An Evaluation of TRAMO/SEATS and Comparison with X-12-ARIMA

Catherine C. Hood and David F. Findley
Catherine C. Hood, US Census Bureau, Rm 3000-4, Statistical Research Division, Washington, DC 20233-9100

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

The Census Bureau's Import/Export time series are a difficult set of series to adjust. There are 270 individual series. In addition to adjustments at the individual level, the series must also be summed to produce aggregate series, the three most important being Total Imports, Total Exports, and Balance of Trade.

These series are difficult to adjust for several reasons. First, in the production setting, there is not enough time to analyze all 270 series in detail. Therefore, relying on a seasonal adjustment program's default options for most of the series is necessary, though a limited number of important series will receive detailed analysis. Second, though many series show strong seasonality, some series show no signs of seasonality and other series are very erratic. Third, in the past, getting acceptable adjustments at the aggregate level when the series are adjusted at the individual level has been problematic. Residual seasonality in the Balance of Trade series is particularly troublesome.

To see if we could improve the adjustments for this group of series and the aggregate series, we looked at TRAMO (Time series Regression with ARIMA noise, Missing observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series). TRAMO and SEATS are linked programs developed by Victor Gomez and Agustin Maravall to seasonally adjust time series using ARIMA-model-based signal extraction techniques. We compared the performance of seasonal adjustments from TRAMO/SEATS with 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 compared the quality of the seasonal adjustment based mainly on residual seasonality in the seasonally-adjusted series and irregular component, but we also looked at the revisions of the seasonal adjustment estimates. We also note the ease of use of each program and the influence of each program's diagnostics or lack of diagnostics on the adjustment options chosen.

Summary

We compared default adjustments for both X-12-ARIMA and TRAMO/SEATS for all 270 series. In default mode, more adjustments from TRAMO/SEATS had residual seasonality than adjustments from X-12-ARIMA. In particular, we found residual seasonality when we summed the TRAMO/SEATS adjustments to get the indirect adjustment for Total Imports. There was no residual seasonality in the indirect adjustments for any of the aggregate series with X-12-ARIMA. On an individual level, 12 adjustments from TRAMO/SEATS and four series from X-12-ARIMA had residual seasonality. Our method for determining when residual seasonality is present is given in Section 1.2. More details on the results of the default adjustments are in Section 3.1.

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.

For this set of series, not all the series should be adjusted. Some series showed no signs of seasonality. Diagnostics from TRAMO/SEATS were sometimes misleading when identifying series that should not be adjusted, so we relied on the diagnostics in X-12-ARIMA. Our criteria for deciding whether or not to adjust a given series is in Section 3.2.

In practice, we don't have time in a production setting to work on a detailed analysis of all 270 series. We looked in detail at the series with 1) residual seasonality in the seasonally adjusted series or in the irregular component, and 2) large-value series that also had problematic diagnostics. More details on our selection criteria for the series that needed more attention are in Section 3.3.

After a detailed analysis of some series, we still had residual seasonality in four of the individual series adjusted with TRAMO/SEATS. We also found X-12-ARIMA had smaller revisions for most of the series. Revisions for the indirect seasonal adjustment of the aggregate series in SEATS could not be done automatically, so we could not compare revisions of the indirect adjustment we computed with X-12-ARIMA. Details on the results after detailed analysis are given in Section 3.4.

X-12-ARIMA has many features that make it easy to do seasonal adjustment for a large number of series: a variety of diagnostics to test for a variety of problems in the series, different types of output and diagnostic files, and the ability to customize output files and log files. X-12-ARIMA is also linked to a flexible graphics package, X-12-Graph (Hood, 1999). More information on the ease of use of the programs is given in Section 4.

TRAMO and SEATS are lacking in diagnostics. Using the diagnostics readily available in X-12-ARIMA is essential not only to compare TRAMO/SEATS with X-12-ARIMA, but also to improve the TRAMO/SEATS adjustments. Examples of series where X-12-ARIMA-type diagnostics helped us see problems with the TRAMO/SEATS adjustment are given in Section 5.

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 Maravall (1993) and 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 (1999).

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. Both programs have automatic procedures for model, outlier, and calendar effect identification. However, X-12-ARIMA has more predefined regressors including a stock trading day variable and a ramp. X-12-ARIMA also has a wider range of ARIMA models available; X-12-ARIMA allows higher order models and models with skipped lags.

TRAMO and X-12-ARIMA also contain similar model diagnostics. These include diagnostics to aid in model selection, such as Autocorrelation Function (ACF) values and plots, Partial Autocorrelation Function (PACF) values and plots, and the Ljung-Box Q to test for uncorrelated residuals. Both programs also include tests for normality of the residuals using skewness, kurtosis, and a Chi-square test. (Although no study supports that normality of the residuals is necessary for seasonal adjustment, normality may be necessary when computing standard errors for the seasonal adjustment.) Both programs have a Ljung-Box Q for the squared residuals to test for independence. Both programs also include the AIC and BIC to aid in model comparisons. In addition to these diagnostics, X-12-ARIMA contains spectral diagnostics for the model residuals to aid in model selection, and diagnostics for forecast errors to aid in model comparisons.

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. As stated above, however, the two programs have very different approaches to decomposition. In SEATS, the model is used in the decomposition of the series. In X-12-ARIMA, the model is used to provide statistically efficient estimates of the regression coefficients, and, usually, also to forecast or perhaps backcast the series. Forecast extension allows X-12-ARIMA to use symmetric seasonal and trend filters, and generally results in smaller revisions to the initial seasonal estimates. See Dagum (1988) and Bobbitt and Otto (1990).

