Revision of Seasonally Adjusted Labor Force Series

* Robert J. McIntire

The purpose of seasonal adjustment is to eliminate from economic time series the influence of weather, holidays, the opening and closing of schools, and other such seasonal events in order to make it easier to observe and analyze the cyclical and other nonseasonal movements in the series. The seasonality which the adjustment process endeavors to eliminate is represented by seasonal factors. The seasonal factors used for current adjustment are estimates of how much the original unadjusted values can be expected to deviate from underlying trend-cycle levels due to annually recurring behavior as projected from average seasonal patterns in the recent past. Even though seasonality involves regularly recurring patterns, it does tend to change over time, creating a need for periodic reestimation of factors and revision of recently adjusted estimates. By including more recent data in the estimation process, the revision process can provide better estimates of how much the original, unadjusted estimates actually deviated from underlying trend-cycle levels during the recent period, thereby improving the historical seasonally adjusted data for that period. In addition, the new information is incorporated to produce the new projected factors to be used for current seasonal adjustment.

Therefore, at the end of each calendar year, the Bureau of Labor Statistics reestimates the seasonality of the unemployment, employment, and other labor force series derived from the Current Population Survey by including another full year of data in the estimation process. Based on this annual reestimation, BLS issues the projected factors for the first 6 months of the new year as well as revised estimates of historical seasonally adjusted data for the last 5 years. Each year's data are generally subject to five revision cycles before the values are considered final. The fifth and final revisions in the earliest of the 5 years are usually quite small, while the first-time revisions in the most recent year are generally much more substantial, although even these rarely alter the essential trends observed in the initial major estimates. This year's revisions incorporate data through December 1989 and provide revised estimates for January 1985 through December 1989 for all seasonally adjusted labor force series. In addition, seasonally adjusted data for a few persons-at-work series were revised back through January 1978 because of the first-time application of a special new extension of the seasonal adjustment procedure that provides a correction for the moving-holiday effects of Easter in such series. This extension is discussed in more detail later in this article.

Table 1 contains the new projected seasonal factors to be applied during the first 6 months of 1990 to the 12 component series used in the computation of the seasonally adjusted civilian labor force and unemployment rate. (See the section on aggregation procedures later in the article.) Projected factors for the last 6 months of 1990 will be estimated in early July, based on data through June 1990, and published in the July issue of this publication.

Effect of revisions

One of the criteria used to evaluate alternative methods of seasonal adjustment is how close initial estimates are to later revisions. Policymakers and analysts must make determinations based on current information, and so it is important that the initial estimates of current factors for the seasonal adjustment of major economic series produce estimates of level and change that are as close as possible to the improved estimates that will be made after more data have become available. Even though the revisions currently being released for the 1989 seasonally adjusted data are not final, the first revisions are, as mentioned above, usually the most substantial of the five that will be made and often indicate the direction any subsequent revisions will take. Therefore, it is appropriate to compare these first revisions with the initial estimates. Table 2 shows the overall and civilian unemployment rates for 1989 as first computed and as revised, as well as the change due to revision. Both of the rates, rounded to one decimal place as published, changed in only 4 or 5 months of the year, with the absolute effect of the changes never exceeding 0.1 percentage point in any of those months. Compared to the initial estimates, the revised rates suggest slightly smoother behavior for the rates during the year.

Adjustment methods and procedures

The official seasonal adjustment procedure for the labor force series is the X-11 ARIMA program, which was developed at Statistics Canada during the 1970's as an ex-

* Robert J. McIntire is an economist and Chief of the Data Development Staff in the Office of Employment and Unemployment Statistics, Bureau of Labor Statistics.

