Seasonal Adjustment of Vital Statistics by Electronic Computer

HARRY M. ROSENBERG

THE MONTHLY variation in the frequency of vital events has long been the subject of scientific study and popular conjecture (*1–4*). In this country there has been remarkable stability during the past decade in the rhythmic yearly variation of the birth rate, the marriage rate, and the death rate. These seasonal patterns, while of great intrinsic interest, pose a problem for the analyst who is interested in determining the underlying trend of a time series.

Knowledge of the trend on a monthly basis, free from distortion due to large seasonal fluctuations, facilitates comparisons between different types of series to study; for example, the influence of business cycles on marriage and birth rates (*5*), the relation between excess mortality associated with influenza epidemics and the risk of fetal deaths (*6*, *7*), and the epidemiology of heat waves (*8*). Death rates and deaths due to influenza and pneumonia, free from seasonal factors, serve as quantitative indices of the impact of pneumonia-influenza epidemics (*9*, *10*).

In general, information about the current trend in birth, marriage, and death rates provides a foundation on which to base estimates of future developments in time series of vital statistics. To establish the level of trend from monthly data, the task of the analyst is to estimate the seasonal factor and to eliminate it from the original observations.

Several methods have been developed to do this mostly in connection with economic data (*11*). "Same-month-year-ago" comparisons are perhaps the most familiar method used. This involves eliminating the seasonal factor by dividing the figure for a given month by the figure for the same month in the previous year, and expressing the result as some percentage change. This method, discussed later, has serious limitations. Two other techniques, the monthly means and the link relatives, were among the first "refinements"; but they too give approximate results. The ratio-to-moving-average method, with its advantages of greater flexibility and precision, has been adopted by many groups engaged in large-scale seasonal adjustment work despite the fact that it is quite laborious.

In 1954 the Bureau of the Census, and more recently the Bureau of Labor Statistics, developed electronic computer programs for seasonally adjusting time series, using an adaptation of the standard ratio-to-moving-average method (*12–14*). The application of these programs to time series of vital statistics yields useful information about their characteristic seasonal patterns and underlying trend. This paper describes the method developed by the Bureau of the Census (version X–9 of Census Method II) and its application to national vital statistics collected by the Division of Vital Statistics, National Center for Health Statistics, Public Health Service.

Method

The ratio-to-moving-average method, of which the Bureau of the Census method is an adaptation, proceeds on the assumption that the total variation of a time series can be "decomposed," or broken down, into four com-

Mr. Rosenberg is with the Division of Vital Statistics, National Center for Health Statistics, Public Health Service.

ponents, the trend (T), the cycle (C), the seasonal component (S), and the irregular component (I).

The cycle, under this mode of time series analysis, is defined as short-run movements made up of alternating periods of increase or decrease. In economic time series, cycles usually last for several years, from 2 to 10, depending upon the type of series. The trend consists of still longer run movements and ordinarily has little effect on month-to-month movements of a series. In short-term analysis, it has been convenient to combine trend and cycle, referring to the component as the trend-cycle, following the economist's usage, or just as the trend.

The seasonal factor consists of intra-year movements which are repeated more or less regularly each year. Irregular fluctuations, under this scheme, are those that remain after systematic fluctuations have been accounted for. These residual fluctuations, are, by definition, short in duration and exhibit no regular seasonal pattern.

A centered 12-month moving average is the preliminary basis of the ratio-to-moving-average method for eliminating the seasonal and irregular components from time series. The 12-month moving average is a series of averages which embraces, first, the first 12 months of a series; next, the 2d to 13th months; then the 3d to 14th months; and so on. A centered 12-term moving average is a 2-term average of a 12-term average. The centered average is taken as an estimate of $T \times C$, because it smoothes out by averaging the seasonal movement, S, and, for the most part irregular movements, I, since the latter are largerly movements of small amplitude and short duration. Figures 1–3 show both the 12-month moving average and the crude monthly data for time series of birth rates, marriage rates, and death rates (deaths from all causes), demonstrating the effectiveness with which the moving average describes the underlying trend. The birth and marriage trends are smoother than the mortality trend, which in its "swell" reflects the inclusion of high mortality rates associated with the influenza epidemic in 1957–58.

The ratio of the original data to the 12-month

Figure 1. U.S. birth rates by month, on annual basis, and 12-month moving average of the birth rates, 1957–60