The programs have very different seasonal adjustment diagnostics. SEATS diagnostics consist mainly of model-fit diagnostics. X-12-ARIMA diagnostics include spectral plots, the M and Q statistics in X-11-ARIMA (Lothian and Morry, 1972), and two kinds of stability diagnostics: sliding spans (Findley, Monsell, Shulman, and Pugh, 1990) and revision histories (Findley, Monsell, Bell, Otto, and Chen, 1998). The closest diagnostic to a stability diagnostic in SEATS is the standard errors of the components. The spectral plots and the stability diagnostics are discussed below.

1.2 Judging the quality of the adjustments

How do we judge between the two adjustments? This has been a much debated question in the statistical literature; see for example Hillmer and Bell (1984).

Although the model-fit diagnostics in TRAMO and X-12-ARIMA are very similar, they can't be used to judge the quality of the adjustment. SEATS uses filters based on the model. X-12-ARIMA uses the model only to estimate regression coefficients and to forecast the series before adjustment with the X-11-type filters.

For the purposes of this empirical evaluation, we will judge the adjustments based primarily on 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 1.2.1, to look for the presence of residual seasonal effects.

Because revisions in the seasonally adjusted series are important at the Bureau, we also looked at revisions for the seasonally-adjusted series. By revisions we mean the change or percent change from the initial estimate to the final estimate for any given point. The revision diagnostics we used are discussed in Section 1.2.2.

Diagnostics to check for residual seasonality and trading day effects and diagnostics to measure the revisions are readily available in X-12-ARIMA. Explicit diagnostics for residual effects and revisions do not exist in SEATS. Therefore, we wrote a SAS® interface (SAS Institute, Inc., 1990) to TRAMO/SEATS that would calculate the same diagnostics as in X-12-ARIMA. To check for residual seasonality and trading day effects in the seasonally adjusted series, we saved the seasonally adjusted series and the irregular component from SEATS and input them into X-12-ARIMA to compute spectral plots. We also used the interface to run TRAMO/SEATS repeatedly for the revision history diagnostic. We used the output from SEATS to compute averages of the revisions like those in X-12-ARIMA.

The SAS interface also outputs files from SEATS in a format compatible with X-12-Graph. We can then input adjustments from TRAMO/SEATS into X-12-Graph and compare diagnostics and adjustments across programs, including graphs from the revision history diagnostic.

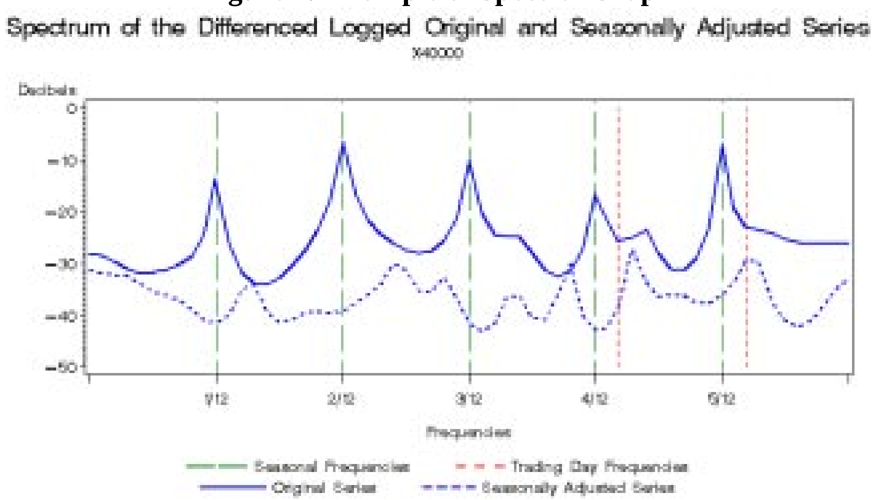
Note: The SAS interface can also call an ICON program (Griswold and Griswold, 1998) to compute sliding spans diagnostics for TRAMO/SEATS adjustments. Though sliding spans were not used in this study to compare the two adjustments, they were useful when deciding how to improve individual adjustments.

1.2.1. 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 $1 \leq k \leq 5$. Trading day frequencies are marked by vertical lines at 0.348 and 0.432 cycles/month. (See Cleveland and Devlin, 1980.) Visually significant peaks at any of the seasonal or trading day frequencies for either the seasonally adjusted series or the irregular is a signal of possible residual seasonality or trading day effect.

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



1.2.2. Revision history diagnostics for revisions to the seasonally adjusted series

The revision history procedure computes a sequence of runs from truncated sets of data. This allows us to compare revisions from the initial estimate to the most recent estimate.

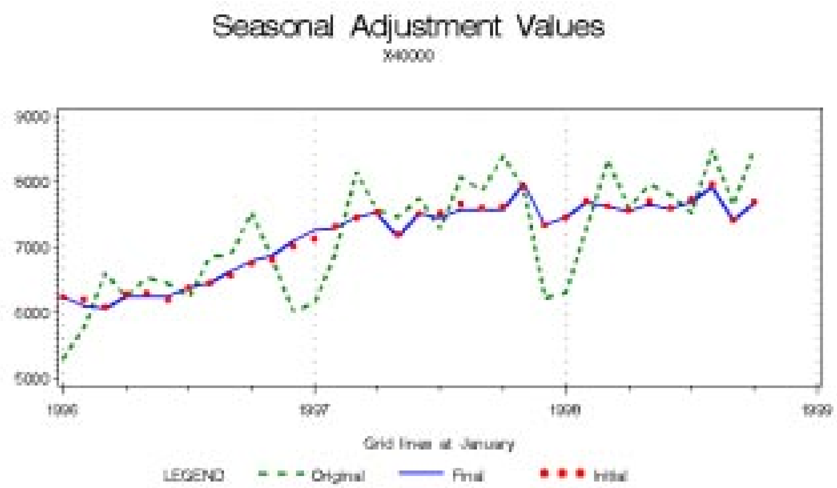
Let X_t be a time series defined for t=1,2,...,N. Let $A_{n|T}$ be the seasonal adjustment of X for observation n calculated using $X_1, X_2, \ldots, X_T$, where $n \leq T \leq N$. Define $A_{n|n}$ to be the *initial* or concurrent seasonal adjustment—the first seasonal adjustment for observation n. Define $A_{n|N}$ to be the *final* or full-series seasonal adjustment—the seasonal adjustment for observation n including all the data up to observation N.

