A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Method

by
J. Lothian and M. Morry

Seasonal Adjustment and Time Series
Analysis Staff
Statistics Canada.

78-10-005 E

October, 1978.

A SET OF QUALITY CONTROL STATISTICS FOR THE X-11-ARIMA SEASONAL ADJUSTMENT METHOD

I. INTRODUCTION

The X-11 Variant of the Census Method II Seasonal Adjustment Method (J. Shiskin et al 1967) contained a summary measures table denoted by F.2. The purpose of the F.2 Table was to give a set of statistics pertaining to the estimated trend-cycle, seasonal and irregular components. These statistics gave information about the average percent changes in each component with and without regard to sign over different spans, relative contribution of components to percent changes in the original series, average duration of run etc.

The X-11 method was modified in 1973 at Statistics Canada to include two statistics, Q_1 and Q_2 which provided an indication of the amount and nature of the irregular and the seasonal component respectively. A description of these statistics and their basic assumptions are discussed by Huot and de Fontenay (1973).

Considerable research has been carried out since the first set of guidelines was developed and it has now been reduced to only one Q statistic which results from the combination of eleven other measures. Most of them are obtained from the summary measures in Table F.2. The values of the eleven statistics range from 0 to 3, low values indicating good quality with 1 being the cut-off point for the test. A weighted sum of the eleven statistics makes up the final Q value. If Q exceeds 1, the series fails the guidelines, i.e., the quality of the seasonal adjustment is considered unacceptable. The sections to follow give a detailed description of each of the statistics.

II. THE FIRST SEVEN QUALITY CONTROL STATISTICS, M1 TO M7

(1) The relative contribution of the irregulars over three months span (M1).

In the Summary Measures section of the X-11-ARIMA program, Table F2.B contains the relative contributions to the variance of the percent change (difference) in the components of the original series. In table F2.B under D13, D12 and D10, the contribution of the irregular, trend-cycle and seasonal components can be found over the spans 1 to 12 (or 4 for quarterly series).

For example, the value for span 1 under the heading D13 denoted here by $R_{I(1)}$ is calculated as follows:

$$\bar{I}(1)^2 = \sum_{t=2}^{N} (I_t - I_{t-1})^2/(N-1), \tag{2.1}$$

where I_t are the final irregulars from Table D13 and N is the number of points in the series, and

$$\bar{O}'(1)^2 = \bar{I}(1)^2 + \overline{TC}(1)^2 + \bar{S}(1)^2 \tag{2.2}$$

where $\overline{TC}(1)^2$ and $\bar{S}(1)^2$ are calculated from Tables D12 and D10 respectively, according to the formula (2.1). The estimated relative contribution of the irregular to the variance of the percent change in the original series over span 1, equals

$$R_{I(1)} = \frac{\bar{I}(1)^2}{\bar{O}'(1)^2} \times 100\% \qquad (2.3)$$

From the point of view of seasonal adjustment, it is important to know the proportion of the irregular contribution relative to the seasonal contribution. If the irregular variation is too high when compared to the variation in the seasonal component, the two components cannot be separated successfully.

Applying differencing (over span 1, span 2, etc.) has the effect of removing a linear trend from the original series in an attempt to make it stationary (it is necessary to have a stationary series otherwise the variance is not defined). Unfortunately, differencing effects the variance of the other components as well.

In order to find out how much of the cycle, seasonal and irregular is removed by lag one, lag two, etc. differencing, the transfer functions of the differencing operators were studied. The following assumptions were made when examining the effect of the transfer functions:

(a) The irregular component I_t's are independent identically distributed random variables, i.e. their contribution to the variance is constant at all frequencies.

(b) The seasonal component shows typical behaviour in the distribution of power over the fundamental and harmonic frequencies.

(c) The cycle is distributed evenly over the very low frequencies.

Using these assumptions, the following information was extracted from the transfer functions and tabulated in Table I.

TABLE I. Percentage of the power (variance) left after applying differencing

Over span	T	C	S	I
1	0	6	24	134
2	0	18	64	122
3	0	36	112	112
4	0	60	147	116

Renormalizing the above table into a form where the relative power of I equals 100%, we obtain Table II.