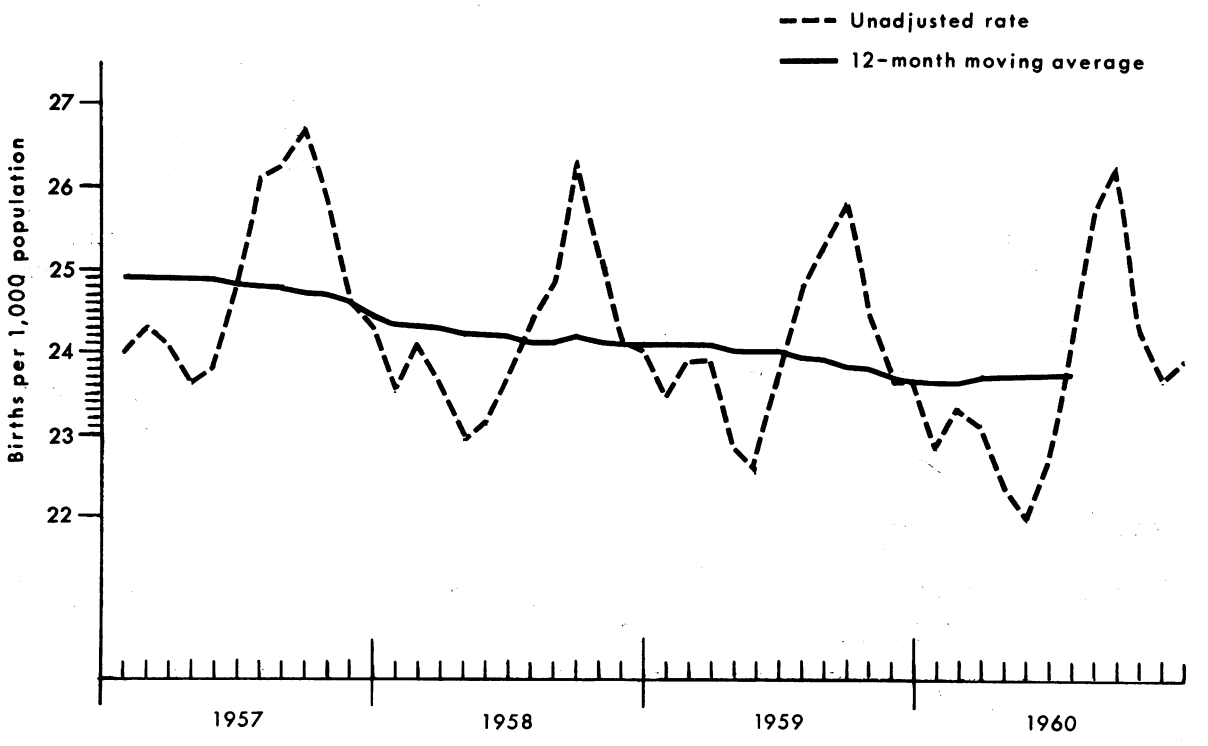
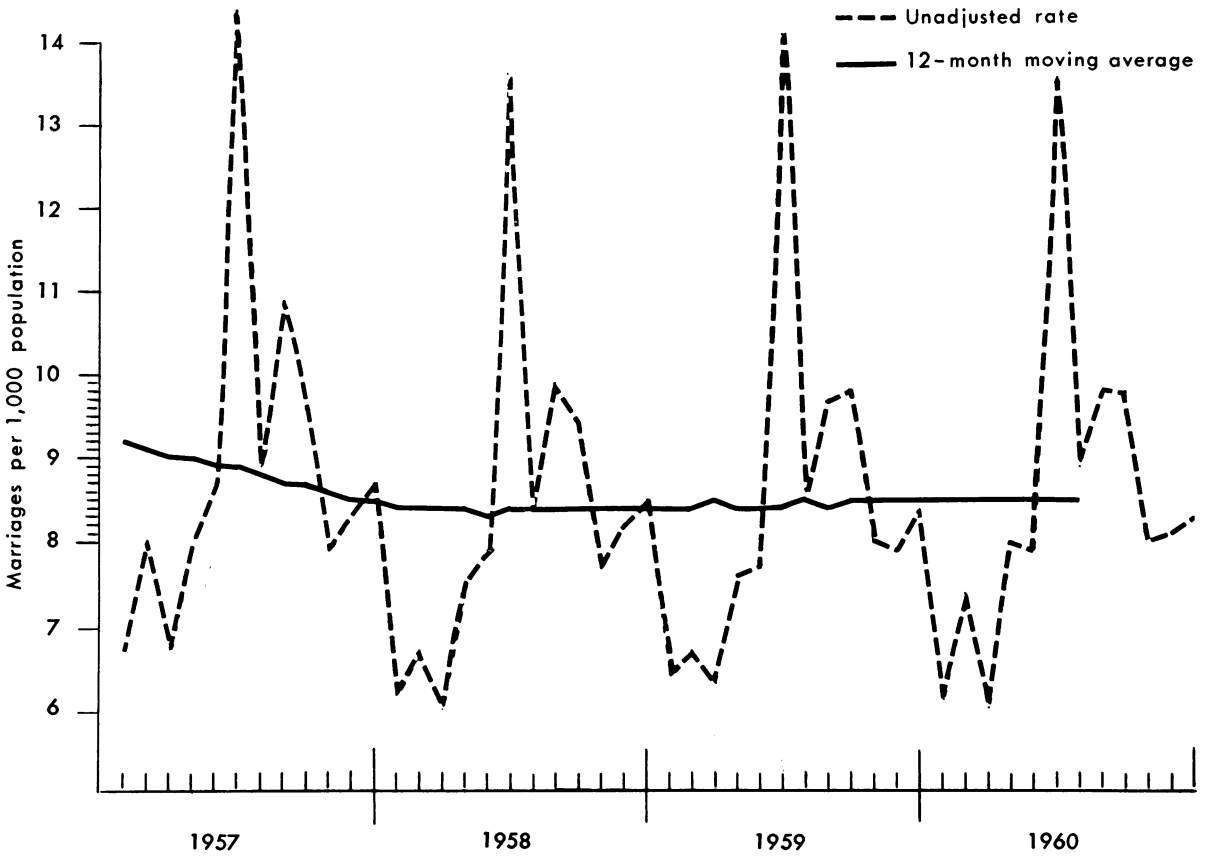


Figure 2. U.S. marriage rates by month, on annual basis, and 12-month moving average of the marriage rates, 1957–60



moving average for each month is an estimate of the seasonal and irregular components for that month, thus:

$$\frac{T\times C\times S\times I}{T\times C}=S\times I$$

Isolation of the seasonal component alone is accomplished by a number of additional averaging and smoothing operations. The resulting schedule of seasonal adjustment factors, S, when divided into the original observations for each month, produces a series, $T\times C\times I$, relatively free from the distortion of seasonal fluctuations.