Table 1. Current seasonal adjustment factors for the 12 major civilian labor force components, January–June 1990

Procedure and series	January	February	March	April	May	June
Multiplicative Adjustment (Divide factor into original value)						
Agricultural employment:						
Men, 20 years and over	.699	.694	.937	1.002	1.039	1.068
Women, 20 years and over	.863	.900	.923	.943	1.049	1.095
Men, 16 to 19 years	.612	.617	.711	1.022	1.130	1.569
Women, 16 to 19 years	.651	.735	.962	.978	.973	1.486
Nonagricultural employment:						
Men, 20 years and over	.988	.989	.992	.995	1.001	1.008
Women, 20 years and over	.997	.998	1.001	1.001	1.000	.993
Unemployment:						
Men, 20 years and over	1.211	1.186	1.125	1.006	.939	.937
Women, 20 years and over	1.049	1.013	.969	.929	.969	.998
Additive Adjustment (Subtract factor from original value)						
Nonagricultural employment:						
Men, 16 to 19 years	-317	-344	-293	-264	-137	426
Women, 16 to 19 years	-241	-250	-263	-216	-143	303
Unemployment:						
Men, 16 to 19 years	16	12	-22	-87	-19	184
Women, 16 to 19 years	-44	-50	-64	-64	-9	266

Table 2. Seasonally adjusted unemployment rates in 1989 and change due to revision

Month	As first computed		As revised		Change due to revision	
	Overall	Civilian	Overall	Civilian	Overall	Civilian
January	5.4	5.4	5.3	5.4	-0.1	0
February	5.1	5.1	5.1	5.2	0	0.1
March	4.9	5.0	5.0	5.0	.1	0
April	5.2	5.3	5.2	5.3	0	0
May	5.1	5.2	5.1	5.2	0	0
June	5.2	5.3	5.2	5.3	0	0
July	5.2	5.2	5.2	5.3	0	.1
August	5.1	5.2	5.2	5.3	.1	.1
September	5.2	5.3	5.3	5.3	.1	0
October	5.2	5.3	5.2	5.3	0	0
November	5.3	5.4	5.3	5.3	0	-.1
December	[1]5.3	[1]5.4	5.3	5.3	0	-.1

[1] These rates reflect the use of seasonal factors projected for December 1989 and published in the July 1989 issue of *Employment and Earnings* and were subject to revision before regular publication of December data.

tension of and improvement to the widely used X-11 method developed at the U.S. Bureau of the Census in the 1960's.[1] The X-11 ARIMA method improves current estimates for most series by allowing recent observations, especially the last 6 months, to weigh more heavily in the estimates of current and recent seasonal factors than did the X-11 alone. The method provides this improvement through the use of ARIMA models to extend the data series by 12 months. The X-11 algorithm for seasonal adjustment is then applied to the extended series.

ARIMA models. ARIMA projections are based only on the past experience observed in a series itself. ARIMA models have proved to have good properties for short-term projection or extrapolation of a large class of time series, especially in a seasonal adjustment context, since the extrapolations tend to track intra-year movements quite well. The ARIMA models in the X-11 ARIMA program used to seasonally adjust the labor force series are of the Box-Jenkins type.[2] They can

[1] The primary documentation for the X-11 ARIMA procedure is in *The X-11 ARIMA Seasonal Adjustment Method*, by Estela Bee Dagum (Statistics Canada Catalogue No. 12-564 E, January 1983). (ARIMA is an acronym for Auto-Regressive Integrated Moving Average.) The X-11 method is described in *The X-11 Variant of the Census Method II Seasonal Adjustment Program*, by Julius Shiskin, Alan Young, and John Musgrave (Technical Paper No. 15, Bureau of the Census, 1967).

[2] For a more detailed discussion of ARIMA models, refer to previously cited Dagum (1983) and to G.E.P. Box and G.M. Jenkins, *Time Series Analysis, Forecasting and Control* (San Francisco, Holden Day, 1970); and C.W.J. Granger and P. Newbold, *Forecasting Economic Time Series* (New York, Academic Press, 1977).

generally be described with the notation:

(p,d,q)(P,D,Q) TRANSFORMATION,

Where:

(1) p is the number of regular (nonseasonal) autoregressive parameters
(2) d is the number of regular differences
(3) q is the number of regular moving average parameters
(4) P is the number of seasonal autoregressive tab parameters
(5) D is the number of seasonal differences
(6) Q is the number of seasonal moving average parameters
(7) TRANSFORMATION may be NONE, LOG, or POWER(n).

While the lettered elements within the parentheses of the model specifications can theoretically take on many values, in practice only small values are useful.