Revisions can be quantified by the mean and maximum absolute percent difference between the initial and final estimate for the seasonal adjustment and between the initial and final estimate of the percent change for the seasonal adjustment. We looked at graphs of the initial and final adjustments.

The revisions history diagnostic was very helpful in selecting options in X-12-ARIMA. Graphs of the initial and final estimates for the last years of the series allowed us to look at individual months to see how various models and options affected the revisions to the final seasonally adjusted series and the month-to-month change in the series.

Figure 2 below shows the original series, the final seasonally adjusted series, and the initial estimates for every month from January 1996 to October 1998. For this series, we have very good agreement between the initial estimates of the seasonally adjusted series and the final estimate of the seasonally adjusted series.

Figure 2. Example of Seasonal Adjustment Revision Graph (Final and Initial) from X-12-ARIMA



The history procedure can also generate historical likelihood statistics and historical forecast errors from the regARIMA model estimation. This helps us make difficult decisions about transformations and regression effects.

Note: When running the revision history diagnostic (and sliding spans) for SEATS adjustments, we followed the advice of Gomez and Maravall (1997b) to reduce instability. Before we ran the revisions history program, we fixed the transformation choice, the model, and the existing outliers. To be consistent, in X-12-ARIMA we also fixed the transformation choice, the model, the existing outliers, and the X-11 options. We allowed the model parameters to be reestimated as in TRAMO/SEATS.

2. METHODS

Our basic procedure for running both programs:
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.

For more information on our specific methods for looking at X-12-ARIMA diagnostics and changing X-12-ARIMA options, please see Findley and Hood (1999).

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, obvious outliers, or series that shouldn't be adjusted because of a large number of missing values or zeros.

Step 2. Run programs in default mode.

<u>TRAMO/SEATS Default</u>

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, and
  - 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)

An example input file:

```
X40000 Exports of Textile Apparel
   143 1987 1 12
 795  925 1149 1058 1123 1106 1082 1104 1199 1338 1245 1067
1092 1219 1483 1378 1525 1426 1327 1442 1476 1607 1533 1505
1439 1651 1964 1839 1933 1963 1799 2028 1941 2099 1918 1699
1784 1870 2366 2170 2218 2257 2055 2244 2282 2630 2411 2101
1969 2342 2843 2918 2878 2929 2736 3022 3180 3458 3225 2751
2748 3177 3652 3573 3724 3632 3436 3747 3999 4362 3976 3322
3445 3904 4449 4104 4293 4435 4131 4438 4530 4755 4457 3788
3903 4090 4934 4487 4893 4932 4611 5206 5254 5575 5102 4370
4369 4996 5973 5506 5645 6001 5810 6444 6060 6486 5832 5122
5297 5783 6606 6246 6533 6442 6235 6848 6888 7523 6805 6019
6149 6917 8142 7559 7462 7743 7286 8070 7856 8388 7878 6218
6311 7305 8326 7598 7959 7795 7515 8485 7650 8501 7493
$INPUT RSA=6 $
```

X-12-ARIMA Default

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
- 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)
- run default X-11 (including the X-11-ARIMA/88 automatic seasonal filter selection)

Because we are in the process of changing the automatic modeling procedure in X-12-ARIMA, we ran default X-12-ARIMA runs with the model selected by TRAMO instead of using X-12-ARIMA's current automatic modeling procedure.

An example .spc file:

```
series{
  comptype = add
  title = "X40000 Exports of Textile Apparel"
  period = 12
  span = (1987.1,)
  name = X40000
  file = X40000.dat
  format = '1L'
  decimals = 2
  savelog = peaks
}
transform{function=auto}
regression{aictest=(td easter) savelog=aictest}
automdl{savelog=amd}
estimate{}
check{print=all savelog=lbq}
outlier{types=all}
forecast{maxlead=24 print=none}
x11{savelog=(m7 m10 m11 q2 fd8)}
slidingspans{savelog=percent}
```

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 and the M7 statistic from X-12-ARIMA. Once the decision was made not to adjust the series, it was not adjusted with either TRAMO/SEATS or X-12-ARIMA.

Note: M7 is one of the M statistics developed at Statistics Canada. It is a function of the difference and ratio of F-tests assessing the significance of stable seasonality and moving seasonality in the final detrended series.

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.

We also looked in detail at the series with the largest values. Based on the total import/export value in 1997, we selected the top 10 percent overall and the top 10 percent in each of ten import and export groups. This gave us an 41 large-value series.

For the large-value series, we used the diagnostics in TRAMO/SEATS to select series that needed additional analysis with TRAMO/SEATS, and we used the diagnostics in X-12-ARIMA to select the series that needed additional analysis with X-12-ARIMA. To look for inadequacies in the default TRAMO/SEATS adjustments for the large-value series, we looked at the normality diagnostics for the residuals in TRAMO and SEATS: a Chi-square test, a skewness test, and a kurtosis test. We also looked at the Ljung-Box Q statistics given at lag 24 associated with the various ACF and PACF outputs. To look for inadequacies in the default X-12-ARIMA adjustments for the large-value series, we used the M and Q statistics. We also looked at the sliding spans diagnostic and graphs of the seasonal factors by month and the SI Ratio plots from X-12-Graph. We noted warning messages produced by the program regarding residual trading day and seasonal peaks in model residuals.