The procedure developed by the Bureau of the Census uses the capacities of the electronic computer for making statistical computations. The method outlined below describes the X–9 version of Census Method II.

$S\times I$ ratios are computed by the conventional ratio-to-moving-average technique, described above. Among the ratios for each month, for the years of the series, extreme values are identified and replaced; and a smooth curve is fitted to the adjusted ratios. The resulting ratios are a first estimate of the seasonal component alone, S. The preliminary seasonal factors are divided into the corresponding figures of the original series, month by month; that is, the seasonal factor for January 1957 is divided into the original observation for January 1957, and so on. This yields the preliminary seasonally adjusted series. A weighted 15-month moving average of the preliminary seasonally adjusted series is computed as an estimate of the trend-cycle, $T\times C$. Ratios of the original observations to the 15-month moving average are computed. Once again extreme values are identified and replaced, and a curve is fitted to the adjusted ratios. These smoothed figures constitute the final seasonal adjustment factors, S.

Results

The birth rate for 1957–60, broken down into its seasonal trend-cycle and irregular components, is shown in figure 4. The seasonal component clearly emerges as the strongest component of the monthly fluctuations, tracing a bimodal pattern with a major peak in September and minor peak in February. For 1954–63 there was an average seasonal difference of more than 14 percent in the birth rate from the lowest to the highest month (table 1). In terms of numbers, this represents a difference of 53,000 live births, from 323,000 in April to 376,000 in September.

The amplitude of the seasonal pattern for the marriage rate (fig. 5) is greater than that for births and deaths, rising more than 60 percent above the annual average in June and declining 25 percent below the annual average in March.

The seasonal pattern of the death rate in figure 6 describes a movement similar to a sine curve which peaks during the winter months and troughs during the autumn. The variation

Figure 4. U.S. birth rates by month, raw data, and seasonal, cyclical, and irregular components, 1957–60

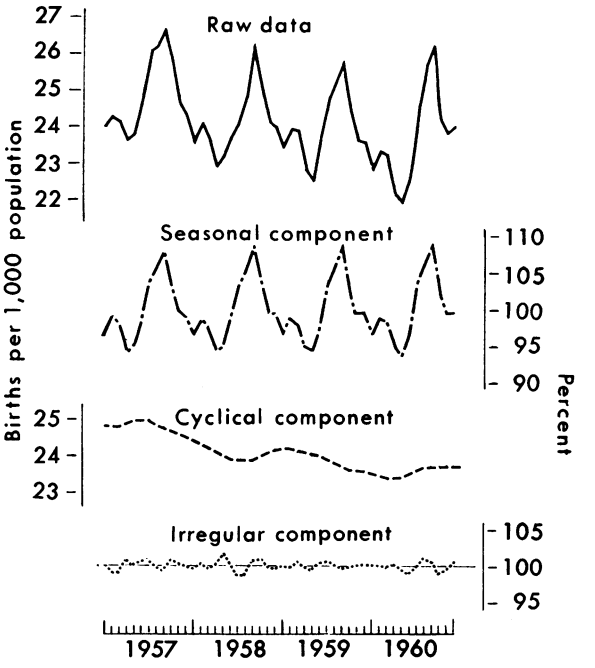


Figure 3. U.S. death rates by month, on annual basis, deaths from all causes, and 12-month moving average of the death rates, 1957–60

