AN EMPIRICAL EVALUATION OF THE PERFORMANCE OF TRAMO/SEATS ON SIMULATED SERIES

Catherine C. Hood, James D. Ashley, and David F. Findley, US Census Bureau
Catherine C. Hood, US Census Bureau, Rm 3110-4, ESMPD, Washington, DC 20233

Key Words: seasonal adjustment, regression models with ARIMA noise, outliers, revisions

TRAMO (Time series Regression with ARIMA noise, Missing observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series) are linked programs developed by Victor Gomez and Agustin Maravall to seasonally adjust time series using ARIMA model-based signal extraction techniques. We will evaluate the performance of TRAMO/SEATS on some simulated economic time series, including series with a large irregular component, series with complex trends, and short series.

The Bureau is moving to a new classification system for establishment based surveys. The former Standard Industrial Codes (SIC) are being replaced by the North American Industrial Classification System (NAICS). When this happens, there is a possibility that some series at the Census Bureau will need to be adjusted with only a few years of data. So we will look at adjusting our simulated series with only four years of data.

We evaluated the performance of seasonal adjustments from TRAMO and SEATS as compared to adjustments from X-12-ARIMA. X-12-ARIMA is the Census Bureau's latest program in the X-11 line of seasonal adjustment programs. We looked at three different versions of SEATS, including a version of X-12-ARIMA that has access to the SEATS algorithm hereafter called X-12-ARIMA-SEATS. X-12-ARIMA-SEATS allows us to compute similar diagnostics for both programs so we can compare adjustments between model-based adjustment and the X-12 method.

1. BACKGROUND

1.1 TRAMO/SEATS and X-12-ARIMA

TRAMO/SEATS and X-12-ARIMA are based on two different methods for seasonal adjustment. SEATS uses signal extraction with filters derived from an ARIMA-type time series model that describes the behavior of the series. This method is based on work by Burman (1980) and Hillmer and Tiao (1982), among others. See also Gomez and Maravall (1997a).

X-12-ARIMA uses signal-to-noise ratios to choose between a fixed set of moving-average filters, often called X-11-type filters. X-12-ARIMA is based on the well-known X-11 program (Shiskin, 1967) and Statistics Canada's X-11-ARIMA and X-11-ARIMA/88 (Dagum, 1988). Major improvements in X-12-ARIMA over X-11-ARIMA/88 are discussed in Findley, Monsell, Bell, Otto, and Chen (1998). See also U.S. Census Bureau (2000).

TRAMO and X-12-ARIMA are designed to preadjust a series before the seasonal adjustment by removing some deterministic effects such as trading day, moving holidays (Easter for example), and outliers. TRAMO and X-12-ARIMA also contain similar model diagnostics.

SEATS and X-12-ARIMA are similar in that both programs use the preadjusted series to decompose the series into seasonal, trend-cycle, and irregular components. SEATS uses the model for the decomposition of the series. X-12-ARIMA uses the model to provide statistically efficient estimates of the regression coefficients, and, usually, also to forecast the series. Forecast extension allows X-12-ARIMA to use symmetric seasonal and trend filters.

1.2 Judging the quality of the adjustments

For the purposes of this evaluation, we will judge the adjustments of the simulated series based primarily on how close we can come to the actual seasonally adjusted series. In Section 1.2.1., we discuss the accuracy measures we used.

We also looked for the presence of residual seasonal or calendar (trading day or holiday) effects. For an adjustment to be acceptable, there should be no residual seasonal or calendar effect present in the seasonally adjusted series or in the irregular component. We used the spectral graphs, discussed below in Section 1.2.2, to look for the presence of residual seasonal effects.

Diagnostics to check for residual seasonality and trading day effects are readily available in X-12-ARIMA. Explicit diagnostics for residual effects and revisions do not exist in SEATS. However, we can take the seasonally adjusted series from SEATS and use X-12-ARIMA to get spectral diagnostics.

X-12-ARIMA-SEATS allows us to compare diagnostics from X-12 adjustments and SEATS adjustments with ease. We can look at the same spectral diagnostics, revision history diagnostics, and sliding spans diagnostics. X-12-ARIMA-SEATS also outputs graphics files in a format compatible with X-12-Graph, the companion graphics package for X-12-ARIMA. We can then input adjustments from both X-12 and SEATS into X-12-Graph and compare diagnostics and adjustments across programs, including graphs from the revision history diagnostic.

