$

The theorem follows immediately, using Fourier transforms and the convolution theorem, from a formula due to Isserlis [3]:

$$\text{cov}\,(X_tY_t,\, X_sY_s) = E[X_tX_s]\, E[Y_tY_s] + E[X_tY_s]\, E[Y_tX_s] \;.$$

Thus, if x_t, y_t are independent series,

$$f_z(\omega) = \int_{-\pi}^{\pi} f_x(\omega-\lambda)\, f_y(\lambda)\, d\lambda \;.$$

Returning to (3.2), it should be noted that

$$E[y_t] = E\left[\frac{N_t}{N}\right] E[x_t]$$

under the assumption of independence, and so

$$E[y_t] = E[x_t] = m \;, \text{ say} \;. \tag{3.4}$$

If $f_x(\omega)$, $f_N(\omega)$ are the power spectra of the series $\{x_t\}$, $\{N_t\}$ respectively, then, as

$$\frac{N_t x_t}{N} = \left(\frac{N_t}{N} - 1\right)(x_t-m) + m\left(\frac{N_t}{N} - 1\right) + (x_t-m) + m \;,$$

---

*I wish to thank Professors E. Parzen and E. J. Hannan for pointing out this theorem to me.

it follows from the theorem that the power spectrum of the series y_t is given by

$$f_y(\omega) = \frac{1}{N^2} \int_{-\pi}^{\pi} f_x(\omega-\lambda)\, f_N(\lambda)\, d\lambda + f_x(\omega) + \frac{m^2}{N^2} f_N(\omega) . \tag{3.5}$$

Further, the variance of y_t is given by

$$\sigma_y^2 = \sigma_x^2 + \frac{m^2}{N^2} \sigma_N^2 + \sigma_x^2 \frac{\sigma_N^2}{N^2} . \tag{3.6}$$

Whereas the effect was slight for instantaneously recorded series, equations (3.5), (3.6) indicate that the fact that the number of working days in a month is not constant can have important implications for accumulated series. This can be particularly well illustrated by considering some extreme cases. If $\sigma_x^2 = 0$ (i.e., $x_t = m$, all t), then $\sigma_y^2 = \frac{m^2}{N^2} \sigma_N^2$ and so the effect is to move from a series with no variance to one in which the variance could be considerable. If $\{x_t\}$ is white noise with mean comparable in size with the variance, the power spectrum of the resulting series y_t is $f_y(\omega) = a + \frac{m^2}{N^2} f_N(\omega)$, which may be very different from the "true" spectrum $f_x(\omega) = \text{const.}$

The effect is seen to be of particular importance when the mean of the underlying series x_t is large compared to the variance of x_t. Discussion of the effect in the real world is delayed until the next section.

The effect on cross-spectral analysis may also be important.

Suppose we have two "basic" series $x_t^{(1)}$ and $x_t^{(2)}$ with means m_1, m_2; power spectra $f_{x,1}(\omega)$, $f_{x,2}(\omega)$, and power cross-spectrum $c^x(\omega) = c_o^x(\omega) + iq^x(\omega)$. The two observed series can now be approximated by

$$y_t^{(1)} = \frac{N_t}{N} x_t^{(1)}$$

$$y_t^{(2)} = \frac{N_t}{N} x_t^{(2)} ,$$

assuming the same series N_t to operate in both cases.

Consider

$$E[(y_t^{(1)} - m_1)(\bar{y}_{t-\tau}^{(2)} - m_2)] = \frac{1}{N^2} E[N_t \bar{N}_{t-\tau}] \; E[x_t^{(1)} \bar{x}_{t-\tau}^{(2)}] - m_1 m_2$$

$$= \frac{1}{N^2} \rho_\tau^{(N)} \mu_x(\tau) + \mu_x(\tau) + \rho_\tau^{(N)} \frac{m_1 m_2}{N^2} \; . \qquad (3.7)$$