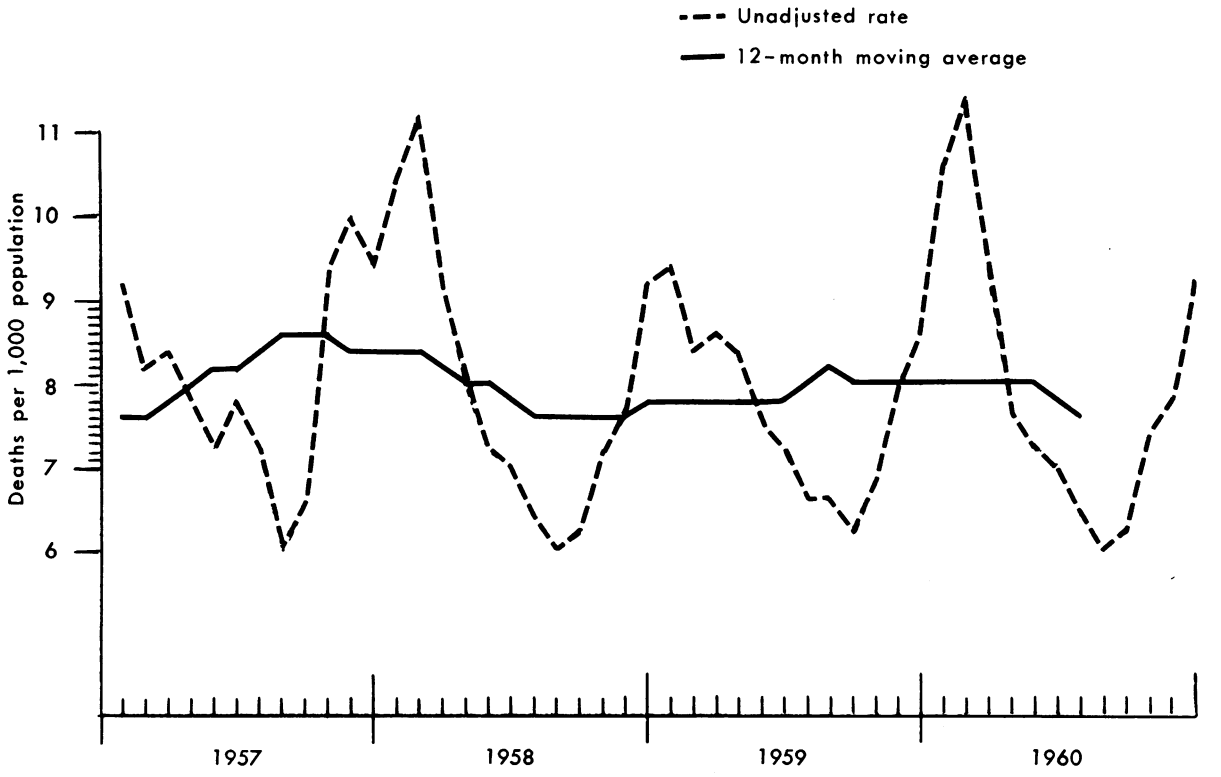


Table 1. Average seasonal adjustment factors for U.S. birth rates, marriage rates, and death rates for selected periods

Month	Series and years		
	Birth rate, 1954–63	Marriage rate, 1953–61	Death rate, 1951–60
January	96.4	73.9	107.9
February	99.1	85.2	105.0
March	97.6	73.4	103.8
April	94.4	91.9	101.8
May	94.7	93.1	98.5
June	97.7	160.5	98.4
July	103.9	100.9	95.7
August	106.3	112.7	92.2
September	108.5	114.8	92.2
October	103.1	94.7	96.4
November	99.5	96.8	101.4
December	99.3	102.3	106.7

NOTE: The data have been corrected for the varying number of days in each month before seasonal adjustment. Marriage data include additional adjustments for variation in the Lenten season and for the number of Saturdays in each month. Mortality data were modified to exclude high mortality rates that occurred during major influenza epidemics.

between the highest and lowest seasonal factors was about 16 percent during the period 1951–60 (table 1).

An irregular component is present in each series, contributing relatively little to the monthly fluctuations of the marriage rate, somewhat more for the birth rate, and accounting for about a fourth of the total variation in the death rate, as computed from data in table 2, which shows the average absolute monthly change in amplitude of the components. While the absolute magnitude of the irregular component of the marriage rate is largest of the series, it is small in relation to the seasonal component, only a tenth as large as the seasonal component ($\bar{I}/\bar{S}=0.12$). For births the irregular fluctuations are about a third as large, on the average, as the seasonal; and for deaths more than half as large as the percentage change in amplitude of the seasonal component. In time series of births, marriages, and deaths, major but unmeasurable factors contributing to the irregular component are random, such as the vicissitudes of the data collection and processing systems.

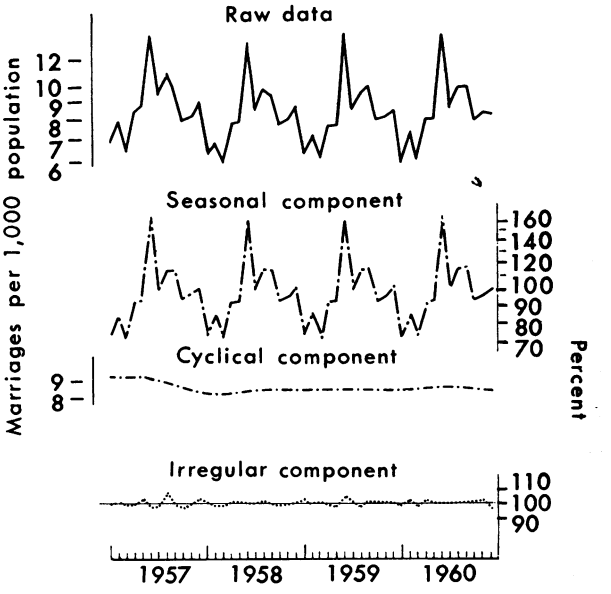
The underlying trends of the three series, reflected in the trend-cycle components, trace relatively smooth paths (figs. 4–6). In comparison with the other components, the trend-cycle contributes little to monthly fluctuations. For births, marriages, and deaths, the average absolute monthly percentage change in the trend-cycle is less than 0.5 percent.

Analysis of Current Trend

Using monthly vital statistics rates for a period of about 10 years, the seasonal adjustment program developed by the Bureau of the Census can decompose historical data into seasonal, trend-cycle, and irregular components, providing information on the stability or shifts in the seasonal pattern over the years and on intra-year turning points in the trend. Quantitative information on the components of monthly time series is useful for evaluation of the current trend, as well as the past trend, offering an alternative to "same-month-year-ago" comparisons in analysis of current data.

