Abstract #W2113
Submitted to SUGI 2000

SAS Programs to Get the Most from X-12-ARIMA's Modeling and Seasonal Adjustment Diagnostics

Catherine C. Hood, U.S. Census Bureau, Washington, DC

KEYWORDS

Time series, Seasonal Adjustment, Graphical diagnostics

ABSTRACT

SAS®'s PROC X11 is based on the Census Bureau's X-11 seasonal adjustment program written in the early 1960's. Future SAS/ETS® Versions will contain PROC X12, based on the Census Bureau's new X-12-ARIMA program, a publicly-available program used worldwide for time series modeling and seasonal adjustment. Major advances in X-12-ARIMA include 1) the ability to use regARIMA models – regression models with ARIMA errors, 2) a new user interface, and 3) new modeling and seasonal adjustment diagnostics, including graphics.

This paper will briefly review some of the basic issues of seasonal adjustment and advances in X-12-ARIMA over X-11. I will focus on the diagnostics available in X-12-ARIMA, particularly the graphical diagnostics including spectral graphs, forecast error history graphs, and revision graphs. The Census Bureau has developed two SAS programs to provide graphical diagnostics and easier access to the diagnostics: X-12-Graph, the companion graphics package to X-12-ARIMA and a SAS interface to X-12-ARIMA. X-12-Graph uses SAS/AF®, SAS/GRAPH®, and SAS Macros and is available for the PC and Unix. The Bureau's SAS interface uses SAS/AF. This paper is intended for anyone interested in time series analysis.

INTRODUCTION

Time Series and Seasonal Adjustment

A time series is any group of repeated measurements of the same concept over consecutive equal reporting periods, such as every month or every quarter. The data must be comparable over time, so they must be consistent in the concept being measured and the way the concept is measured.

Some important features of economic indicator series include direction, turning points, and consistency between indicators. Seasonal movements can make these features difficult or impossible to see. So economists often want seasonally adjusted time series for econometric models.

The basic goal of seasonal adjustment is to decompose a time series into a trend or trend/cycle, a seasonal component, and the residual or irregular component.

The trend/cycle is the basic level of the series. It can also be thought of as the long-term tendency of the series to grow or decline. It may also include turning points and look very cyclical, or it may show steady signs of growth or decline.

The seasonal component contains any effect that is reasonably stable in terms of annual timing, direction, and magnitude. Some possible causes of seasonality are natural factors (the weather), administrative or legal measures (end of fiscal-year spending), and social/cultural/religious traditions (fixed holidays such as Christmas). The values of the estimated seasonal component are called seasonal factors.

The irregular component is anything not included in the trend or seasonal effect. The irregular is unpredictable in terms of timing, impact, and duration. Some possible causes for the irregular are sampling error, non-sampling error, unseasonable weather and natural disasters, strikes, and socioeconomic changes.

There are two basic decompositions: additive and multiplicative. The additive decomposition is appropriate when the size of the seasonal movements is independent of the level of the series. Often for economic time series, the size of the seasonal oscillations increases as the level of the series increases. For these series, if the series has only positive values, a multiplicative decomposition is usually appropriate. Under a multiplicative decomposition, the seasonally-adjusted series is obtained by dividing the original series by the estimated seasonal component. Seasonal factors for the multiplicative decomposition are centered about one.

Other calendar effects that are often estimated are trading day effects and moving holiday effects. Because they are persistent, predictable, calendar-related effects, they are included with the seasonal effects and removed from the original series with the seasonal effects.

A trading day effect relates to the number of days of the week in a given month. For example, for a series with higher activity on the weekend such as grocery sales, the number of weekends in the month will have an effect on the series for the month. So we expect a month with five weekends to have more activity than a month with only four weekends.

Inventory or stock series, where the measurement is made on an amount left in inventory on a given day of the month, have a different kind of trading day than a sales or flow series. We call these two kinds of series stock and flow. Most economic series are flow series.

Another calendar effect is moving holidays. Probably the most important moving holiday in the United States is Easter. Although we know when Easter will occur every year, it can alternate between March and April. This makes it different from a fixed holiday, such as Christmas, that can be included in the seasonal effect for the month.

X-12-ARIMA and Companion SAS Programs

X-12-ARIMA is the Census Bureau's new seasonal adjustment program. 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). X-12-ARIMA estimates seasonality with signal-to-noise ratios to choose between a fixed set of moving-average filters, often called X-11-type filters.