For each labor force series which has been extended based on an ARIMA model, the model has been specifically chosen as well suited to the particular series, based on a set of established criteria. The criteria essentially require a model to: (1) fit the series well, (2) have low average forecasting errors in the last 3 years prior to the projected year, and (3) produce residuals (the differences between the observed values and the values forecast by the model for the observed period) which follow a random pattern. Acceptable ARIMA models have been identified and were used for 166 of the 192 labor force series which were directly adjusted at the end of 1989, including all 12 major civilian labor force components, whose ARIMA models are shown in table 3. The models for two of those major components—agricultural employment for women 16 to 19 years and unemployment for women 20 years and over—are different from those used last year. The 26 remaining series for which acceptable models have not been identified were simply run through the X-11 part of the program without any ARIMA extrapolations.

X-11 procedures. The procedures used for adjusting the labor force series within the X-11 part of the process were the same as those followed last year. A 10-year time period, including data from January 1980 through December 1989, was used for the adjustment of all the labor force series except those being run through the new moving-holiday extension for the first time.

The X-11 method of seasonal adjustment contained in the X-11 ARIMA procedure assumes that the original series, including the 12 extrapolated observations if an ARIMA model has been applied, is either the product or the sum of three components—trend-cycle, seasonal, and irregular. The method uses either a ratio-to- or difference-from-moving-average approach to estimate the components, depending on whether the multiplicative or additive model is used. The seasonally adjusted series values are computed by dividing each month's original value by the corresponding seasonal factor if the multiplicative model is used, or by subtracting the factor if the additive model is used. Of the 12 major civilian labor force components, the 4 teenage unemployment and nonagricultural employment series were adjusted using the additive model, and the other 8 series with the multiplicative model. Of all the 192 directly adjusted series, 37 were adjusted with the additive model, primarily those involving teenage employment and unemployment, for which the seasonal component seems to be fairly independent of the trend-cycle.

Table 3. ARIMA models used in end-of-1989 seasonal adjustment for the 12 major civilian labor force components

Series	Model	Transformation
Agricultural employment:		
Men, 20 years and over	(1,0,0) (0,1,1)	LOG
Women, 20 years and over	(1,0,0) (0,2,2)	LOG
Men, 16 to 19 years	(0,1,2) (0,1,1)	NONE
Women, 16 to 19 years	(2,1,2) (0,1,1)	NONE
Nonagricultural employment:		
Men, 20 years and over	(0,1,2) (0,1,1)	LOG
Women, 20 years and over	(0,1,1) (0,1,1)	LOG
Men, 16 to 19 years	(2,1,0) (0,1,1)	NONE
Women, 16 to 19 years	(0,1,1) (0,1,1)	NONE
Unemployment:		
Men, 20 years and over	(2,1,2) (0,1,1)	NONE
Women, 20 years and over	(0,1,2) (0,1,1)	LOG
Men, 16 to 19 years	(0,1,2) (0,1,1)	NONE
Women, 16 to 19 years	(0,1,2) (0,1,1)	NONE

Moving-holiday adjustment. Three of the series directly adjusted with multiplicative models were seasonally adjusted using the new moving-holiday extension of X-11 ARIMA which has been developed at BLS. These three series, all relating to persons at work, had tested as having significant and well-defined effects in their April data related to the timing of Easter. The series are: (1 and 2) at work on part-time schedules for noneconomic reasons, usually work part time, all industries and nonagricultural industries, both of which are published monthly in seasonally adjusted form; and (3) at work on full-time schedules, nonagricultural industries, a seasonally adjusted series formerly published and still maintained.

The standard X-11 seasonal-factor moving averages in X-11 ARIMA assume that seasonal behavior in any given month is relatively consistent from year to year, even though it may evolve over a number of years. Holidays are one source of seasonality, and most affect economic data in a way sufficiently consistent from year to year to allow the standard X-11 assumption to work reasonably well. Easter, however, can occur anywhere from March 22 to April 25 and is the principal example of a potential source of the seasonal or calendar effects generally called "moving-holiday effects" in economic series. When these effects are present and not controlled for, they tend to confound seasonal component estimation for the affected month or months.