Removal of the seasonal component from current data by dividing original observations by seasonal adjustment factors yields a relatively smooth series representing the trend-cycle and irregular components. Applying to these values a moving average of appropriate length suppresses the irregular component, leaving an ap-

Figure 5. U.S. marriage rates by month, raw data, and seasonal, cyclical, and irregular components, 1957–60



proximation of the trend-cycle. For retrospective data, the trend-cycle is estimated by a weighted 15-month moving average, described earlier. Figure 7 illustrates how information on time-series components is used to seasonally adjust current natality data. The original monthly birth rates are represented by a dashed line, the seasonally adjusted series by a heavy line, and the underlying trend—a 4-month moving average of the seasonally adjusted series—by a dotted line.

The original observations, representing $T \times C \times S \times I$, have been divided by the seasonal adjustment factors, S, in table 3, for the appropriate months and years. Adjustment factors for 1964 were estimated by adding to the factors for 1963 one-half the trend between 1962 and 1963, permitting the seasonally adjusted series to be kept up to date as new data become available each month. The choppiness of the resulting series, $T \times C \times I$, reflects to some extent errors in the seasonal adjustment factors that were estimated for the current year; however, the largest part of this residual movement is due to the presence of the irregular component.

Table 2 shows the size of the irregular component in relation to the trend-cycle, expressed in terms of the average absolute month-to-month percentage change in the components. The ratios of the irregular component to the trend-cycle ($\bar{I}/\bar{C}$) show that in series of the birth rate, the marriage rate, and the death rate, fluctuations in the irregular component are from 4 to 8 times as large as the underlying trend. In order to suppress the irregular component, we can determine the number of months that must go by before the trend-cycle, which is cumulative in the short-run, dominates the irregular factor, which is not cumulative (*12a*); that is, the number of months that must be averaged for the $\bar{I}/\bar{C}$ ratio to fall below 1 and stay there. For the birth rate 4 months are required, thus:

Months' span	*$\bar{I}/\bar{C}$ ratio*
1	3.38
2	2.21
3	1.46
4	.94

Comparisons of the seasonally adjusted birth rate over a 4-month period, therefore, are required for the trend-cycle to dominate the irregular component, indicating that comparisons of month-to-month percentage change are less reliable than comparisons of 4-month periods for birth rates.

Figure 6. U.S. death rates by month, raw data, and seasonal, cyclical, and irregular components, 1957–60

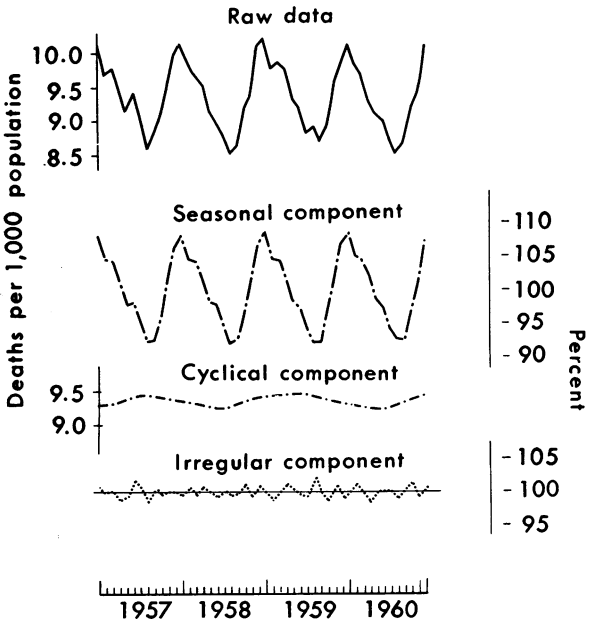


Table 2. Percentage change in average monthly amplitudes for time series components and relations among components of the U.S. birth rates, marriage rates, and death rates for selected years

Components and their relations	Series and years		
	Birth rate, 1954–63	Marriage rate, 1953–61	Death rate, 1951–60
Components:			
Raw data ($\bar{O}$)	2.96	19.96	3.24
Irregular component ($\bar{I}$)	.75	2.43	1.54
Trend-cycle ($\bar{C}$)	.25	.40	.20
Seasonal component ($\bar{S}$)	2.82	19.60	2.71
Relations between components:			
$\bar{I}/\bar{C}$	3.00	6.08	7.70
$\bar{I}/\bar{S}$	.27	.12	.57
$\bar{S}/\bar{C}$	11.28	49.00	13.55

NOTE: The averages represent month-to-month percentage change without regard to sign. The relationship among the components can be approximated by the formula from reference 15: $\bar{O}^2 \doteq \bar{C}^2 + \bar{S}^2 + \bar{I}^2$.

In taking 4-month moving averages of seasonally adjusted birth rates, 2 months are lost. For marriage rates, a 6-month moving average of the seasonally adjusted values is necessary to suppress irregular movements, losing 3 months' currency. The larger irregular component in mortality series is also smoothed by taking a 6-month moving average.

For all three types of data, smoothing of the seasonally adjusted data yields an approximation of the underlying trend on a current basis, relatively free from the distortion of irregular and seasonal fluctuations.

Discussion

The merits of seasonal adjustment by electronic computer are apparent to those already engaged in making these computations by the standard, time-consuming manual methods. Electronic computer programs developed for this purpose have been adopted by a number of organizations both here and abroad, and are used by the Federal Government to adjust many current economic indicators. (For example, *Economic Indicators*, prepared monthly for the Joint Economic Committee by the Council of Economic Advisors, and *Business Cycle Developments*, prepared monthly by the Economic Research and Analysis Division, Bureau of the Census.)