TABLE II. Percentage of power (variance) left relative to I equals 100, after differencing

Over span	T	C	S	I
1	0	4.5	17.9	100.0
2	0	14.8	52.5	100.0
3	0	32.1	100.0	100.0
4	0	51.7	126.7	100.0

Since the main concern is to get a clear idea of the relative variation of the seasonal versus that of the irregular component, obviously a first difference removes too much of the seasonal relative to the irregular component. A lag 3 difference, however, appears optimal because it preserves the original proportions between the seasonal and the irregular. It has the minor disadvantage of not removing the cycle completely. Still, differencing over span 3 provides the best measure for comparing the contribution of the irregular against that of the seasonal component. Three months span corresponds to one quarter, thus in quarterly series testing is carried out on lag 1 differences.

The maximum acceptable contribution of the irregular to the total variance was set at 10% in the lag 3 difference. Thus if

$$R_{I(3)} > 10\% \tag{2.4}$$

the series fails the test statistics M1. Renormalizing $R_{I(3)}$ yields

$$M1 = \frac{R_{I(3)}}{10} \tag{2.5}$$

Thus if M1 is greater than 1 the contribution of the irregular to the variance is considered too high.

II.2) The relative contribution of the irregular component to the variance of the stationary portion of the series (M2).

This measure is similar to M1. The only difference is in the trend remover used to make the series stationary. Instead of lag three differencing, a line is fitted to the trend-cycle values in Table D12 to obtain a trend estimate (or an exponential growth is fitted and all the components are logarithmically transformed if the series is multiplicative). This trend estimate is removed from the original series to obtain a stationary raw series B1'. Table D12 is transformed as well by removing the same trend from it to get a new Table D12'. The relative contribution of the components appearing in Table F2.F are calculated as follows:

$$\text{contribution of I} = \frac{\text{variance of table D13}}{\text{variance of table B1'}} \tag{2.6}$$

$$\text{contribution of C} = \frac{\text{variance of table D12'}}{\text{variance of table B1'}} \tag{2.7}$$

$$\text{contribution of S} = \frac{\text{variance of table D10}}{\text{variance of table B1'}} \tag{2.8}$$

If the contribution of I is greater than 10%, the series fails the M2 test, where

$$M2 = 100 \times \frac{\text{contribution of I}}{10} \tag{2.9}$$

Thus if M2 exceeds 1 the variation of the irregular component contributes too much to the total variation of the series.

II.3) The amount of month-to-month change in the irregular as compared to the amount of month-to-month change in the trend-cycle (M3).

The purpose of a seasonal adjustment procedure is to extract the seasonal component from the raw data in order to estimate a seasonally adjusted series. Because of the iterative nature of the X-11 program it is important that in the steps leading up to the final seasonal adjustment, not only must the seasonal be well identified, but the trend-cycle and irregular component be properly estimated as well. If the month-to-month movement of the irregulars is dominant in the CI series, it is difficult to separate these two components and the overall quality of the seasonal adjustment suffers.

The statistic measuring this relationship between the irregular and the trend-cycle is the $\bar{I}/\bar{C}$ ratio where $\bar{I}$ and $\bar{C}$ are the mean absolute change from tables D13 and D12 respectively. This $\bar{I}/\bar{C}$ ratio can be found at the top of Table D12, and also in Table F2.H. If it exceeds 3, the amount of irregular movement is considered too high. The corresponding test statistic is the following:

$$M3 = (\bar{I}/\bar{C} - 1)/2. \tag{2.10}$$

The formula for quarterly series is:

$$M3 = (\bar{I}/\bar{C} - .33)/.67 \tag{2.11}$$

If M3 exceeds 1, the series fails this test.

II.4) The amount of autocorrelation in the irregular as described by the average duration of run (M4).

One of the basic assumptions of the statistical F-tests in the X-11 method is that the irregular component is a purely random process with constant variance and zero covariance when the relationship among the trend-cycle, seasonal and irregular is additive (multiplicative).

The program uses the Average Duration of Run statistic (ADR) to test for the randomness in the final estimated residuals obtained from Table D13 and prints it in Table F2.D under I. This non-parametric test, developed by W.A. Wallis and G.H. Moore (1941), is constructed on the basis of the number of turning points (a turning point occurs in a time series when the sign of the month-to-month change reverses). It is designed to test the randomness of the residuals against the alternative hypothesis that the errors I_t follow a first order autoregressive process of the form $I_t = \rho I_{t-1} + e_t$, where ρ is the autocorrelation coefficient and e_t is a purely random process.