This paper reports the general results of research undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform parties of research and to encourage discussion.

1.2.1 Measures of Accuracy

We judged the accuracy of the adjustments to the simulated series based primarily on how close we can come to the actual seasonally adjusted series. Because we created the series by combining components from existing decompositions, discussed below in section 2.1, we were able to calculate the actual seasonally adjusted series, x_t. For each adjustment we calculated a relative root mean squared deviation (RRMSQD) and a relative mean absolute deviation (RMAD), as shown below.

$$RRMSQD = \sqrt{N^{-1} \sum_{i=1}^{N} \frac{(x_t - \hat{x}_t)^2}{x_t^2}}$$

$$RMAD = N^{-1} \sum_{i=1}^{N} \frac{|x_t - \hat{x}_t|}{x_t}$$

where N is the number of data points in the series, x_t is the actual seasonally adjusted series, and $\hat{x}_t$ is the is the estimated seasonally adjusted series.

We then used these measures of accuracy to compare adjustments between the two methods. When comparing two adjustments we determined that one was more accurate if both measures (RRMSQD and RMAD) favored one adjustment. If one measure favored X-12 and the other favored SEATS, then we declared it a tie.

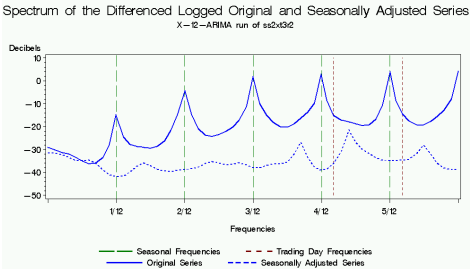
If we had found discrepancies between the measures often (ties), we would have dropped the RRMSQD and used only the RMAD. Fortunately, we did not have many ties.

1.2.2 Spectral Graphs for Residual Seasonality

X-12-ARIMA automatically estimates three spectra whenever seasonal adjustment is requested: the spectrum of the differenced original series, the spectrum of the differenced seasonally adjusted series, and the spectrum of the final irregular component. Seasonal frequencies are marked by vertical lines at k/12 cycles/month for $1 \le k \le 5$. A visually significant peak at any of the seasonal frequencies for either the seasonally adjusted series or the irregular is a signal of possible residual seasonality.

See Figure 1 for an example of a spectral graph with seasonal peaks in the original series (solid line). Notice the seasonal peaks are suppressed in the spectrum of the seasonally adjusted series (dotted line).

Figure 1. Example of Spectral Graph



2. METHODS

2.1 Constructing the Series

We constructed simulated series by combining trend and seasonal components produced by decompositions in X-12-ARIMA and SEATS with randomly assigned irregular components.

We took three different "real" economic series and adjusted them with both SEATS and X-12-ARIMA. Because both the trends and the seasonal factors can look very different even with the same original series, we wanted to use trends and seasonal factors from both programs. So with three original series, we now had six trends and six sets of seasonal factors. To limit the number of series somewhat, we decided to combine trends from SEATS with seasonal factors from X-12-ARIMA and vice versa.

We took the irregular components from three other series. We adjusted the three series with both SEATS and X-12-ARIMA. We combined the irregular factors from both programs into three files, one file for each series. Then we sampled with replacement from each of the three files to get three different sets of irregular factors.

To insure a variety of simulated series, we chose components with a variety of characteristics. The three trend components from each program consisted of one complex, very cyclical trend; one very smooth trend; and one trend somewhere in between.