where

$$E[N_t \bar{N}_{t-\tau}] = \rho_\tau^{(N)}$$

$$E[x_t^{(1)} \bar{x}_{t-\tau}^{(2)}] = \mu_x(\tau) \; ,$$

then taking Fourier transforms and using the convolution theorem, the power cross-spectrum between the series $\{y_t^{(1)}\}$, $\{y_t^{(2)}\}$ is found to be given by

$$c^y(\omega) = c^x(\omega) + f_N(\omega) \frac{m_1 m_2}{N^2} + \frac{1}{N^2} \int_{-\pi}^{\pi} f_N(\lambda) \; c^x(\omega-\lambda) \; d\lambda \; ,$$

and thus the co-spectrum and quadrative spectrum are given by

$$c_o^{\,y}(\omega) = c_o^{\,x}(\omega) + f_N(\omega) \frac{m_1 m_2}{N^2} + \frac{1}{N^2} \int_{-\pi}^{\pi} f_N(\lambda) \; c_o^{\,x}(\omega-\lambda) \; d\lambda$$

$$q^y(\omega) = q^x(\omega) + \frac{1}{N^2} \int_{-\pi}^{\pi} f_N(\lambda) \; q^x(\omega-\lambda) \; d\lambda \; .$$

Generally, the convolution terms will be negligible compared to the others, but the term $f_N(\omega) \; m_1 m_2/N^2$ need not be negligible. If we consider the extreme case when the series $\{x_t^{(1)}\}$, $\{x_t^{(2)}\}$ are independent, then $c_o^{\,x}(\omega) = q^x(\omega) \equiv 0$ and

$$c_o^{\,y}(\omega) = f_N(\omega) \frac{m_1 m_2}{N^2} \; , \quad q^y(\omega) = 0 \; .$$

Thus the coherence will not vanish, and one effect of a varying number of working days is seen to be that a spurious coherence can arise, thus apparently indicating a relationship between the two series.

The quantity

$$A(\omega) = \frac{c_o^{\,2}(\omega) + q^2(\omega)}{f_1(\omega) \; f_2(\omega)} \; ,$$

called the coherence, is a direct measure at a given frequency of the amount by which the two series under consideration are related at that frequency. Similarly, the quantity

$$\varphi(\omega) = \tan^{-1} \frac{q(\omega)}{c_o(\omega)}$$

is a direct measure of the lag (or lead) between the two series at a given frequency.*

Using the notation and results proved above, but ignoring the convolution terms as being probably negligible, we find

$$A^y(\omega) = \frac{(c_o^x(\omega) + f_N(\omega)\, m_1 m_2/N^2)^2 + (q^x(\omega))^2}{[f_1(\omega) + f_N(\omega)\, m_1{}^2/N^2][f_2(\omega) + f_N(\omega)\, m_2{}^2/N^2]} \qquad (3.8)$$

which may be considerably different from the coherence for the undisturbed series

$$A^x(\omega) = \frac{(c_o^x(\omega))^2 + (q^x(\omega))^2}{f_1(\omega)\, f_2(\omega)} \quad .$$

Similarly, we have

$$\varphi^y(\omega) = \tan^{-1} \left[\frac{q^x(\omega)}{c_o^x(\omega) + f_N(\omega)\, m_1 m_2/N^2} \right] \qquad (3.9)$$

instead of

$$\varphi^x(\omega) = \tan^{-1} \left[\frac{q^x(\omega)}{c_o^x(\omega)} \right] .$$

As $f_N(\omega)\, m_1 m_2/N^2 \geq 0$ (it being reasonable to assume the mean of an accumulated series to be positive), it is seen, for instance, that when $c^x(\omega)$ is positive the lag between the two series is reduced.

Clearly the effect under discussion can be of a considerably disrupting nature when spectral and cross-spectral methods are being used on accumulated data. It will be especially important when the ratio between

*For a description of the use of these and related statistics on both stationary and non-stationary economic series see Granger [2].

the mean and the standard deviation of the undisturbed series are not small. If the effect is important in spectral analyses it will be at least as important in alternative methods of investigating time series, such as the fitting of autoregressive models, but the nature in which it disrupts the method is likely to be more subtle and difficult to investigate.