Given a purely random process of infinite length, the ADR statistic would equal 1.50. For a series of 120 observations, the 99% confidence interval for the ADR extends from 1.30 to 1.75. Values greater than 1.75 indicate positive autocorrelation and those smaller than 1.30, negative autocorrelation of the residuals. The test statistic M4 is based on the normal approximation formula given by Bradley (1968).

$$M4 = \frac{\left| \frac{N-1}{ADR} - \frac{2(N-1)}{3} \right|}{\left(\frac{16N-29}{90} \right)^{\frac{1}{2}}} \times \frac{1}{2.58} \qquad (2.12)$$

where the value 2.58 is the 1% limit value of the standard normal distribution in two-sided tests. If M4 is greater than 1, there is significant autocorrelation present in the residuals and the series fails this test.

II.5) The months (quarters) for cyclical dominance statistic (M5).

This statistic measures the number of months (quarters) it takes the average absolute change in the trend-cycle to dominate that in the irregular. This value is printed in Table F2.E. This measure is similar to M3, namely it examines the relative size of the changes in the irregular and trend-cycle components. The $\bar{I}(k)/\bar{C}(k)$ ratios are computed for spans k equal to 1 to 12 (or 1 to 4 for quarterly series), and the MCD is derided from them:

$$MCD = k \quad \text{if} \quad \bar{I}(k)/\bar{C}(k) \leq 1 \quad \text{and} \quad \bar{I}(k-1)/\bar{C}(k-1) > 1 \tag{2.13}$$

For example, given the following $\bar{I}/\bar{C}$ ratios for spans 1, 2, 3, 4:

$$\bar{I}(1)/\bar{C}(1) = 1.82$$
$$\bar{I}(2)/\bar{C}(2) = 1.10$$
$$\bar{I}(3)/\bar{C}(3) = 0.81$$
$$\bar{I}(4)/\bar{C}(4) = 0.72$$

then the MCD equals 3, indicating that it takes 3 months on average for the absolute change in the trend-cycle to become higher than that of the irregular component.

The MCD statistic takes integer values only. Therefore, it is not capable of distinguishing between an $\bar{I}/\bar{C}$ ratio that just fell below 1 after, say, 3 months span and one that exceeded 1 by only a minimal amount after 2 months span and became much less than 1 after a 3 months span. To remedy this problem, a new statistic MCD' was calculated by linearly interpolating the $\bar{I}/\bar{C}$ ratio to find its intersection with 1. In the example quoted above, the MCD' value is the following:

$$\text{MCD}' = 2 + \frac{1.10 - 1.00}{1.10 - 0.81} = 2.34 \tag{2.14}$$

An MCD statistic of 6 or over has been traditionally considered to be unacceptable. Thus the final M5 statistic takes the form:

$$\text{M5} = \frac{\text{MCD}' - 0.5}{5.0} \tag{2.15}$$

The quarterly equivalent of this test is as follows:

$$\text{M5} = \frac{\text{QCD}' - 0.17}{1.67} \tag{2.16}$$

M5 values greater than 1 fail the test for months (quarters) for cyclical dominance.

II.6) The amount of year-to-year change in the irregular as compared to the amount of year-to-year change in the seasonal (M6).

As mentioned earlier, it is very important from the point of view of seasonal adjustment that the seasonal factors are properly identified. One of the steps in the X-11 seasonal adjustment is the application of the 7-term moving average (3 x 5) weights to the SI ratios (differences) in order to separate the irregular from the seasonal component. Experience has shown that when the year-to-year change in the irregular is too small, compared to the year-to-year change in the seasonal factors, as described by a low $\bar{I}/\bar{S}$ ratio, the (3 x 5) moving average is not flexible enough to follow the seasonal movement. On the other hand, when the $\bar{I}/\bar{S}$ ratio is too high, the (3 x 5) seasonal filter proves too flexible and the resulting seasonal factors are contaminated with some of the irregular movement.