Figure 2. Trend #1 from SEATS

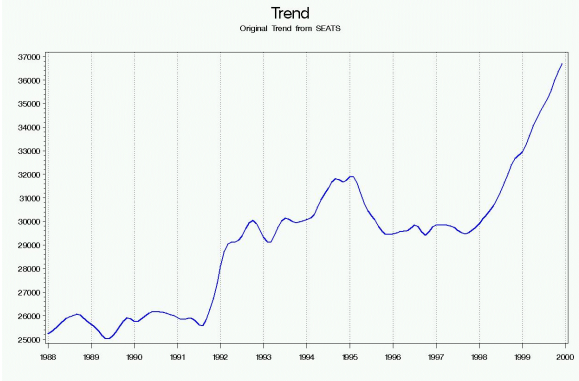


Figure 3. Trend #2 from SEATS

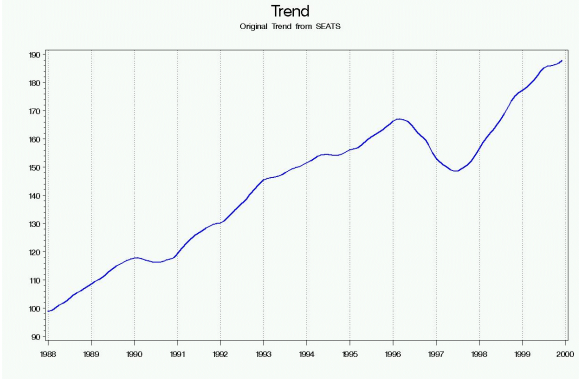
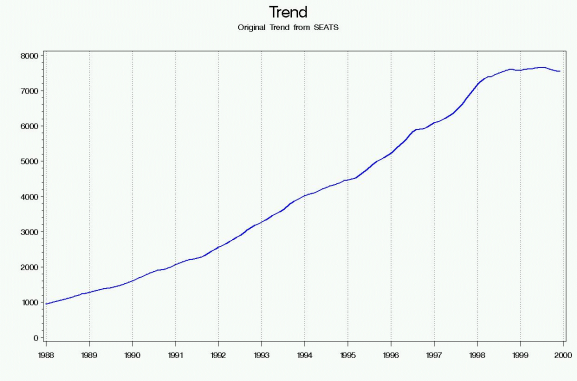


Figure 4. Trend #3 from SEATS



We chose series with different ranges for the seasonal factors and for the irregular.

Table 1.

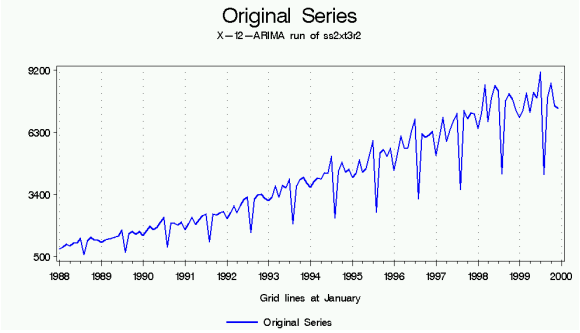
Ranges for Seasonal Factors and Irregular	
Seasonal Factors	
Very small	0.96 - 1.06
Large	0.5 - 1.2
In between	0.85 - 1.12
Irregular	
Very small	0.95 - 1.04
Large	0.4 - 1.9
In between	0.9 - 1.13

To construct the series, we combined the three trend components from SEATS with the three seasonal components from X-12-ARIMA for each of the three irregular components. Next we combined the three trend components from X-12-ARIMA with the three seasonal components from SEATS for each of the three irregular components. These combinations resulted in 54 simulated series.

Series 1 = SEATS' Trend 1 * X-12's SF 1 * Irregular 1
Series 2 = SEATS' Trend 1 * X-12's SF 1 * Irregular 2
Series 3 = SEATS' Trend 1 * X-12's SF 1 * Irregular 3
Series 4 = SEATS' Trend 1 * X-12's SF 2 * Irregular 1
Series 5 = SEATS' Trend 1 * X-12's SF 3 * Irregular 1
...
Series 52 = X-12's Trend 3 * SEATS' SF 3 * Irregular 1
Series 53 = X-12's Trend 3 * SEATS' SF 3 * Irregular 2
Series 54 = X-12's Trend 3 * SEATS' SF 3 * Irregular 3

As an example, Figure 5 shows Series #49, constructed using Trend #3 from X-12-ARIMA, Seasonal Factors #2 from SEATS, and Irregular #2.

Figure 5. Series #49



2.2 Running SEATS and X-12-ARIMA for the full series

Our basic procedure for running both programs on long series:

Step 1. Graph the series.
Step 2. Run programs in default mode.
Step 3. Decide which series should not be seasonally adjusted.
Step 4. Decide which series need further analysis, i.e., series with deficiencies in the adjustment.
Step 5. Correct problems for the series identified in Step 4.

Step 1. Graph the series.

Before we ran either program, we graphed the series to look for obvious problems with the series, such as changes in the seasonal pattern or obvious outliers.

Step 2. Run programs in default mode.