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. Therefore, though we used only TRAMO/SEATS diagnostics to decide which series needed detailed analysis with TRAMO/SEATS, we used some of the diagnostics in X-12-ARIMA to help us decide how to improve the adjustment.

For the selected problematic series in both TRAMO/SEATS and X-12-ARIMA, we reviewed the choice of transformation and changed it if necessary, reviewed the ARIMA model selection, reviewed choices for trading day and Easter effects, reviewed choices for outliers, and then selected the final regARIMA model. For X-12-ARIMA, we had the additional step of reviewing the x11 spec options. These include options for filter length and for the X-11 extreme value procedure.

We used X-12-Graph to produce graphs for both the TRAMO/SEATS adjustments and the X-12-ARIMA adjustments. We looked at graphs of the original series with the seasonally adjusted series and the trend, seasonal factors by month, and the revisions. For X-12-ARIMA adjustments, we also looked at SI Ratio plots.

3. RESULTS

3.1. Default-mode adjustments

For the three aggregate series, we summed the individual series to get an indirect seasonal adjustment for the aggregate. One series, Total Imports, had residual seasonality in the indirect adjustment from TRAMO/SEATS. With X-12-ARIMA, none of the aggregate series had residual seasonality in the indirect adjustments.

On the individual level, 12 series had residual seasonality in the SEATS adjustment. Four series had residual seasonality in the X-12-ARIMA adjustment.

One of the twelve series with residual seasonality in the SEATS adjustment showed signs of a problem with the model, based on the Ljung-Box Q statistics. None of the other eleven series showed any sign of a problem in the TRAMO or SEATS output files.

3.2. Series that should not be adjusted

Model-fit diagnostics and diagnostics for the normality of the model residuals are inadequate to decide whether a series should or should not be adjusted. Diagnostics in X-12-ARIMA allow us to look for series with no seasonality in the original series or series with serious problems, such as moving seasonality.

We selected 17 series from the original 270 series that we felt should not be seasonally adjusted. Our criteria for deciding we should not adjust the series: 1) no seasonal (or trading day) spectral peaks in the original series, and 2) an M7 statistics greater than 1.2. Recall that M7 is a measure of the strength of the moving seasonality relative to the strength of the stable seasonality.

Although the cutoff for M7 is usually 1.0, we felt that small deviations above 1.0 could possibly corrected, so we set the limit higher than the normal cutoff value. All 17 series had M7 statistics greater than 1.5 although our cutoff for M7 was 1.2. We felt confident that changes to the program options could not correct this statistic and that the series should not be adjusted.

Besides M7, all 17 series also had F-statistics for stable seasonality that were less than 4.0 and failed sliding span statistics for both X-12-ARIMA and SEATS.

Again, the model-fit diagnostics in SEATS were not adequate for deciding whether or not the original series had seasonality that could be adjusted. Of the 17 series, four showed signs of problems with the normality of the residuals or problems with the ACF of the model residuals, but other series with significant seasonality had similar problems.

When we recomputed the indirect adjustments of the aggregate series, we added in the original series for the 17 series instead of adding in the seasonally adjusted series. Now we found no residual seasonality in any of the aggregate estimates for either the SEATS adjustments or the X-12-ARIMA adjustments. Some of the residual seasonality in the default adjustments was present because SEATS had chosen seasonal models for series with no seasonality and was inducing seasonality into the irregular and/or seasonal components. Three of the 17 series had residual seasonal peaks in the SEATS seasonal adjustment. Some examples of series that were not adjusted are shown in Section 5.3.

3.3. Series selected for further analysis

We decided to focus our attention on any series with residual seasonality in the seasonally adjusted series or in the irregular component. We also looked at 41 large-value series selected based on value in 1997.

We selected series with residual seasonality based on X-12-ARIMA diagnostics because TRAMO/SEATS contains no diagnostics for residual seasonality. To select large-value series for detailed analysis, we used TRAMO/SEATS diagnostics to select series for TRAMO/SEATS, and we used X-12-ARIMA diagnostics to select series for X-12-ARIMA.

For TRAMO/SEATS, we selected nine large-value series for further adjustment, some of which had residual seasonality. We also looked at six additional series with residual seasonality.

For X-12-ARIMA, we selected 19 large-value series for further adjustment, one of which had residual seasonality. We also looked at three other series with residual seasonality. More series were selected for review by X-12-ARIMA, not because the adjustments were worse, but because X-12-ARIMA has so many more diagnostics to help us find problematic series.

3.4. Detailed analysis adjustments

Once we selected the series for detailed analysis, we used any available diagnostic to improve the adjustments for both programs. So for TRAMO/SEATS, we used the spectral graphs, revisions history diagnostic, and the sliding spans diagnostic from the SAS interface.

For example, although we used the kurtosis diagnostics in TRAMO/SEATS to identify series with potential problems, we did not always choose the adjustment with favorable values for kurtosis. Often the series with more outliers have a kurtosis value closer to three but worse revisions. Therefore, we chose the adjustment with the better revisions. For an example of a series with large revisions when outliers are added, see Section 5.2.

We were not able to eliminate the residual seasonality in the TRAMO/SEATS adjustment for four of the 12 individual series with residual seasonality in the default adjustment. Details on one of the series with residual seasonality are given in Section 5.1.