The problem of improving the data to reduce the effect will be discussed in section 5.

4. Consideration of the Actual Effect.

The series consisting of the number of days in a month is primarily the numbers 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31 continually repeated. The only change from year to year is due to the lengthening of February by one day each leap year. Clearly, the major component of such a series is the annual component. This component is, however, such that all the harmonics are of roughly equal importance. Taking into account the leap years, the average month-length is 30.437 days with variance 0.70. The first autocorrelation is approximately -0.42 and the second autocorrelation is positive. These facts allow us to discuss the more obvious properties of the series ϵ_t introduced in section 2. We see that $\sigma_\epsilon^2 = \frac{0.70}{(30.437)^2} = 0.00076$ and that the spectral density of the series will have a high value at the Nyquist frequency, will generally be larger for high frequencies than for low frequencies, and will have peaks at frequencies corresponding to harmonics of the annual component. With $\sigma_\epsilon^2 = .00076$, the term $\exp\{-\omega^2\sigma_\epsilon^2\}$ changes from 1 at $\omega = 0$ to 0.991 at $\omega = \pi$, and so the effect of this term in equation (2.8) is to emphasize slightly the low frequencies, but the effect is so small as to be negligible.

To consider the effect on accumulated series of the variable number of working days, series were constructed of the actual number of days worked per month in the United Kingdom in the 30-year period 1933 to 1962. One series assumed a 5-day working week and the other a $5\frac{1}{2}$-day week. Weekends and all public holidays were taken into account but not annual (summer) holidays or strikes. The resulting series had the following characteristics:

<u>5-day week</u>

$$\text{mean} = 21.15 \text{ days, variance} = 1.51 \text{ days}$$

$$\rho_1 = -0.51 \quad , \quad \rho_2 = 0.15$$

<u>$5\frac{1}{2}$-day week</u>

$$\text{mean} = 23.27 \text{ days, variance} = 1.46 \text{ days}$$

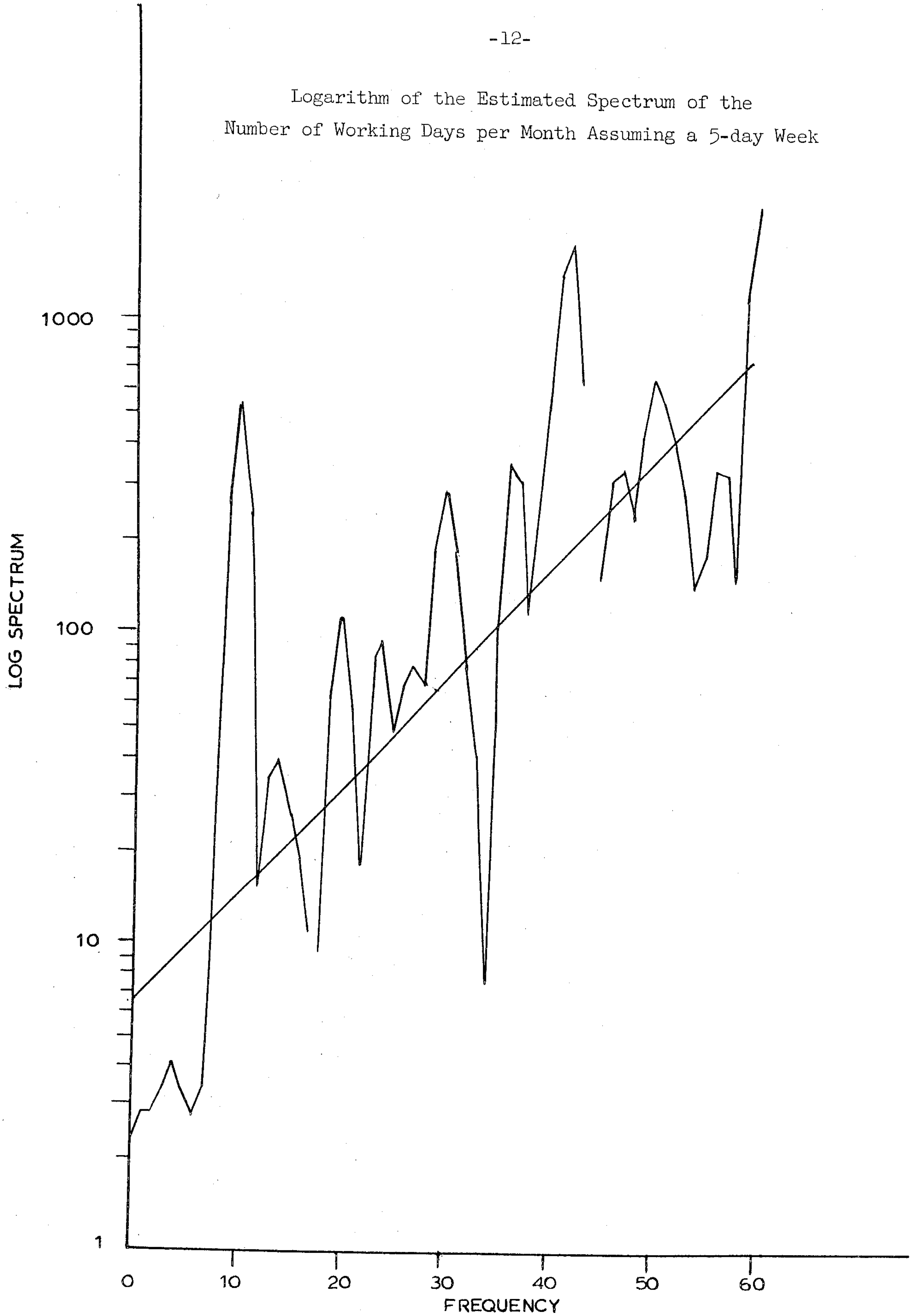
$$\rho_1 = -0.49 \quad , \quad \rho_2 = 0.23$$