Adjustment of vital statistics for seasonal variation by this method, however, has generally been limited to retrospective studies (*16*) and to the annual publication of U.S. marriage and birth statistics. An exception is the development of a control chart procedure for comparing weekly numbers of deaths from 108 U.S. cities with an expected weekly number by the Communicable Disease Center of the Public Health Service. The expected number of deaths, based on several years' data, is estimated by a least squares line representing secular trend combined with a sine curve fitted to the average seasonal pattern of mortality (*17*).

Usually, monthly birth, marriage, and death statistics are analyzed by comparing the data with levels prevailing in the same month of the previous year. This method of eliminating the seasonal factor involves dividing the figure for

Figure 7. U.S. birth rates by month, unadjusted and adjusted for seasonal variation, 1961–64

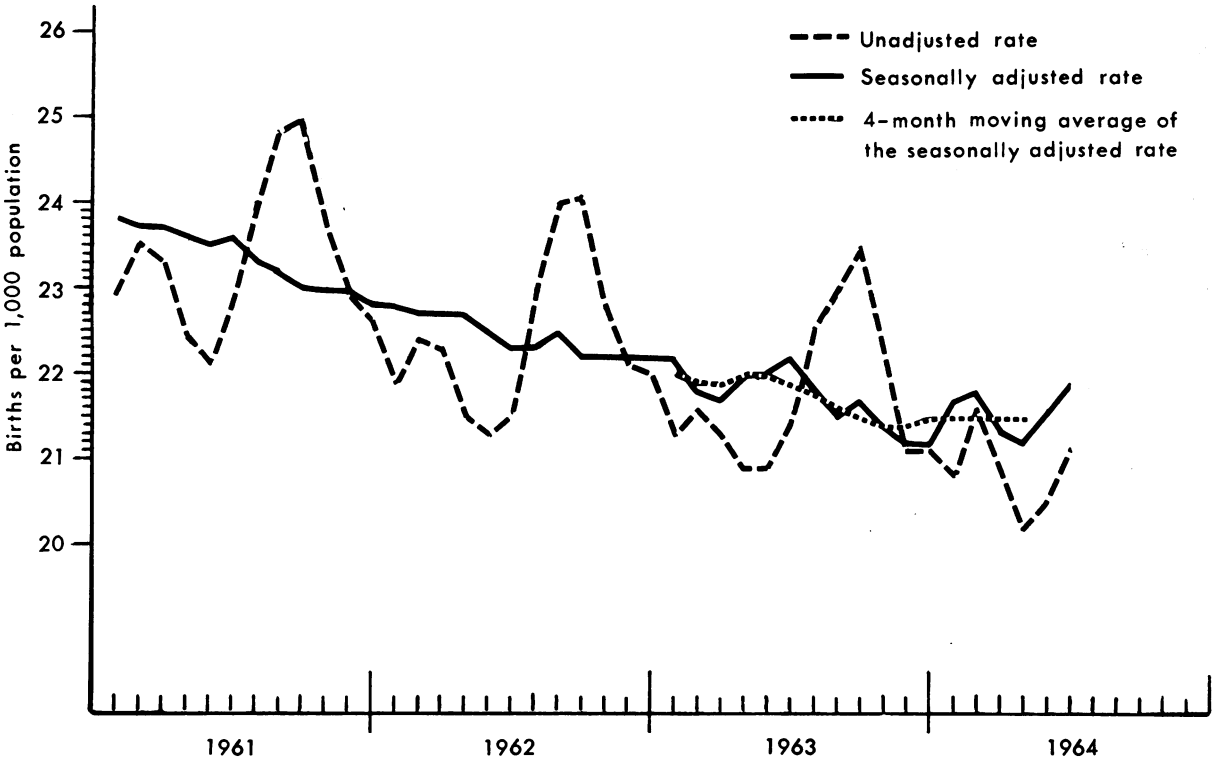


Table 3. Seasonal adjustment factors of U.S. birth rates, marriage rates, and death rates for selected years

Series and years	January	February	March	April	May	June
Birth rate:						
1963	95.9	98.9	98.2	95.1	94.9	96.4
1962	96.1	98.9	98.3	95.1	94.6	96.4
1961	96.4	98.9	98.4	94.9	94.1	96.6
1960	96.5	98.9	98.3	94.7	93.9	96.9
1959	96.4	98.8	98.0	94.5	94.2	97.6
1958	96.4	98.9	97.4	94.4	94.8	98.2
1957	96.4	99.0	97.0	94.2	95.0	98.5
1956	96.5	99.2	96.8	93.9	95.0	98.6
1955	96.5	99.3	96.7	93.7	94.9	98.6
1954	96.4	99.3	96.7	93.5	94.8	98.7
Marriage rate:						
1961	73.0	84.3	73.7	90.9	93.0	163.5
1960	73.2	84.2	73.5	91.0	93.1	163.4
1959	73.5	84.0	73.1	91.5	93.2	163.2
1958	74.0	84.0	72.8	91.7	93.4	162.7
1957	74.3	84.4	72.8	92.0	93.4	161.6
1956	74.5	85.2	73.0	92.2	93.5	159.8
1955	74.4	86.1	73.4	92.6	93.2	158.4
1954	74.5	86.7	73.6	92.9	92.9	157.6
1953	74.5	87.1	73.7	93.1	92.6	157.3
Death rate:						
1960	108.4	105.0	104.6	102.0	98.7	97.2
1959	108.4	104.8	104.5	102.1	98.5	97.5
1958	108.2	104.7	104.5	102.2	98.3	97.9
1957	108.0	104.4	104.3	102.1	98.0	98.2
1956	107.9	104.4	104.0	101.7	98.1	98.2
1955	108.0	104.4	103.6	101.3	98.4	98.3
1954	108.0	104.9	103.4	101.2	98.6	98.9
1953	107.7	105.3	103.3	101.4	98.7	99.2
1952	107.4	105.8	103.0	101.7	98.7	99.2
1951	107.2	106.0	102.8	101.9	98.8	98.8