3.5. Combining adjustments

We believe that if we can combine TRAMO/SEATS adjustments and X-12-ARIMA adjustments to get the aggregate series, we can improve revisions at the aggregate level. Revisions were generally smaller for X-12-ARIMA adjustments than for the TRAMO/SEATS adjustments. However, there were three large-value series where we preferred the adjustment from TRAMO/SEATS because of smaller revisions. It might be possible in the future to use TRAMO/SEATS for those series that have smaller revisions with TRAMO/SEATS and use X-12-ARIMA for the other series.

We checked for residual seasonality in the combined aggregate series. Again, without the adjustment for the 17 series and with TRAMO/SEATS adjustments for a few select series, there was no residual seasonality in any of the aggregate series.

Unfortunately, we can't look at revisions for the aggregate series – diagnostics for the revisions of aggregate series with TRAMO/SEATS are not programmed into the interface.

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 has an excellent automatic model selection procedure, which is why we used the model selected by TRAMO in the default X-12-ARIMA runs instead of the model selected by X-12-ARIMA's current automatic modeling procedure.

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 of the output from TRAMO/SEATS. Seeing which series need additional attention by paging through many different long output files is difficult, especially when there are two from each series – one from TRAMO and another from SEATS. Even a novice user of X-12-ARIMA can use options built into the program to customize the output and write important diagnostics to a log file. Options in TRAMO and SEATS to customize the output or print to a log would also make TRAMO and SEATS easier to use. Running TRAMO/SEATS in batch mode for many series also gives the user a very limited number of files that are available for the user to save. For example, if the user runs TRAMO/SEATS in batch mode, the user can save the file that contains the seasonally adjusted series, but not the files that contain the seasonal or trading day factors.

5. DIAGNOSTICS

TRAMO and SEATS are missing some very useful diagnostics. Below we list a few examples of how diagnostics already present in X-12-ARIMA would be useful when selecting seasonal adjustment options in TRAMO/SEATS.

5.1. Residual effects in the seasonally adjusted series

For four of the individual series, we were not able to eliminate the residual seasonality in the TRAMO/SEATS adjustment. One of the series, Imports of Passenger Cars (M3000R), had residual seasonality in both the default TRAMO/SEATS adjustment and the default X-12-ARIMA adjustment. For this series, both TRAMO/SEATS and X-12-ARIMA chose the airline model.

In the X-12-ARIMA output, we noticed a residual seasonal peak in the spectral plot of the regARIMA model residuals and a large spike in the ACF plot at Lag 4. We changed the model to include fixed seasonal regressors with an ARIMA model of (0 1 [1 4])(1 0 0), where [1 4] represents MA terms for Lags 1 and 4 with no terms for Lags 2 and 3. This model worked very well for forecasting, and with no other changes, the residual seasonality was gone. See Figure 3.

For TRAMO/SEATS we couldn't use the same model. TRAMO does not allow AR or MA terms greater than three and does not allow skipped lags. We tried every possible combination of terms for the nonseasonal part of the model, including (0 1 3)(0 1 1). For every possible model, there was still a visually significant seasonal peak at the 1/12 frequency in the spectral graph. See Figure 4.

Figure 3. Spectral plot for M3000R after detailed analysis — X-12-ARIMA

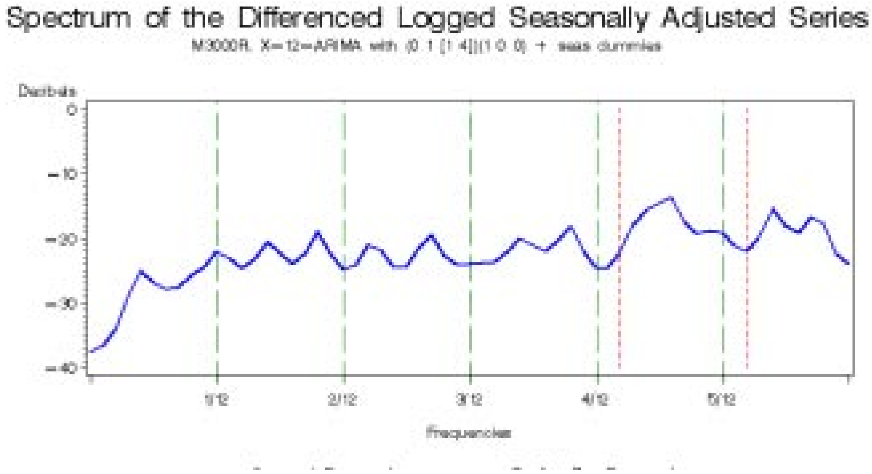
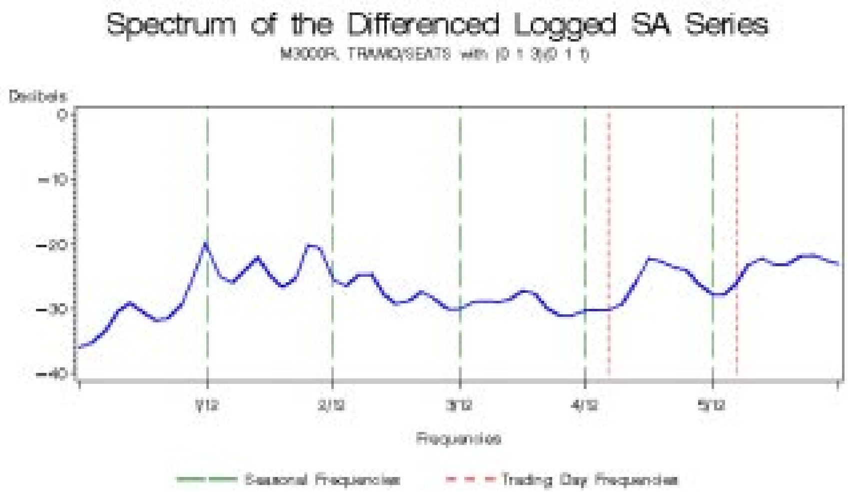