(If similar series had been constructed for the United States, the means would have been lower and the variances greater, as there are more public holidays in the United States than in Britain.)

The power spectra of each of the series were calculated on an IBM 650 computer using the Tukey estimate (see Blackman and Tukey [4] or Granger [2]) over 60 frequency bands. As the major features of the two resulting curves were similar, only the estimated spectrum for the series assuming a 5-day week is illustrated. (See diagram on following page.)

The diagram illustrates the plot of $\log \hat{f}(\omega)$ against frequency ω. (The gap at frequency $\frac{44\pi}{60}$ is due to the estimating procedure giving a negative estimate.) As would be expected, the higher frequencies are considerably more important than the lower frequencies, and there are important peaks at frequencies $\frac{2\pi}{12}$ (due to annual component) and π (due to ρ_1 being large in magnitude and negative). The high peak at frequency $\frac{42\pi}{60}$ can be explained as corresponding to the alias of the weekly cycle. It was noted above that the average number of days in a

Logarithm of the Estimated Spectrum of the
Number of Working Days per Month Assuming a 5-day Week



month is 30.437 and so there are 4.348 weeks in the average month. As we are recording monthly, a frequency 4.348 will "alias" into the frequency 0.348. Such a frequency corresponds to the 42nd point when a spectrum is estimated over 60 frequency bands as $2 \times 60 \times 0.348 = 41.76$ and 42 is the integer nearest 41.76.

If, for a moment, we ignore the peaks around the 10th and 42nd frequency bands, the actual effect on an accumulated series of the varying number of working days can be quickly seen. Equation (3.6) shows that the effect makes an important difference to the variance if $m^2\sigma_N^2/N^2$ is not negligible compared to σ_x^2. Using the figures for a 5-day week, this means that $0.0034\ m^2$ is not negligible compared to σ_x^2 or that m/σ_x is of the order of 17. Thus, for example, if $m/\sigma_x = 5.4$ the effect would increase the variance of the accumulated series by approximately 10 percent.