TRAMO/SEATS have an option, called RSA, for "routine treatment of perhaps a very large number of series." (Gomez and Maravall, 1997a) We used the RSA parameter set equal to six. This allows TRAMO/SEATS to

- test for a possible log transformation
- search for additive outliers, level shifts, and temporary change outliers
- search for an ARIMA model with
  - regular differences up to and including order 2
  - seasonal differences up to and including order 1
  - regular polynomials up to and including order 3
  - seasonal polynomials up to and including order 1
- replace the model in SEATS when the model chosen by TRAMO does not accept an admissible decomposition
- test for possible trading day effects (using six regression variables to get a coefficient for all seven days of the week)
- test for possible Easter effects (for six days before Easter)

We also ran X-12-ARIMA in default mode. By default mode, we mean letting the program choose options for us as much as possible. For example, the default transformation is the log transformation. But we used the setting **`transform{function=auto}`** to allow the program to choose the transformation for us.

Specifically, we allowed the program to

- test for a possible log transformation
- search for additive outliers, level shifts, and temporary change outliers
- search for an ARIMA model using new automatic modeling procedure
  - regular differences up to and including order 2
  - seasonal differences up to and including order 1
  - regular polynomials up to and including order 3
  - seasonal polynomials up to and including order 1
  - preference for balanced models (we did this so the models selected by this procedure would have a better chance of matching the model from TRAMO)
- test for possible trading day effects (using six regression variables to get a coefficient for all seven days of the week)
- test for possible Easter effects (for either one, eight, or 15 days before Easter)
- forecast 24 months
- run default X-11 (including the X-11-ARIMA/88 automatic seasonal filter selection)

Step 3. Decide which series should not be seasonally adjusted.

For this study, decisions on whether or not to seasonally adjust a series were based on the spectral graph of the original series. If TRAMO chose a nonseasonal ARIMA model, then we considered that as evidence that TRAMO/SEATS thought the series should not be adjusted. For X-12 we also used the M7 statistic, see Lothian and Morry (1978).

Step 4. Decide which series need further analysis, i.e., series with deficiencies in the adjustment.

We looked at spectral graphs of the seasonally adjusted series and the irregular component. We flagged any series with residual seasonal effects for further analysis.

Step 5. Correct problems for the series identified in Step 4.

Once we had identified the series that needed detailed analysis in Step 4, we used any diagnostics available to us to help us improve the adjustment.

2.3 Running SEATS and X-12-ARIMA with only four years of data

Our basic procedure for running both programs with only four years of data was slightly different.

- We didn't attempt model selection. We set the ARIMA model to the airline model for both SEATS and X-12-ARIMA.
- We didn't test for trading day and Easter effects.

Because of possible problems with estimating ARIMA coefficients with only four years of data, we looked at some alternative approaches to seasonally adjusting the short series.

We suspected that the MA coefficients might be overestimated, and we observed that many of the series had Θ_{12} values of 0.98. We fixed the coefficient values in SEATS to see if this gave better results. We tried various combinations for both the θ_1 and Θ_{12} parameters. We looked at θ_1 values of 0.2, 0.5, and 0.8. We looked at Θ_{12} values of 0.5, 0.8, 0.9, and 0.98. It spite of all our efforts, there was no clear preference for any of these models.

When running in default mode with 24 forecasts, X-12 generally chose 3x5 seasonal filters. Since these filters require seven years of data, we tried forecasting 36 months ahead so we would have seven years of data. Also, because of possible problems when forecasting ARIMA models with only four years of data, we also looked at running X-12-ARIMA without the 24 forecasts. X-12-ARIMA with 24 forecasts gave us the best adjustments based on RRMSQD and RMAD.

2.4 Versions of the Programs We Used

We began this research with the version of SEATS currently on the web site at the Bank of Spain, the beta DOS version dated June 1998.

We also compared the results from the DOS version of SEATS to SEATS/TRAMO for Windows beta version dated June 2000. Dr. Maravall sent us this program in July 2000. All results match the DOS version exactly.

We used version 0.3 of X-12-ARIMA, a version with the prototype of the new automatic modeling procedure.

3. RESULTS

3.1 X-12-ARIMA's New Automatic Modeling Procedure

We compared results from the new procedure to the former procedure that was based on Statistics Canada's automatic model selection procedure in X-11-ARIMA. The new procedure performed better. Models from the new procedure gave us seasonal adjustments that were closer to the correct adjustment 29 out of 54 times. The old procedure was closer for 19 series.