Figure 4. Spectral plot for M3000R after detailed analysis — TRAMO/SEATS



5.2. Revision diagnostics

TRAMO/SEATS users often lower the critical value for outlier detection to improve the diagnostics for normality of the residuals found in TRAMO/SEATS (a Chi-square test, skewness, kurtosis, Ljung-Box Q statistics for the residuals and the squared residuals). Often the series with more outliers is smoother but the revisions are much worse. The practice of adding outlier variables to improve normality diagnostics is problematic in TRAMO/SEATS because users can't get information about the consequences for revisions. Large revisions when extra outliers are added is common in TRAMO/SEATS adjustments. We show one example below for Exports of Passenger Cars (X3000R).

Figure 5, on the next page, shows the revisions for the default run of TRAMO/SEATS. In default mode, the kurtosis value is 4.47. We lowered the critical values for outliers until the kurtosis was within the standard error from 3.0. This occurred after we had added six additional outliers. We fixed the outliers and ran the revisions history program again. Notice the much larger revisions for most of the span shown in Figure 6. Figure 7 shows the same series with default X-12-ARIMA for comparison purposes. The revisions are smaller for the default X-12-ARIMA adjustment than for the default TRAMO/SEATS adjustment.

Figure 5. Seasonal Adjustment Revision Graph (Final and Initial) for Default TRAMO/SEATS

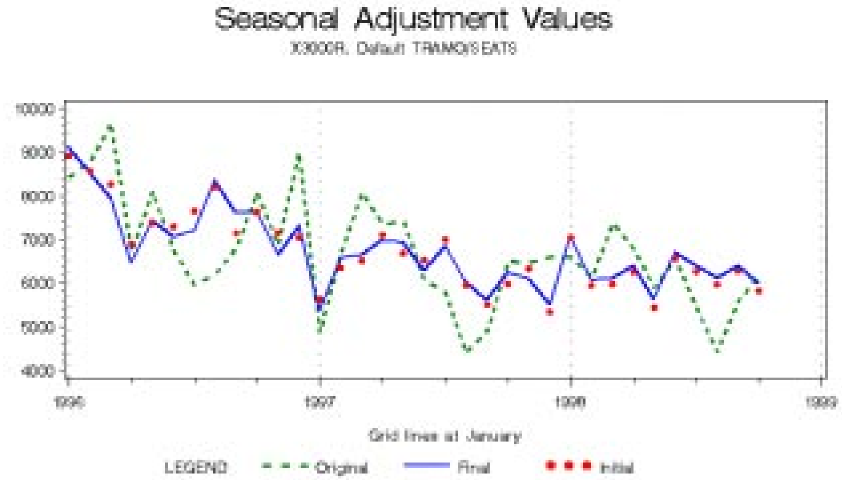


Figure 6. Seasonal Adjustment Revision Graph (Final and Initial) for TRAMO/SEATS with six additional outliers

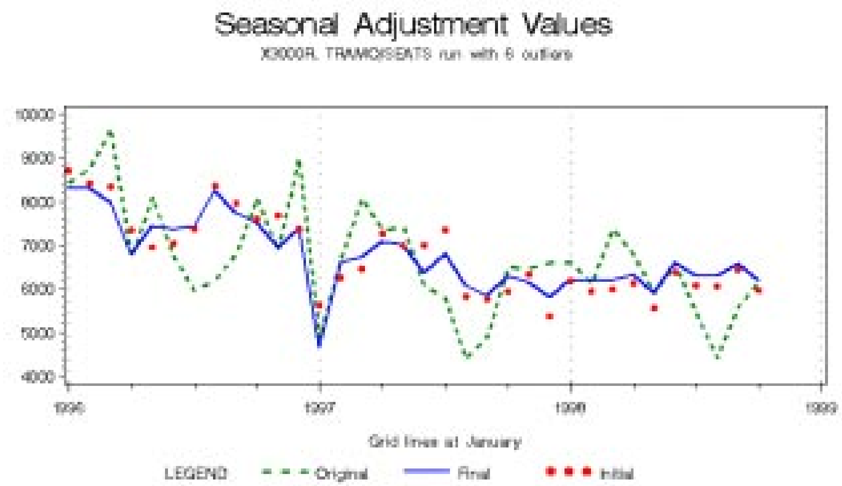
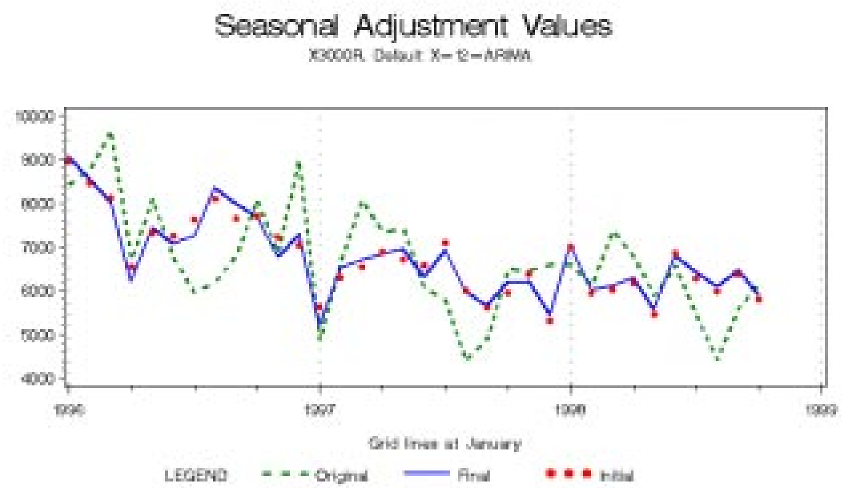


Figure 7. Seasonal Adjustment Revision Graph (Final and Initial) for Default X-12-ARIMA



5.3. Diagnostics for series that shouldn't be adjusted

For some series, it is obvious from a graph that the series shouldn't be adjusted. One such series is shown below. Figure 8 is the graph of the original series. Figure 9 focuses on the smaller values; there are gaps in the graph corresponding to the peaks seen in Figure 8. Figure 10 is the spectral graph of the original series.