When the ratio m/σ_x is such that the effect is not negligible, equation (3.5) indicates that the major change in the power spectrum will be an increase in the high frequencies. One might note that the convolution term will, in general, also chiefly affect the high frequencies. Experience indicates that the majority of economic series have a large peak around zero frequency even when an efficient method of trend removal has been used. This peak will be moved to the Nyquist frequency on convolution due to the major peak at the Nyquist in $f_N(\omega)$. Thus, most aspects of the effect are inclined to raise the higher frequencies of the spectrum of the accumulated series. The effects due to the other peaks in $f_N(\omega)$ are less easily determined, as they chiefly depend on the shape of $f_x(\omega)$. One would certainly expect the harmonics of the annual component to be emphasized in the spectrum of the observed series y_t.

Similarly, equations (3.8) and (3.9) suggest that the coherence and angle diagrams will be chiefly affected at the high frequencies.

It is unusual for a series to be chiefly disturbed at the higher frequencies rather than at the low frequencies. In a sense the results are comforting, as it is the important low frequency components which chiefly concern policy makers. However, when cross-spectral techniques are being used and particularly when short-run prediction is being attempted, the effect being studied could be of considerable importance if the series is such that the ratio m/σ_x is large.

5. Methods of Reducing the Effect.

As, for instantaneously recorded series, the effect is slight, and as there appears to be no obvious way of improving the series once it has been recorded, no further discussion is required for this type of series.

Although the approximation used for discussing the effect on accumulated series is probably a reasonably good one for most purposes, unfortunately it helps only slightly in discussing methods of improving the data. Given the series y_t and assuming that the series N_t is also known, the approximation suggests that we can reconstruct x_t merely by considering

$$z_t = \frac{N}{N_t} y_t \; . \qquad (5.1)$$

If the approximation were exact, $z_t \equiv x_t$, but as it is only an approximation, the transformation of the series y_t in equation (5.1) will not entirely remove the effect. However, this transformation will certainly considerably reduce the effect and will be well worth making for many series.

Some published data has already been so transformed; the Board of Trade in London, for instance, uses it on some import and export series. When the data is "raw" and the effect is likely to be serious the transformation should certainly be made, especially when an electronic computer is being used, as the extra calculation involved will be of little consequence.

However, to use equation (5.1) one needs the series N_t. The number of days that ought to have been worked in any month for a given factory, say, can be easily deduced by subtracting a day from the total in the month for every Sunday and a further day or half-day for each Saturday. For any given country only a little more effort is required to take into account public holidays, and so, in theory, a series N_t can be constructed based on either a 5-day or a $5\frac{1}{2}$-day week. Unfortunately the problem is more complex, as annual holidays and strikes must be taken into account. The "rules" for public holidays are not entirely uniform. If, for instance, a public holiday of not first class importance falls on a Saturday, do the workers get only the half-day off or is a further half-day also not worked? In the latter case does this lie on the Friday or the Monday, which could be in a different month? Are the local semi-official public holidays observed as well as the national ones?

Even more important in complicating the construction of the series N_t is the problem of aggregation. If the series y_t is the steel production for the United States, this will be the sum of individual series from many sources, each of which will have its own series for the number of days worked in every month. The "true" N_t would need to be a weighted average over all the individual series. In such a case, the approximation used in section 3 will be less satisfactory, and any method of reducing the effect by using a transformation of type (5.1) will also be less effective.

Clearly, no definite rule can be stated as to whether or not the transformation (5.1) should be used; it must be left to the investigator to decide if it is appropriate to his individual problem. It would, perhaps, not be asking too much to ask the producers of statistics to consider if the series of an accumulated form could not be so improved at source.

REFERENCES

[1] Cramér, H. On harmonic analysis in certain functional spaces. *Ark. Mat. Astr. Fys.* 28 B (1942).

[2] Granger, C.W.J. *Analysis of economic time series*. (In preparation.)

[3] Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12, pp. 134-139, 1918.

[4] Blackman, R. B. and Tukey, J. W. *The measurement of power spectra*. New York, 1959.