Most of the seasonally adjusted labor force series are not significantly affected by Easter, and previous work on moving-holiday effects has generally disregarded the

possibility of such effects occurring in such reference-period-based series (focusing instead on series which are accumulations over all days of each month, such as series on sales). The at-work series on voluntary part-time and full-time schedules are affected, however, because significant numbers of employed persons shift from being at work to being "with a job but not at work" in the observed data when the CPS reference week is the week immediately before or after Easter. This shift stems primarily from vacations and can show up only in the data for April, the only month in which the CPS reference week—the week containing the 12th of the month—can be the week before or after Easter. In years in which one of these intersections between Easter and the CPS reference week occurs, the number of persons at work in these categories in April tends to decrease substantially. In other years, when the CPS reference week is further removed from Easter, the number of persons at work in April tends to increase, reflecting the pattern for employment in general. The standard X-11 method averages these two very different sets of behavior together and produces April seasonal factors that usually do not adjust very well for either type of behavior. The new BLS moving-holiday extension provides a means of averaging the different types of behavior separately and allowing the differences to be reflected in the final seasonal factors.

The moving-holiday procedure works as follows for each series to which it is applied. First, an initial decomposition into trend-cycle, seasonal, and irregular components is done with X-11 ARIMA (without ARIMA extrapolation). The resulting initial seasonal component reflects the standard X-11 averaging process. The deviations from this average April seasonal component that are due to the timing of Easter are reflected in the April values of the resulting initial irregular component, which tend to be low for the years when the Easter/CPS-reference-week intersection occurred and high in the years when it didn't. The second stage of the procedure computes averages of those April irregular values (including a correction for any extremes) within sets of years grouped in ranges of Easter dates. The most appropriate ranges of Easter dates for any given series are selected based upon empirical analysis of the initial irregular component and upon expert knowledge of how the economic phenomenon as measured in the series is affected by the proximity or timing of Easter. The average of the April irregulars within each set becomes the estimate of the holiday effect (more precisely the holiday effect net of the average seasonal) in April for each of the years participating in that set. These holiday effects then become "holiday prior-adjustment factors," used to prior adjust the April observations for the third stage of the procedure, which is a second run of the series through X-11 ARIMA (this time with ARIMA extrapolation if a model is specified or requested). Finally, combined seasonal factors, including the projected factors, are computed as the product of the seasonal factors produced by the second X-11 ARIMA run, which essentially are the average seasonals reestimated with the net moving-holiday effect removed, and the holiday prior-adjustment factors derived in the first two stages. One benefit of this approach is that seasonal factors for months other than those with moving-holiday effects are only minimally affected.

Because this procedure involves creating subsets of the irregular values for the months affected by moving holidays, it is usually necessary to use more years for the estimation of the holiday effect in the first two stages than typically need to be used for the final seasonal adjustment in the third and final stages. The extra years help to ensure that a reasonable number of years are available for each subset. For the estimation of the holiday effect in the three at-work series, data back to 1967 were used for the two part-time series and data back to 1970 were used for the series on full-time schedules. Three subsets were used for estimating the moving-holiday effects for these series, based on the following ranges of Easter dates: (1) April 13-19, defining years in which the CPS reference week for April is the week immediately before Easter, the timing when this holiday effect is usually strongest; (2) April 6-12, defining years in which the CPS reference week for April is the week after (and containing) Easter, when the average effect is similar to but not as strong as the effect for the week before; and (3) all other dates of Easter, defining years in which the unadjusted data are not affected by Easter.

For this year's final seasonal adjustment of these three at-work series, data back to 1975 were used in the third stage of the procedure because of the need to revise the historically seasonally adjusted data back to 1978, a year in which the confounding effect of the moving holiday had been fairly substantial in the existing seasonally adjusted data. (The more usual practice will be to use the regular 10-year period of data for the final seasonal adjustment of these series and revise only the most recent 5 years.) In addition to revising the three directly adjusted series back to 1978, the aggregated seasonally adjusted series on the total at work in non-agricultural industries was recomputed back to 1978 to be consistent with the revised directly adjusted components.