Studies with 421 series presently seasonally adjusted by X-11 at Statistics Canada indicated (see Lothian (1978)) that when the $\bar{I}/\bar{S}$ ratio fell between 1.5 and 6.5, the (3 x 5) moving average worked relatively well. Beyond that range, the use of a shorter seasonal filter (for too low $\bar{I}/\bar{S}$ values) or a longer moving average (for too high ratios) would have been necessary to separate the two components correctly. Incidentally, of the 421 series 2% had $\bar{I}/\bar{S}$ values less than 1.5 and 2% had values exceeding 6.5. The M6 measure is based on the cut-off points 1.5 and 6.5 and is formulated the following way:

$$M6 = \left| \frac{\bar{I}/\bar{S} - 4.0}{2.5} \right| \tag{2.17}$$

If it exceeds 1.0, the statistics fails, but the problem may be remedied by adjusting the series with a (3 x 1) moving average if the $\bar{I}/\bar{S}$ ratio as shown in Table F2.H is less than 1.5 or using the stable seasonality option if this ratio is greater than 6.5.

II.7) The amount of stable seasonality present relative to the amount of moving seasonality (M7).

The 1967 version of the X-11 program contained a one-way-analysis F-test applied to the final SI ratios in Table D10 to measure the amount of stable seasonality present. At Statistics Canada, a companion F-test was developed by J. Higginson (1975) to signal if there is moving seasonality in the series.

These two F-values were combined into one statistic denoted by T that was designed to indicate whether the seasonality present in the series is 'identifiable' by X-11 or not. Here the seasonality is called identifiable if the absolute error (or distortion) introduced in the final seasonal factor estimates is not too high. It was found that this distortion depended on both the F_S (F-value from the stable seasonality test) and F_M values (F-value from the moving seasonality test). Low F_S values suggested high distortion, while high F_M-values indicated further distortion was introduced due to movement. The cut-off point was based on 10-year monthly series and it corresponds to a combination of F_S and F_M values that indicate 50% distortion in the seasonal factor estimate. Thus the test statistic took the following form:

$$M7 = T = \sqrt{\frac{1}{2}\left(\frac{7}{F_S} + \frac{3F_M}{F_S}\right)} \qquad (2.18)$$

For a detailed description on how the test was derived, the reader is referred to Lothian and Morry (1978). M7 values exceeding 1 indicate that the seasonality in the series is not identifiable.

III. THE LAST FOUR QUALITY CONTROL STATISTICS DESCRIBING THE YEAR-TO-YEAR MOVEMENT IN THE SEASONAL COMPONENT, M8 TO M11

The seasonal filters of X-11 work well only on constant seasonals in the first and last three years of the series, while in the middle years they can reproduce a line or a constant. Thus only a constant seasonal component can be optimally estimated for the whole length of the series. If the original seasonals contain year-to-year movement, the seasonal factor estimates will have considerable error.

We distinguish between two types of movement; one that exhibits quasi random fluctuations and the other where changes appear in the same direction throughout the years. The size of the first type of movement can be measured from the average absolute year-to-year change in the seasonal factors, while the simple arithmetic mean of the changes gives an indication of the size of systematic (linear) movement. Random fluctuations are measured by statistics M8 and M10. Statistics M9 and M11 describe the size of linear movement. M8 and M9 are calculated using all the data in Table D10.

Since users are mostly interested in the quality of seasonal adjustment in the recent years, statistics M10 and M11 were introduced to describe the seasonal movement at the end of the series. It is especially important to know if there is significant linear movement in the seasonal factors of the last years because the seasonal factor estimates will then be considerably distorted by the end-weights of the seasonal filters. It is this same distortion that prevents the use of the last three years' seasonal factors to measure the amount of seasonal movement. Instead, the changes of the three years before the last three years are examined in the hope that the seasonal movement remains unaltered in the end years.

Obviously an average absolute change of .5 in multiplicative seasonal factors ranging from 98.0 to 102.0 within a year does not have the same significance as an average change of .5 coming from additive seasonal differences in the range −165 to +200. Therefore, it is important to normalize the values in Table D10 before proceeding to calculate the statistics. Thus the measures M8 to M11 were based on the normalized seasonal factors:

$$S'_t = \frac{S_t - \bar{S}}{\text{standard deviation of } S_t} \tag{3.1}$$

III.1) The size of the fluctuations in the seasonal component throughout the whole series (M8).