The major improvements in X-12-ARIMA fall into three general categories: 1) new modeling capabilities using regARIMA models—regression models with ARIMA errors—for estimating calendar effects or outliers with built-in or user-defined regressors; 2) a new user interface; and 3) new diagnostics for modeling, model selection, adjustment stability, and for the quality of indirect as well as direct seasonal adjustment. The article by Findley, Monsell, Bell, Otto, and Chen (1998) gives a detailed overview of the improvements. See also U.S. Census Bureau (1999).

X-12-ARIMA uses regARIMA models to preadjust a series before the seasonal adjustment by removing effects such as trading day, moving holidays, and outliers. Often the model is also used 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). X-12-ARIMA has automatic procedures for model, outlier, and calendar effect identification.

The new user interface allows input specification files written in "English." It also has convenient output files, including log files and diagnostics files. Users can customize log files and output files.

X-12-ARIMA model diagnostics 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. X-12-ARIMA also includes the AIC and BIC to aid in model comparisons. New X-12-ARIMA modeling diagnostics include spectral diagnostics for the model residuals to aid in model selection and diagnostics for forecast errors to aid in model comparisons.

X-12-ARIMA also contains several different seasonal adjustment diagnostics, including the M and Q statistics from X-11-ARIMA (Lothian and Morry, 1972). New X-12-ARIMA diagnostics include spectral plots and two kinds of stability diagnostics: sliding spans (Findley, Monsell, Shulman, and Pugh, 1990) and revision histories (Findley, Monsell, Bell, Otto, and Chen, 1998).

X-12-ARIMA also has diagnostics for aggregate series to judge the quality of indirect as well as direct seasonal adjustment. Additional capabilities in X-12-ARIMA make it easier to adjust large numbers of series and determine which have problematic adjustments.

Graphs are an important part of seasonal adjustment diagnostics. Graphs help the user see changes in the seasonal pattern, outliers, and changes in the level of the series. Other graphs can help the user look at the behavior of the seasonal factors. Some graphical diagnostics, such as spectral graphs and forecast error history graphs, help us with regARIMA model selection. Other graphical diagnostics help us see the size and location of the revisions in the seasonally-adjusted series or trend over time.

To make it easy to use the graphical diagnostics from X-12-ARIMA, we programmed a companion graphics package in SAS called X-12-Graph. X-12-Graph reads in graphics output files from X-12-ARIMA and gives the users a simple way to get SAS graphs, even if they are not familiar with the SAS System.

X-12-Graph comes in both a batch version and an interactive version. At the heart of both versions is the GPLOT and GREPLAY procedures in SAS/GRAPH®, templates, and annotate data sets. The batch program uses SAS Macros and can send selected graphs for any number of series directly to a printer. The interactive program uses SAS/AF® and Screen Control Language (SCL). We developed batch and interactive programs for both PC and UNIX platforms since we support X-12-ARIMA on both platforms. (See Hood, 1999a and 1999b.)

We wrote a SAS interface to X-12-ARIMA to facilitate multiple X-12-ARIMA runs. Seasonal adjustment procedures fall into two basic groups: doing a routine seasonal adjustment procedure on hundreds or even thousands of series, or doing a detailed analysis of just a few series. Often after running in routine mode, we want to identify the series that need more attention based on the available diagnostics. When running X-12-ARIMA for a large number of series, it can be difficult to keep track of all the diagnostics. The SAS interface helps us write the input specification files, and uses X-12-ARIMA's diagnostic and output files to save selected diagnostics in a SAS data set. This helps the user see which series need more attention.

X-12-GRAPH AND GRAPHICAL DIAGNOSTICS

Providing specific time series graphs is one way to help people who are using X-12-ARIMA for seasonal adjustment but are not familiar with the SAS System. There were several useful time series graphs in the literature that we wanted to program into X-12-Graph. So far, we have 14 different types of graphs available.

Although space does not permit details for every type of graph, examples of all the graphs and instructions for producing them in X-12-Graph are available in the User's Guide (Hood, 1999a and 1999b).

How do we get a good general picture of the time series? We will look at some general purpose graphs: overlay graphs and seasonal factor graphs. How do we judge the quality of the seasonal adjustment? We will look at spectral graphs and revision graphs for the seasonally adjusted series. How do we choose between regARIMA models? We will look at spectral graphs again, AIC history, and forecast error history graphs.