Six-month updates. With respect to the seasonal adjustment schedule for the labor force series, the current official practice involves the running of all 192 directly adjusted series through X-11 ARIMA twice each year, after receipt of June and December data, with 6 months of projected factors drawn from each run and historical revisions drawn from the end-of-year run. This practice allows, among other things, the prior publication of seasonal factors, which historically has been regarded by the Bureau of Labor Statistics and other statistical agencies as an important way of ensuring the openness of their seasonal adjustment procedures, especially where very sensitive indicators such as the unemployment rate have been involved. In recent years, a number of research studies, including a 1987 paper on the labor force series,[3] have indicated that the alternative practice of con-

[3] G.R. Methe and R.J. McIntire, "An Evaluation of Concurrent Seasonal Adjustment for the Major Labor Force Series," in the 1987 *Proceedings of the Business and Economic Statistics Section*, American Statistical Association; copies of a more complete version of the paper can be obtained directly from the authors at the Bureau of Labor Statistics.

current adjustment, where the seasonal adjustment procedure is run with all available data each month and factors cannot be published ahead of time, generally produces initial seasonally adjusted estimates requiring smaller revisions than those produced by adjustment using projected factors. The BLS is continuing to evaluate concurrent adjustment for the labor force series, and concurrently adjusted alternative estimates for major labor force series are available to interested users upon request.

Aggregation procedures

BLS maintains and publishes several hundred seasonally adjusted labor force series in addition to the 192 directly adjusted series discussed above. These additional series are produced by arithmetically combining or aggregating the directly adjusted series with each other or, in some cases, with series on population or resident Armed Forces levels, which are not seasonally adjusted because they are not considered to have any significant seasonal variation. For example, the seasonally adjusted levels of total unemployment, civilian employment, and civilian labor force, and the seasonally adjusted unemployment rate for all civilian workers, are all produced by aggregation of some or all of the seasonally adjusted results for the 12 major civilian labor force components. The seasonally adjusted level of total unemployment is the sum of the seasonally adjusted levels of unemployment for the four age-sex groups—men and women 16 to 19, and men and women 20 years and over. Seasonally adjusted civilian employment is the sum of the seasonally adjusted levels of employment for the eight employment components—the same four age-sex groups as noted above employed in nonagricultural and agricultural industries. The seasonally adjusted civilian labor force is the sum of all 12 components. The seasonally adjusted civilian unemployment rate is calculated by taking the total seasonally adjusted unemployment level as a percent of the total seasonally adjusted civilian labor force. For the overall labor force, the resident Armed Forces level is added to the seasonally adjusted civilian labor force, and the seasonally adjusted overall unemployment rate is, of course, calculated by taking total seasonally adjusted unemployment as a percent of that labor force figure.

The principal reason for producing many of the major seasonally adjusted estimates for the labor force by aggregation rather than by direct adjustment is that this approach ensures that the major seasonally adjusted totals will be arithmetically consistent with at least one major set of components. If the totals were directly adjusted along with the components, such consistency would not, in all likelihood, occur, since the X-11 is not a sum-preserving procedure; that is, the sum of the result for two or more directly adjusted series will not generally be the same as the result of directly adjusting the sum of the unadjusted versions of the same series. Another factor is that it would generally be inappropriate to apply seasonal factors computed for an aggregate series to the components of the aggregate. The various labor force components tend to have significantly different patterns of seasonal variation; for example, teenage unemployment tends to peak in June, while unemployment of adult men tends to peak in the winter months of January and February. In order to properly estimate these varying seasonal patterns, it is necessary to directly adjust the components. Of course, one of the implications of producing seasonally adjusted estimates for many major series by aggregation is that exact factors cannot be projected for those series. However, implicit seasonal adjustment factors can be calculated after the fact by taking the ratio of the unadjusted aggregate to the seasonally adjusted aggregate, or, for additive implicit factors, the difference between those two aggregates.

Availability of revised series

This issue of *Employment and Earnings* contains revised data for the last 13 months and quarters for many seasonally adjusted labor force series. Next month's issue will provide the 1985-89 revisions for a few hundred of the seasonally adjusted monthly labor force series most in demand. These revisions replace the data published in the February 1989 issue for 1985-88 and the seasonally adjusted estimates for 1989 published during the past year. Other than for those few persons-at-work series revised back to 1978, seasonally adjusted data for 1984 and earlier years were not revised.

Additional data for any of the several hundred seasonally adjusted labor force series, as well as the January-June 1990 factors for any of the directly adjusted series beyond the 12 major components, can be obtained from BLS upon request. Requests for data or inquiries concerning seasonal adjustment methodology or the availability of machine-readable files of labor force data should be addressed to the Data Development Staff, Office of Employment and Unemployment Statistics, Bureau of Labor Statistics, Washington, DC 20212.