As mentioned before, the fluctuations are measured by the average absolute change.

$$|\overline{\Delta S'}| = \frac{1}{J(N-1)} \sum_{j=1}^{J} \sum_{i=2}^{N} |S'_{Ji+j} - S'_{J(i-1)+j}| \tag{3.2}$$

where N is the number of years and J equals 4 or 12 (for quarterly or monthly series). The maximum acceptable change was set at 10%. Thus the M8 measure took the following form:

$$M8 = 100 \times |\overline{\Delta S'}| \times \frac{1}{10} \tag{3.3}$$

III.2) The average linear movement in the seasonal component throughout the whole series (M9).

Averaging the year-to-year changes for each month, measures the amount of systematic movement. If there are only random fluctuations from year-to-year this average will be very close to zero. If most of the changes are in the same direction per month, the average absolute change will be very close to the average arithmetic change (3.2).

Using the formula:

$$\sum_{i=1}^{N-1} \Delta S'_{Ji+j} = S'_{J(N-1)+j} - S'_{j} \qquad (3.4)$$

and setting the acceptance limit at 10%, we obtain:

$$M9 = 100 \times \frac{\sum_{j=1}^{J} \left| S'_{J(N-1)+j} - S'_{j} \right|}{J(N-1)} \times \frac{1}{10} \qquad (3.5)$$

III.3) The size of the seasonal component fluctuations in the recent years (M10).

This statistic is equivalent to M8 except that only year N–2, N–3, N–4, and N–5 are involved in the calculations. Thus formula (3.2) becomes:

$$\left| \overline{\Delta S'} \right|_R = \frac{1}{3J} \sum_{j=1}^{J} \sum_{i=N-4}^{N-2} \left| S'_{Ji+j} - S'_{J(i-1)+j} \right| \qquad (3.6)$$

The final statistic M10 is of the form:

$$M10 = 100 \times \left| \overline{\Delta S'} \right|_R \times \frac{1}{10} \; . \qquad (3.7)$$

III.4) The average linear movement in the seasonal component in recent years (M11).

This measure corresponds to M9 using data from year N–2, N–3, N–4 and N–5. The formula for calculating M11 is the following:

$$M11 = 100 \times \frac{\sum_{j=1}^{J} \left| S'_{J(N-2)+j} - S'_{J(N-5)+j} \right|}{3J} \times \frac{1}{10} \qquad (3.8)$$

When M11 exceeds 1, there is strong indication that the seasonal factors for recent years are highly distorted due to the flattening effect of the end weights on linear movements.

IV. GENERAL COMMENTS ON THE OVERALL QUALITY OF THE ADJUSTMENT

The eleven statistics each examine a different facet of the adjustment and no one statistic can judge the overall quality of the adjustment. Also, each statistic has been developed for an average series and thus might break down for an unusual series.

It is possible that the series fails the M1 or M2 statistic and the adjustment does not necessarily suffer. These two statistics measure the irregular variation in proportion to the seasonal variation. The average series adjusted has a cycle which contributes about 5 to 10% to the stationary portion of the variance. The threshold level for the M1 and M2 statistics is based on this assumption. If the series contains no cycle, the irregular can contribute 13 to 14% to the total variation (resulting in M1 and M2 values exceeding 1) and still be acceptable. Similarly, if the cycle contributes more than 10%, the threshold level should be lowered.

If a series has a flat (i.e. almost constant) trend-cycle, it is possible to have an $\bar{I}/\bar{C}$ ratio exceeding 3 and thus failing the M3 test, without jeopardizing the quality of the adjustment. Actually, the X-11 program compensates for the lack of trend by applying a 23-term Henderson moving average to estimate the trend-cycle. However, if the user's main objective is business cycle analysis a high M3 value signals a serious problem. It indicates that the final seasonally adjusted series contains a very high

[redacted]

proportion of irregular movement that will prevent users from properly identifying the trend-cycle component.