This series had bad diagnostics in X-12-ARIMA: no seasonal peaks in the spectral graph of the original series, and M7 and Q statistics greater than 2.0. TRAMO/SEATS chose a model with no seasonal terms, (2 1 0)(0 0 0), and had bad model diagnostics, but no model diagnostics worse than other series that should be adjusted.

Figure 8. Graph of the original series for M22200

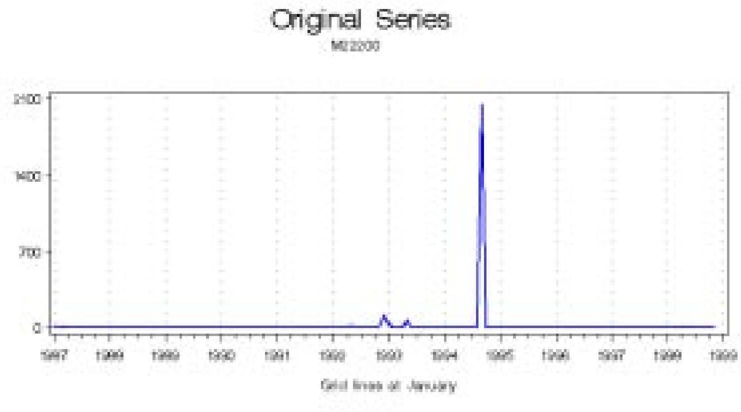


Figure 9. Graph of the original series without the larger values for M22200

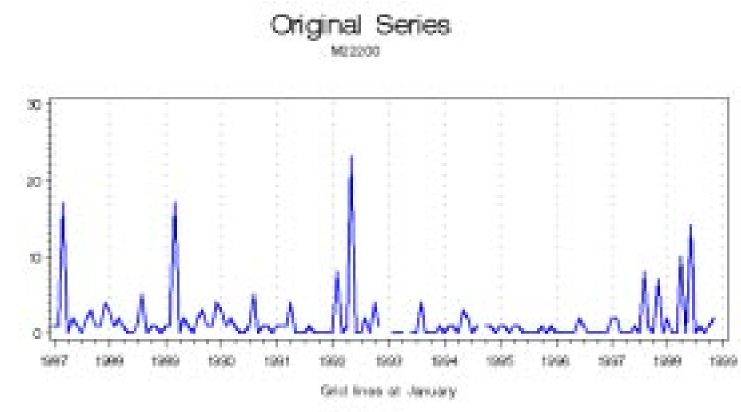
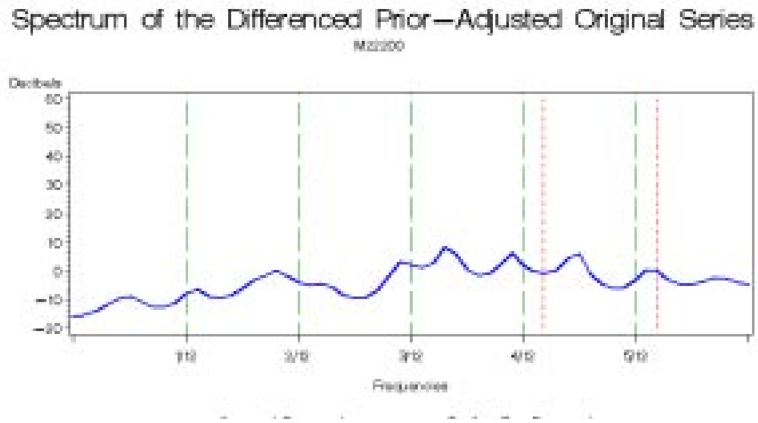


Figure 10. Spectral Graph of the original series for M22200



Another series we chose not to adjust was Imports of Copper (M14220). For this series, it is a little bit more difficult to see by the graph if the series should be adjusted or not. Figure 11 shows the original series and Figure 12 shows the spectrum of the original series.

In X-12-ARIMA, the diagnostics indicated no stable seasonality present: no seasonal peak in the original series, an F-test for stable seasonality less than 3.0, and M7 and Q statistics greater than 1.5. TRAMO/SEATS chose a model of (1 1 1)(1 0 0), and though kurtosis was above 4.0, there was no other indication that there were problems in the series.