General Purpose Graphs

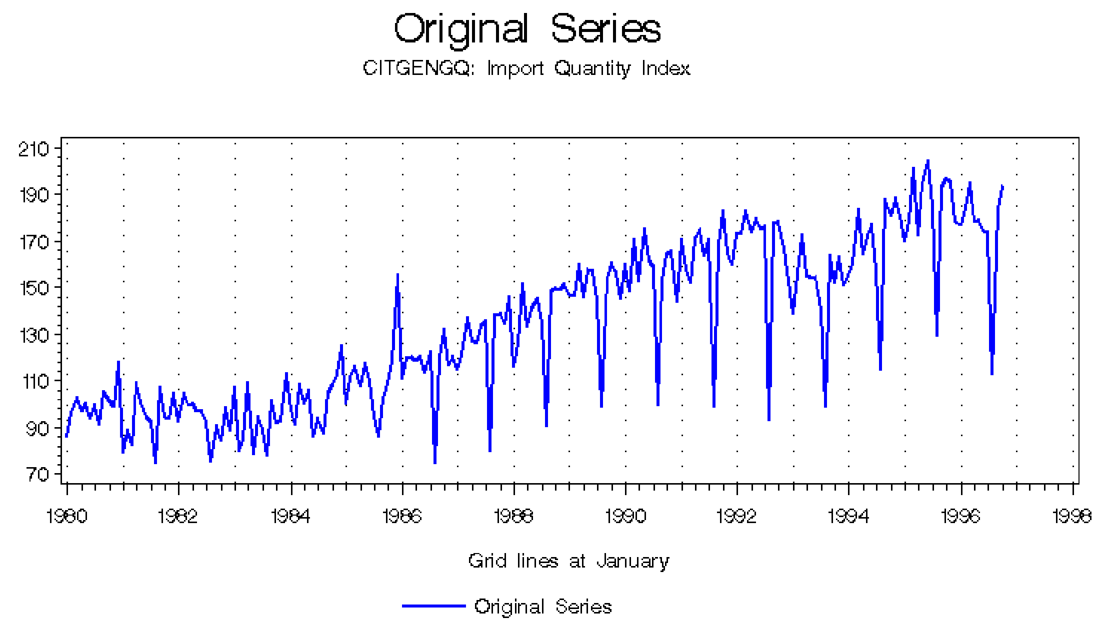
Overlay Graphs

Overlay graphs allow the user to select from one to three different graphical elements to plot above a single axis. Examples of the kind of elements in an overlay graph are the original series, the trend, and the seasonally-adjusted series. Overlay graphs are useful for getting a good general feel for the series and the seasonal adjustment.

We will look at two examples of how overlay graphs can be useful to help find a good data span for modeling or adjustment and to help see the outliers chosen by X-12-ARIMA.

The example below shows a series with a change in the seasonal pattern beginning in January, 1986. For seasonal adjustment, we used only the data span beginning in January, 1986.

Figure 1. Graph of Original Series with a Change in the Seasonal Pattern
Italian Import Quantity Index



We can also use the overlay plot to see the outliers chosen by X-12-ARIMA's automatic outlier procedure. The automatic outlier procedure, run with default options, will search for point outliers and level shifts. X-12-ARIMA also has an option to specify a "ramp" variable when the level of the series is changing over a period of a few months instead of a sudden change in level.

Figure 2 below shows the entire series for a US Export series. Figure 3 below shows the trend, original series, and outliers chosen by X-12-ARIMA for a portion of the same series: point outliers at January 1987 and December 1991; and level shift outliers at November 1987, January 1988, May 1990, July 1990, and March 1992. Figure 4 shows the same portion of the series with user specified outliers: point outliers at January 1987, May 1990, June 1990, and December 1991; a level shift outlier at March 1992; and a ramp variable from October 1987 to January 1988. Notice in particular how much smoother the trend is from April to July, 1990.