Finding significant autocorrelation in the final irregulars as indicated by an M4 value greater than 1, can signal, for example, that the user should have applied trading-day regression and thus the adjustment is not valid. At the same time it is possible that the original irregulars were autocorrelated due to the sampling design. This will not affect the X-11 seasonal adjustment that is based on recognizing characteristic seasonal and trend-cycle behaviour in a series and obtains the irregulars as the residuals of the procedure. Thus the correct seasonal factor can still be well identified. It was found that the measure M4 moved rather independently from the other measures and quite often it was the only statistic that failed or one of the very few that did not fail. Consequently, it was not as related to the quality of seasonal adjustment as the others and was assigned a minimum weight.

In the case of M5, what was said about M3 applies again. It is possible that the irregulars are too high but it is also conceivable that the series contains an almost constant trend-cycle which does not prevent X-11 from isolate out the right seasonal movement.

As pointed out before, M6 is the only statistic where failure can be corrected. The user is advised to rerun X-11 and apply the appropriate seasonal moving average to the SI series in order to improve the quality of seasonal adjustment.

Any series that fails the statistic M7 has either no seasonality or the seasonal estimates are so distorted that the seasonal component is not identifiable as indicated in the message after Table D8. This measure is the most important one in the set of quality control statistics and is, therefore, assigned the highest weight. If the series fails M7, the user is strongly advised not to adjust the series. However, there are exceptions even here. It is possible that due to using an additive option in adjusting a series where the components are related multiplicatively and that has a rapidly growing trend, the F_M-value from the test for moving seasonality is very high. This can result in an M7 value exceeding 1. If the adjustment is rerun multiplicatively, the F_M-value will be reduced significantly and M7 passes the guidelines.

Failing statistics M8 and M10 might not be crucial if M9 and M11 pass the guidelines (their value being less than 1) and the user is only worried about bias in the current seasonal estimates. Similarly, if M9 and M11 exceed 1, but the user is only interested in the historical seasonal factors, those estimates can still be accurate because the central weights of the seasonal moving average can follow any linear movement. However, if one is interested in the current seasonal factors, high M9 and M11 values indicate the presence of significant distortion in the estimates.

From the above discussion, it is obvious no one statistic can assess the quality of the adjustment. If all eleven fail, the adjustment is unacceptable. But what if some fail and others do not? A quality control statistic was developed that is a weighted sum of the eleven statistics. Each statistic was assigned a weight according to its relative importance to the overall quality of the adjustment. One statistic cannot cause the adjustment to be rejected, rather it must be a composite effect of all the statistics. The weights assigned to the eleven statistics appear in Table III.

TABLE III. The Standard Eleven M Weights

Statistics(Mi)	Weight (w_i)
M1	13
M2	13
M3	10
M4	5
M5	11
M6	10
M7	16
M8	7
M9	7
M10	4
M11	4

The eleven statistics can sometimes take values less than zero or greater than three. If this happens the statistic is set to be zero or three respectively. Thus the quality control statistic Q is defined as:

$$Q = \frac{\sum_{i=1}^{11} w_i Mi}{\sum_{i=1}^{11} w_i} \tag{4.1}$$

If the user selects a seasonal moving average different from a (3 x 5) for estimating the seasonal factors, the statistic M6 is not relevant. Thus under these conditions:

$$w_6 = 0. \tag{4.2}$$

If the series is less than 6 years long, or the stable seasonal option is chosen, the statistics M8, M9, M10 and M11 cannot be calculated and the weights are redefined as displayed in Table IV

TABLE IV. Modified M Weights

Statistics(Mi)	Weight (w_i)
M1	17
M2	17
M3	10
M4	5
M5	11
M6	10
M7	30
M8	0
M9	0
M10	0
M11	0

This combination of the eleven statistics were very successful in assessing the quality of adjustment of 421 series tested by the authors. These series were adjusted at Statistics Canada with the X-11-ARIMA program and varied in length from 5 to 30 years. The average value for the eleven M statistics and the Q statistics for the 421 series are given in Table V.