Figure 11. Graph of the original series for M14220

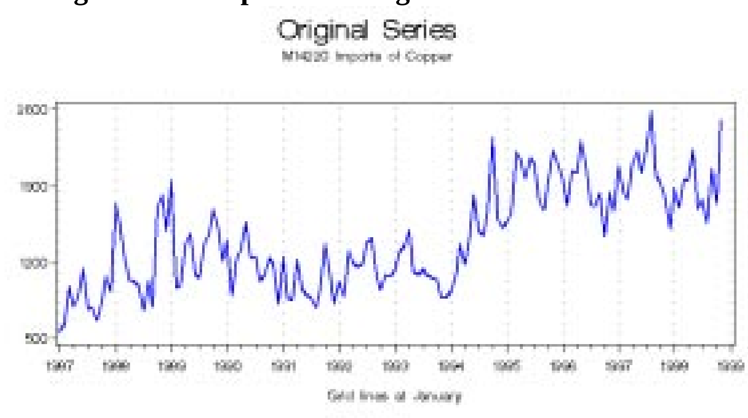
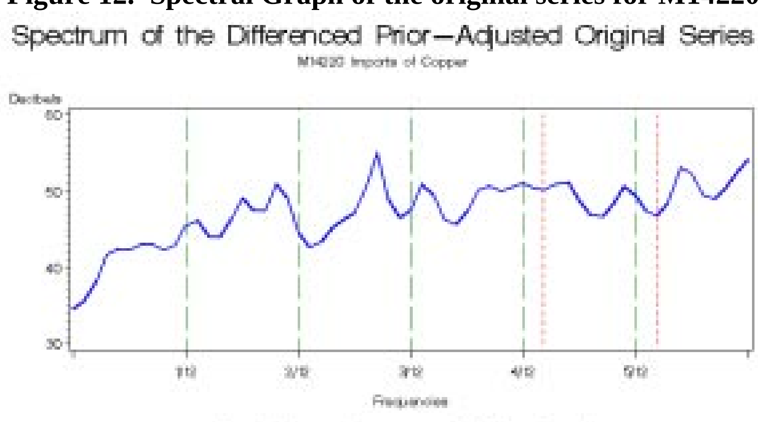


Figure 12. Spectral Graph of the original series for M14220



If TRAMO/SEATS chooses a model with no seasonal component, can we use that as a sign the series should not be adjusted? TRAMO/SEATS does not compute seasonal factors or a seasonally adjusted series in this case. Yet for two of 270 series, TRAMO chose a model with no seasonal component when the original series had visually significant seasonal peaks. For these two series, the resulting irregular component had residual seasonality. In our judgement, these series should be adjusted, and we changed the model to include a seasonal component and therefore cause SEATS to estimate seasonal factors. This eliminated the residual seasonality for the two series.

5.4. Diagnostics for indirect adjustments

In X-12-ARIMA, it is easy to see the effect on the aggregate series when we change options for an individual series. We can look at the residual seasonality in the aggregate series. We can get revision histories, sliding spans, and M and Q statistics. For any of the decisions we make, from whether or not to adjust a series to how long the seasonal filter should be, we can see how the decisions will affect the aggregate series.

While we can get spectral graphs of the aggregates of the TRAMO/SEATS adjustments, we haven't programmed any of the other diagnostics.

6. CONCLUSION

Most of the series we studied had better seasonal adjustments with X-12-ARIMA based on lack of residual seasonality and smaller revisions. The adjustments from the different programs can not be adequately compared without using diagnostics similar to those in X-12-ARIMA.

TRAMO/SEATS are missing very important diagnostics, particularly diagnostics for residual seasonal and calendar effects. Relying on the model diagnostics in the program can be misleading.

TRAMO/SEATS provide no indication of residual seasonal or calendar effects in the seasonally adjusted series or the irregular component. TRAMO/SEATS provide no diagnostics to help you decide if a series should be adjusted or not; users must rely on graphs of the original series which can themselves be difficult to interpret sometimes. TRAMO/SEATS have no diagnostics for looking at revisions, and in fact, improving the available diagnostics in TRAMO/SEATS often leads to larger revisions. TRAMO/SEATS have no diagnostics to allow us to see how the decisions we make effect the indirect adjustment of the aggregate series. For all of these important decisions, we had to rely on diagnostics either in X-12-ARIMA or program the diagnostics ourselves outside of TRAMO/SEATS.

ACKNOWLEDGMENTS

We thank Dr. Maravall for his advice and instruction on how to use TRAMO/SEATS. We also thank him for giving us his TRAMO/SEATS input files for three of the U.S. import/export series. We thank Brian Monsell for writing the sliding spans program in ICON for SEATS. Thanks also to everyone who gave us comments on earlier versions of the paper.

REFERENCES

Bell, W.R. and S.C. Hillmer (1984), "Issues Involved with the Seasonal Adjustment of Economic Time Series" (with discussion), *Journal of Business and Economic Statistics*, 2: 291-349.

Bobbitt, L. and M.C. Otto (1990), "Effects of Forecasts on the Revisions of Concurrent Seasonally Adjusted Data Using the X-11 Seasonal Adjustment Procedure," In *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 449-454, Alexandria VA: American Statistical Association.

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

Cleveland, W.S. and S.J. Devlin (1980), "Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods," *Journal of the American Statistical Association*, 75, 487-496.

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

Findley, D.F. and C.C. Hood (1999). "X-12-ARIMA and Its Application to Some Italian Indicator Series," to appear in *Seasonal Adjustment Procedures—Experiences and Perspectives*, Istituto Nazionale di Statistica (ISTAT), Rome, 231-251.

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.

Findley, D.F., B.C. Monsell, H.B. Shulman, and M.G. Pugh (1990), "Sliding Spans Diagnostics for Seasonal and Related Adjustments," *Journal of the American Statistical Association*, 85:345-355.

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.

Griswold, R.E. and M. T. Griswold (1998), *The Icon Programming Language, Third Edition,* San Jose, CA: Peer-to-Peer Communications.

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. (1999), "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.

Maravall, A (1993), "Unobserved Components in Economic Time Series," *Handbook of Applied Econometrics*, ed. M.H. Pesaran, T. Schmidt, M. Wickens, Oxford Basil Blackwell

SAS Institute, Inc. (1990), *SAS® Language: Reference, Version 6, First Edition*, Cary, NC: SAS Institute, Inc.

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 (1999), *X-12-ARIMA Reference Manual*, *Final Version 0.2*, Washington, DC: U.S. Census Bureau.