Figure 2. Graph of Trend and Original Series, with Outliers Chosen by X-12-ARIMA
US Exports of Electronic Apparatus

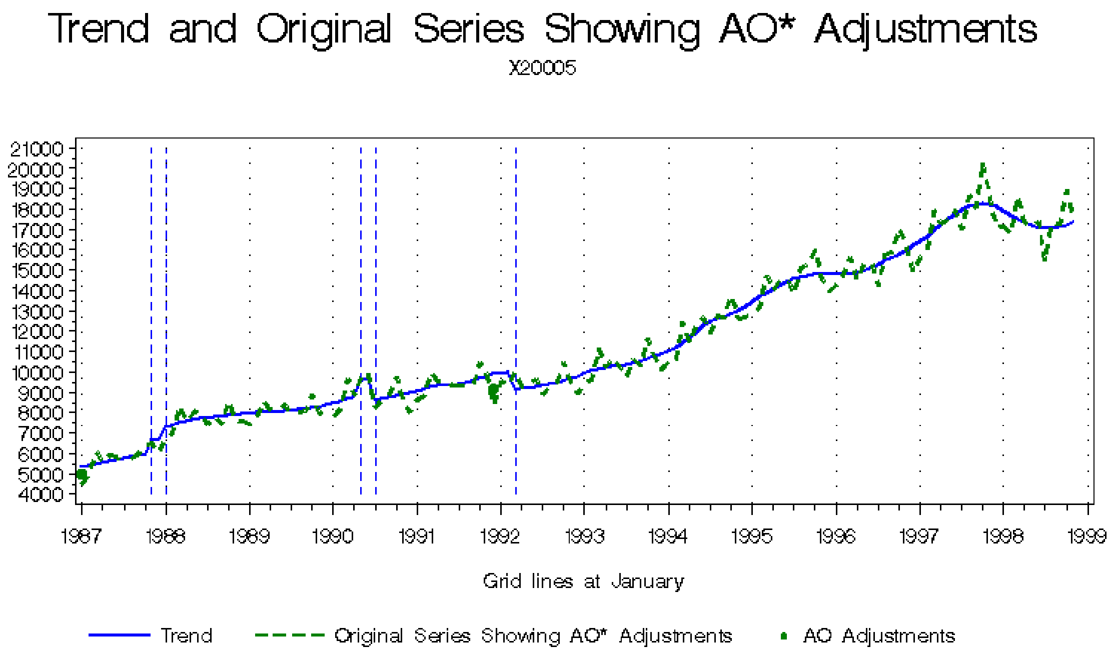


Figure 3. Graph of Trend and Original Series, with Outliers Chosen by X-12-ARIMA
US Exports of Electronic Apparatus, 1987-1992 only