TABLE V. Average Values for the Statistics

Statiatics	Monthly Series	Quarterly Series	All Series
M1	0.719	0.556	0.680
M2	0.605	0.332	0.540
M3	0.485	0.304	0.442
M4	0.424	0.662	0.481
M5	0.593	0.465	0.563
M6	0.380	0.516	0.412
M7	0.403	0.362	0.393
M8	0.640	0.562	0.619
M9	0.394	0.393	0.393
M10	0.724	0.685	0.714
M11	0.684	0.669	0.680
Q	0.529	0.461	0.513

If the Q statistic is greater than 1, the adjustment of the series is declared to be unacceptable. The adjustment is also rejected if the test for identifiable seasonality fails. For quarterly series 11.0% of the series failed the Q statistic and an additional 1% failed the test for identifiable seasonality. For monthly series, 8.4% had Q-values higher than 1 and an additional 3.7% were rejected because they did not pass the test for identifiable seasonality. Overall 12.1% of the 421 seasonal adjustments were rejected. The 51 series that failed were examined in detail by the authors and for all of them, the adjustment was deemed to be unacceptable. The quality control statistics presented here will enable users with a large number of series to quickly assess the quality of all their adjustments as well as enable people with little knowledge of seasonal adjustment to make judgements on the acceptability of the results. The Q statistic provides a general assessment of the quality of the adjustment, but the users should beware of attaching significance to small changes in the statistic. This especially holds for aggregate adjustments as shown in Appendix A.

At the back of each printout produced by the X-11-ARIMA program, appears a summary of the Q statistics for all the series adjusted in that run. Thus, the quality of large numbers of series can be quickly judged. Immediately after the Q statistics are copies of the F2 and F3 tables of all the series run. If any series has produced an unacceptable adjustment, the user can turn to the F2 and F3 tables for that series and further identify the problem. Following this procedure, hundreds of series can be assessed in less than an hour.

APPENDIX A

QUALITY ASSESSMENT OF AGGREGATE SERIES

The new X-11-ARIMA program can automatically produce direct and indirect aggregate seasonal adjustments of several component series. The preceding quality control statistics are produced for both the direct and indirect adjustment. The Q statistics for the two methods can be used to assess the acceptability of the adjustment for both methods. Unfortunately, the Q statistics cannot be used to judge which of the two methods gives a superior adjustment. The Q's for the two types of adjustments are usually very close to each other and small differences in the Q's cannot be interpreted as being significant. The M8 and M10 statistics for the indirect method will generally be greater than for the direct. This will tend to make the Q for the indirect method greater than for direct. This creates a bias in the Q statistic against indirect adjustments.

Additional summary statistics are produced if an aggregate adjustment is requested. These statistics are printed on the same page as the summary of all the Q statistics for the series adjusted. This summary appears right after the printout of the last series.

The comparison test for the direct versus the indirect seasonal adjustment method is based on a paper by Lothian and Morry (1977). The statistic tests the degree of smoothness of the seasonally adjusted series. The standard deviation of the month-to-month (or quarter-to-quarter) changes in the two seasonally adjusted series is computed for the whole series and for the last three years of the series. The standard deviation of the direct differences is then subtracted from the indirect for both the whole series and the last three years. If the resulting differences are positive, the indirect method gives a smoother adjustment than the direct method. If they are negative, the direct method results in a smoother seasonally adjusted series. It is also possible that the results for the last three years disagree with those of the full series. The differences are normalized by dividing by the average value of the seasonally adjusted series and multiplying by 100 to get the percentage difference between the direct and indirect methods.

REFERENCES

Bradley, James V. (1968): Distribution-Free Statistical Tests, Prentice Hall, Englewood Cliffs, N.J.

Higginson, John (1975): "An F test for the Presence of Moving Seasonality when Using Census Method II-X-11 Variant" Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Huot, Guy and de Fontenay, Allain (1973): "General Seasonal Adjustment Guidelines, Statistics Canada version of the X-11 Program", Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Lothian, John (1978): "The Identification and Treatment of Moving Seasonality in the X-11-ARIMA Seasonal Adjustment Program", Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Lothian, John and Morry, Marietta (1978): "A Test for Identifiable Seasonality When Using the X-11-ARIMA Program", Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Lothian, John and Morry, Marietta (1977): "The Problem of Aggregation; Direct or Indirect" Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Shiskin, J., Young, A.H. and Musgrave, J.C. (1967): "The X-11 Variant of Census Method II, Seasonal Adjustment", Technical Paper No. 15, Bureau of the Census, U.S. Dept. of Commerce

Wallis, W.A. and Moore, G.H. (1941): "A Significance Test for Time Series", National Bureau of Economic Research, Technical Paper No. 1.