3.2 Results for Default Adjustments

We compared X-12-ARIMA default adjustments, using the models from the new automatic modeling procedure, to adjustments from SEATS using the RSA

option. SEATS and X-12-ARIMA did equally well on series with 12 years of data. SEATS was closer to the correct seasonally adjusted series on series with large irregular when there was 12 years of data. For series with only four years of data, X-12-ARIMA was closer to the correct seasonal adjustment, even for the series with large irregular.

Table 2. Results for Default Adjustments

	12 years	4 years
All 54 Series		
SEATS closer	25 series	11 series
X-12 closer	25 series	39 series
No preference	4 series	4 series
Total	54 series	54 series
Series with Large Irregular (18 series)		
SEATS closer	16 series	0 series
X-12 closer	2 series	17 series
No preference	0 series	1 series
Total	18 series	18 series

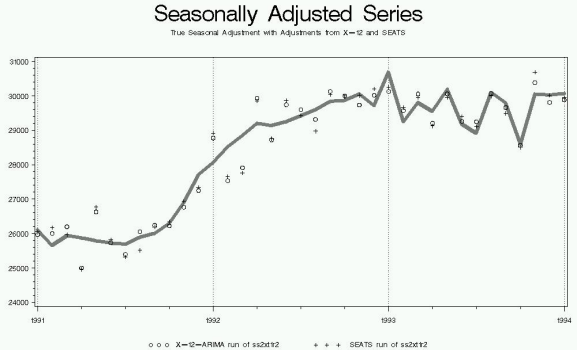
Table 3. Overall Average RRMSQD and RMAD for 12 Years of Data for All 54 Series

	12 years	4 years
RRMSQD		
SEATS	0.023	0.037
X-12	0.025	0.032
RMAD		
SEATS	0.018	0.030
X-12	0.020	0.027

The seasonal adjustments from both programs were very similar. For some series, both SEATS and X-12 were very close to the true value.

Even when the series differed from the true values, they often differed in similar ways. Figure 6 shows a closeup for the years 1991-1993 for one series.

Figure 6. Example of SEATS and X-12 Adjustments



3.3 Series That Should Not Be Adjusted

We identified eight series that should not be adjusted. We based this on the spectral graph of the original series and the M7 statistic from X-12-ARIMA. All eight series identified displayed no visually significant (six star peak in X-12-ARIMA) seasonal peaks in the original series and had M7 statistics greater than 1.0.

With SEATS it is possible to induce residual seasonality into the seasonally adjusted series by adjusting a series that shouldn't be adjusted. For more information on our findings using the spectral diagnostics, please see Section 5 on diagnostics.

3.4 Changes Made from the Default Adjustments

For X-12-ARIMA adjustments with 12 years of data, there were three series with residual seasonal peaks in the seasonally adjusted series. With a few minor changes to the ARIMA models, we were able to eliminate the residual seasonality in X-12.

For SEATS, the only series with problems with the residual seasonality were series that X-12-ARIMA indicated shouldn't be adjusted. So for SEATS, we didn't make any changes to the default models. All problems with residual seasonality for SEATS adjustments were eliminated when we eliminated the eight series with large M7 statistics.

Using the new models for X-12-ARIMA and eliminating the eight series we felt we shouldn't be adjusting, the programs continued to perform equally well based on our accuracy measures RRMSQD and RMAD. For series with the large irregular, SEATS was closer to the correct seasonally adjusted series for all 10 series.

With only four years of data, it is difficult to get accurate spectral diagnostics. None of the short series showed any signs of residual seasonality according to the spectral diagnostics. Therefore, we didn't change any of the default adjustments for the short series.

4. EASE OF USE

The learning curve for running either program in default mode is very small. Learning X-12-ARIMA overall may be more difficult because X-12-ARIMA has many more options than TRAMO/SEATS.

TRAMO/SEATS have a faster running time than X-12-ARIMA. This is not noticeable when running one series at a time, but does add up for a large number of series.

TRAMO/SEATS would be easier to use if the programs had better error messages. Error messages from TRAMO/SEATS are sometimes vague and sometimes nonexistent.