Series and years	July	August	September	October	November	December
Birth rate:						
1963	103.3	106.5	108.2	103.8	99.4	99.3
1962	103.3	106.8	108.4	103.4	99.4	99.3
1961	103.3	106.9	108.7	102.9	99.5	99.4
1960	103.5	106.6	109.0	102.6	99.4	99.6
1959	103.8	105.7	108.9	102.9	99.4	99.6
1958	104.1	105.2	108.7	103.3	99.4	99.3
1957	104.3	105.5	108.3	103.4	99.4	99.0
1956	104.4	106.2	108.1	103.0	99.4	98.9
1955	104.3	106.7	108.1	102.6	99.5	99.0
1954	104.4	106.8	108.2	102.4	99.5	99.1
Marriage rate:						
1961	101.3	116.0	117.0	----	----	----
1960	101.3	115.8	116.5	92.5	96.3	100.1
1959	101.2	115.6	115.5	92.8	96.2	100.6
1958	101.1	115.0	114.6	93.3	96.2	101.5
1957	101.0	114.0	113.8	94.1	96.5	102.4
1956	100.8	112.3	113.5	95.0	96.9	103.5
1955	100.7	110.6	113.5	96.0	97.4	103.9
1954	100.6	109.0	113.7	96.8	97.7	104.2
1953	100.7	108.1	113.8	97.3	97.9.	104.2
Death rate:						
1960	94.2	92.7	92.4	97.0	100.8	106.9
1959	94.5	92.5	92.5	96.8	100.9	106.9
1958	94.7	92.3	92.7	96.8	100.9	106.8
1957	95.5	92.4	92.8	96.5	101.0	106.4
1956	96.4	92.5	92.6	96.4	101.1	106.3
1955	97.4	92.6	92.0	96.1	101.3	106.4
1954	97.2	92.0	91.8	96.1	101.5	106.7
1953	96.4	91.7	91.6	96.1	101.7	106.8
1952	95.7	91.5	91.7	96.2	102.1	106.8
1951	95.5	91.7	91.7	96.3	102.4	106.7

a given month by the figure for the same month of the previous year, and expressing the result as some percent change. Figure 8 shows the trend of the birth rate for the period 1961–64 implied by "same-month-year-ago" comparisons. The 100 percent line is used as a base; the line represents the date at the end of the period of comparison. The difficulty of interpreting the trend in "same-month-year-ago" comparisons arises mainly from sharp month-to-month fluctuations, which obscure the underlying movements of the series (*12b.*)

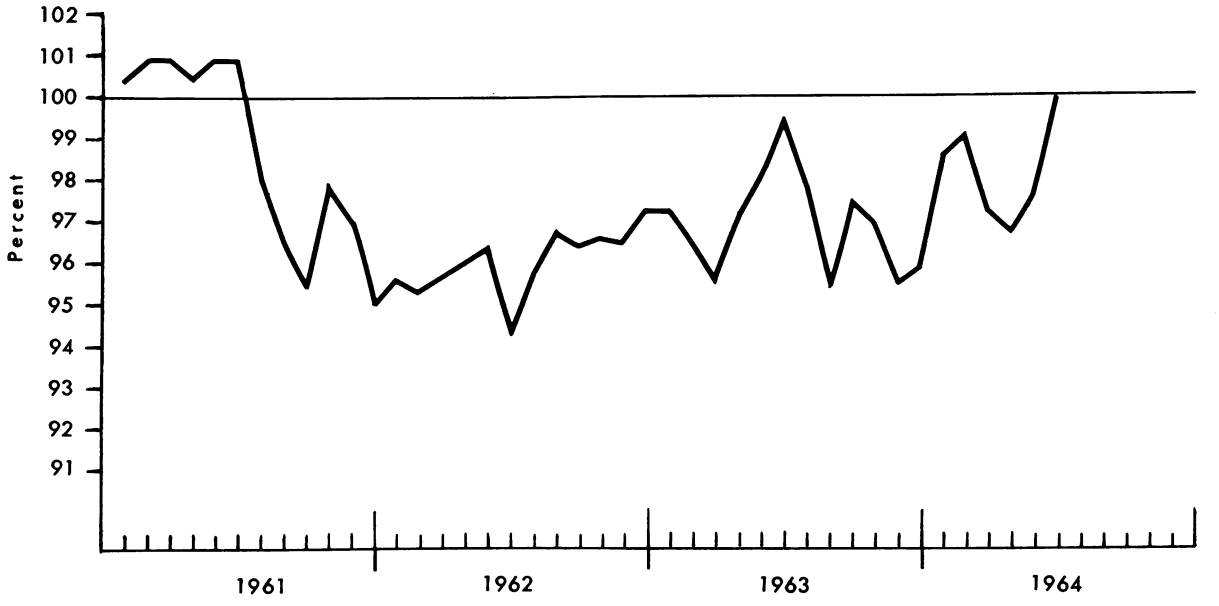
The success achieved with the electronic computer for seasonal adjustment of birth and marriage data has encouraged us to adopt this procedure on a limited scale for current analysis. *Health, Education, and Welfare Indicators*, prepared by the Office of the Assistant Secretary for Legislation, U.S. Department of Health, Education, and Welfare, shows seasonally adjusted birth and marriage rates for the most recent month for which data are available. Seasonal adjustment of mortality statistics still presents a number of conceptual and statistical problems which may be overcome by refinements of present computer programs, or by different methods.