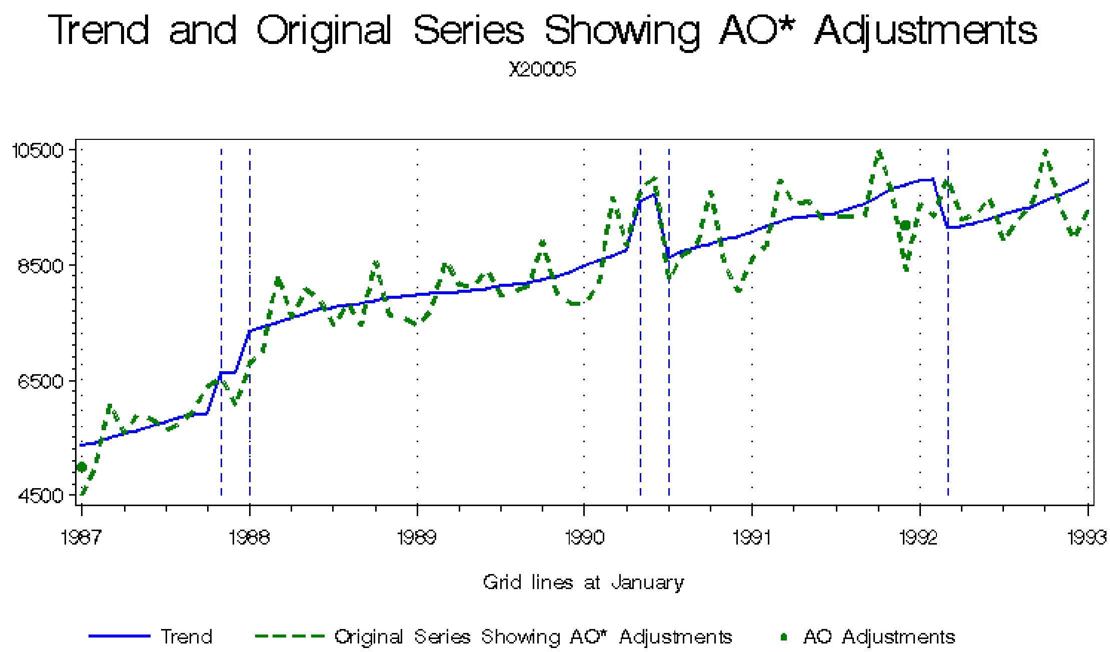
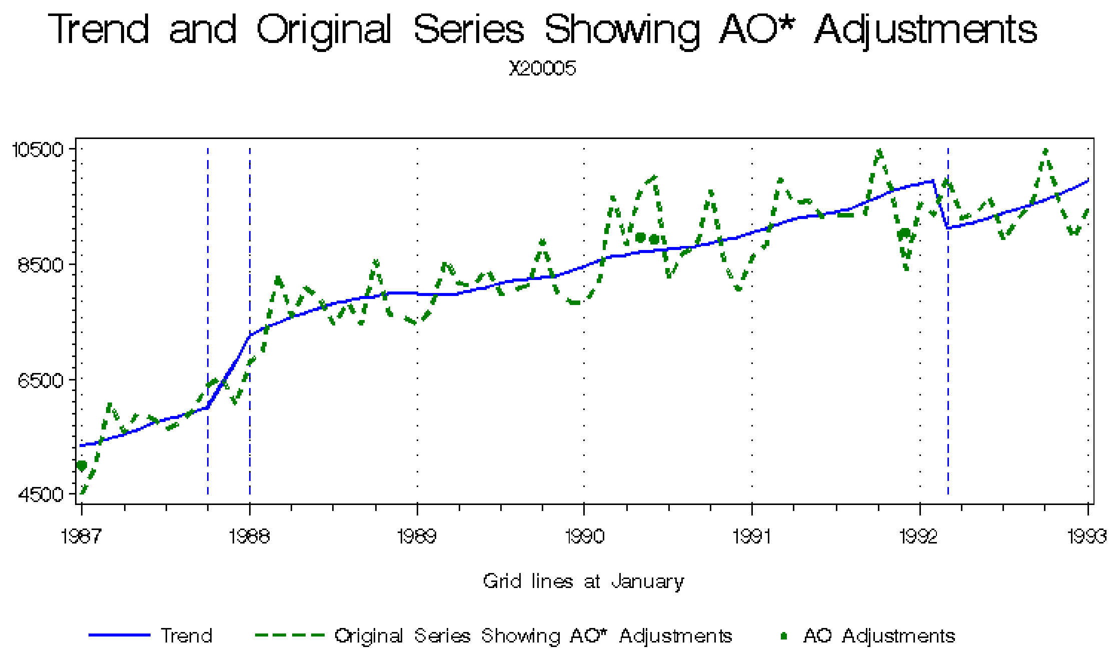


Figure 4. Graph of Trend and Original Series, with User Specified Outliers
US Exports of Electronic Apparatus, 1987-1992 only

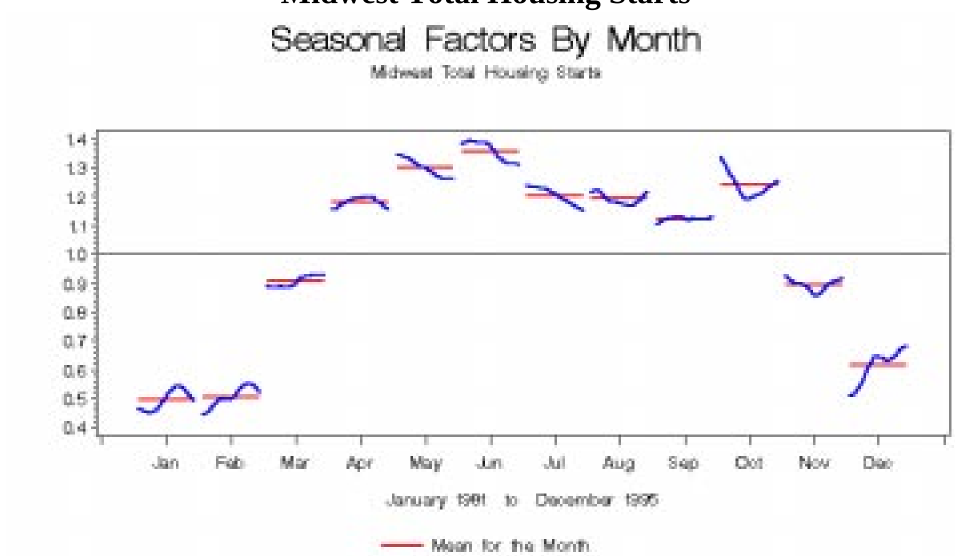


Seasonal Factors by Month or Quarter

Another useful graph is the Seasonal Factor by Month graph, developed by Cleveland and Terpenning (1982). This graph allows the user to see the seasonal pattern clearly. It also gives the user a picture of the size of the adjustments and the stability of the seasonal factors. Each calendar period has a year axis drawn at the level of its factor mean.