For large-scale production use, X-12-ARIMA is much easier to use, partly because of the lack of diagnostics in TRAMO/SEATS, but also because it is easier to customize output in X-12-ARIMA. X-12-ARIMA also has log files specifically designed to help users run large numbers of series.

The Windows version of SEATS has better access to the output files, but the output files themselves haven't changed. The problems listed above are still problems in the Windows version. The graphics package attached to

SEATS is much improved over the DOS version. The graphs are easier to obtain and look much nicer. There is still a problem that all the output files and the graphs refer to the points in the series by observation number instead of the date.

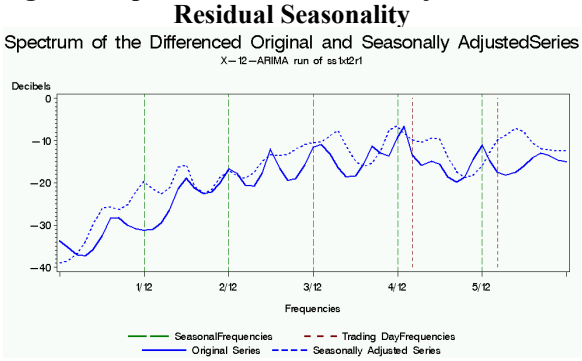
X-12-ARIMA-SEATS output files are in a format consistent with X-12-ARIMA, which makes it simple to do comparison graphs in X-12-Graph. This also means that the output files are labeled with dates.

5. DIAGNOSTICS

TRAMO and SEATS are missing some important diagnostics.

With SEATS you can induce residual seasonality into the seasonally adjusted series by adjusting a series that shouldn't be adjusted. For SEATS adjustments for series with 12 years of data, we found five series with residual seasonal peaks in the seasonally adjusted series where there were not peaks in the original series. For example, in Figure 7, notice that there are not seasonal peaks in the original series (the solid line), but there is a seasonal peak at 1/12 in the seasonally adjusted series (the dotted line).

Figure 7. Spectrum for a SEATS Adjustment with Residual Seasonality



For large-scale production work, diagnostics to warn the user not only of residual seasonality but also of series that shouldn't be adjusted would be extremely useful.

We also feel that revision diagnostics would be helpful when making model and seasonal adjustment comparisons.

6. CONCLUSION

SEATS and X-12-ARIMA both perform similarly for the simulated series using 12 years of data. SEATS does a better job with series that have a large irregular component. For the series with only four years of data, X-12-ARIMA does a better job with the series that have a large irregular component.

We would like to recommend using SEATS adjustments for production work at the Census Bureau, in particular for some of our series with large irregular. We can not make that recommendation, however, until some of the missing diagnostics are built into SEATS. Spectral diagnostics for the seasonally adjusted series and the irregular would be especially useful in SEATS.

ACKNOWLEDGMENTS

We thank Dr. Maravall for his advice and instruction on how to use TRAMO/SEATS. We thank Brian Monsell and Gianluca Capporella for writing X-12-ARIMA-SEATS. Thanks also to everyone who gave us comments on earlier versions of the paper.

REFERENCES

Burman, J.P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society*, Ser. A, 143: 321-337.

Dagum, E.B. (1988), "X-11-ARIMA/88 Seasonal Adjustment Method - Foundations and Users' Manual," Statistics Canada.

Findley, D.F., B.C. Monsell, W.R. Bell, M.C. Otto and B.-C. Chen (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program" (with discussion), *Journal of Business and Economic Statistics*, 16: 127-176.

Gomez, V. and A. Maravall (1997a), "Program TRAMO and SEATS: Instructions for the User, Beta Version," Banco de Espana.

Gomez, V. and A. Maravall (1997b), "Guide for Using the Programs TRAMO and SEATS, Beta Version," Banco de Espana.

Hillmer, S.C. and G.C. Tiao (1982), "An ARIMA-Model-Based Approach to Seasonal Adjustment," *Journal of the American Statistical Association*, 77: 63-70.

Hood, C.C. (2000), "X-12-Graph: A SAS/GRAPH® Program for X-12-ARIMA Output, User's Guide for the X-12-Graph Interactive for PC/Windows, Version 1.2," Bureau of the Census, U.S. Department of Commerce.

Lothian, J. and M. Morry (1978), "A Test of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

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

U.S. Census Bureau (2000), *X-12-ARIMA Reference Manual, Final Version 0.2*, Washington, DC: U.S. Census Bureau.