One difficulty is that the scheme of time-series analysis used was developed for economic data; it is not designed to cope with extreme values of several months' duration, such as the high mortality rates associated with influenza epidemics. Unless such high values are eliminated subjectively prior to seasonal adjustment, the procedure adopted in this paper, they will be allocated mainly to the trend component. A more appropriate representation of mortality time-series components might include a fourth type of fluctuation, an epidemic component (E), along with seasonal, trend-cycle, and irregular components. Also, the 12-month moving average may be too flexible a first estimate of trend for mortality. A least squares line through the original data may be a better estimate.

Summary

Birth rates, marriage rates, and death rates presented on a monthly basis have exhibited relatively stable seasonal patterns over a period of years. A method was developed recently by the Bureau of the Census, using an electronic computer, to establish the nature of these seasonal patterns and to remove the identified seasonal component from time series of monthly data. The method is basically an adaptation of a standard technique, the ratio-to-moving-average method, for seasonally adjusting time series of economic data.

Figure 8. Trend of U.S. birth rates using same-month-year-ago comparisons, 1961–64



The application of the Bureau of the Census method to time series of the birth rate, the marriage rate, and the death rate is useful in the restrospective and current analysis of vital statistics. Good results are indicated for birth and marriage data, but present methods are not entirely suitable for dealing with mortality statistics. Continuing research on this subject should yield better approximations of the trend-cycle and an adequate method for isolating the epidemic component of mortality time series.

NOTE: Specifications of the X–9 version of Census Method II are available upon request from the Office of the Chief Economic Statistician, Bureau of the Census, Washington, D.C., 20203. Specifications of the BLS Seasonal Factor Method are available from the Deputy Associate Commissioner for Systems Analysis, Bureau of Labor Statistics, Washington, D.C., 20210.

REFERENCES

(*1*) Huntington, E.: Season of birth—its relation to human abilities. John Wiley & Sons, New York, 1938.

(*2*) Shapiro, S., and Halpin, H.: Seasonal variation in the crude birth rate. Vital Statistics—Special Reports 23: 329–335 (1947).

(*3*) Lewit, S.: Seasonal variations in marriage licenses. Vital Statistics—Special Reports 33: 204–211 (1952).

(*4*) Kanellakis, A. P.: Seasonal swing in mortality in England and Wales. Brit J Prev Soc Med 12: 197–203 (1958).

(*5*) Kirk, D.: The influence of business cycles on marriage and birth rates. Demographic and economic change in developed countries. Princeton University Press, Princeton, N.J., 1960.

(*6*) Hotelling, H., and Hotelling, F.: Causes of birth fluctuations. J Amer Statist Assoc 26: 135–149 (1931).

(*7*) Hewitt, D.: A study of temporal variations in the risk of fetal malformation and death. Amer J Public Health 52: 1676–1688 (1962).

(*8*) Schuman, S., Anderson, C., and Oliver, J.: Epidemiology of successive heat waves in Michigan in 1962 and 1963. JAMA 189: 733–738 (1964).

(*9*) Collins, S.: Trend and age variation of mortality and morbidity from influenza and pneumonia. PHS Publication No. 544 (Public Health Monograph No. 48). U.S. Government Printing Office, Washington, D.C., 1957.

(*10*) Eickhoff, T., Sherman, I., and Serfling, R.: Observations on excess mortality associated with epidemic influenza. JAMA 176: 776–782 (1961).

(*11*) Croxton, F. E., and Cowden, D.: Applied general statistics. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955, pp. 240–393.

(*12*) Shiskin, J.: Electronic computers and business indicators. Occasional Paper No. 57. National Bureau of Economic Research, Inc., New York, 1957: (*a*) p. 235; (*b*) pp. 225–230.

(*13*) Shiskin, J., and Eisenpress, H.: Seasonal adjustments by electronic computer methods. J Amer Statist Assoc 52: 415–449 (1957).

(*14*) The BLS seasonal factor method (1964). Bureau of Labor Statistics, U.S. Department of Labor, 1964.

(*15*) Bongard, J.: Notes on the characteristics of seasonal variation in Europe. *In* Seasonal adjustment on electronic computers. Organization for Economic Cooperation and Development, New York, 1961, pp. 181–184.

(*16*) Waggoner, D., and Schachter, J.: Seasonal variation of births in the United States, 1948–54. Vital Statistics—Special Reports 47: 127–143 (1959).

(*17*) Serfling, R.: Methods for current statistical analysis of excess pneumonia-influenza deaths. Public Health Rep 78: 494–506 (1963).