Notice in the graph below from a US Housing Starts series the increased activity in the summer and decrease in winter.

Figure 5. Seasonal Factors by Month
Midwest Total Housing Starts



Graphical Diagnostics for the Quality of the Adjustment

Spectral Graphs

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 use spectral graphs to look for the presence of residual seasonal effects.

Graphs of 10 times the $\log_{10}$ of the spectrum amplitudes are similar to those in the X-12-ARIMA output file. 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 $1 \leq k \leq 5$. Cleveland and Devlin (1980) identified the trading day frequencies of this graph as the frequencies most likely to have spectral peaks if a flow series has a trading day component. Trading day frequencies are marked by vertical lines at 0.348 and 0.432 cycles/month. 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.

We will show an example of where this diagnostic was useful.

When run in default mode with X-12-ARIMA, the Italian Export Quantity Index (CETGENGQ) had residual trading day effects in the regression residuals, seasonally-adjusted series, and the irregulars as indicated by the spectra of these series. In default mode, we asked both programs to test for possible trading day effects using six regression variables to obtain coefficients for the seven days of the week. The AIC preferred the model with no trading day. (Most, but not all, of the day-of-week coefficients were statistically insignificant.)

By trying all of the types of trading day models of X-12-ARIMA, we found the end-of-month stock trading day model, tdstock[31], gave the best spectrum results (and also the lowest AIC value if we fixed some negligible coefficient values to be zero). Figures 6 and 7 show the spectrum plot of the irregular series from X-12-ARIMA with no trading day adjustment (Figure 6) and with a stock trading day adjustment (Figure 7).

Figure 6. Spectrum of the Seasonally-Adjusted Series from X-12-ARIMA For CETGENGQ with No Trading Day Variables

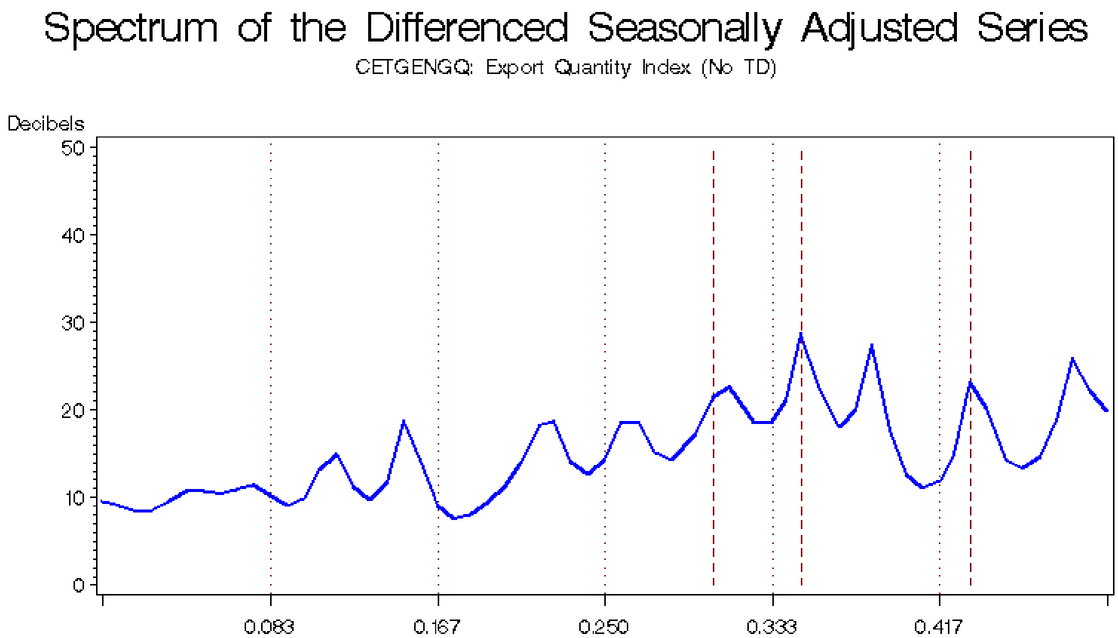
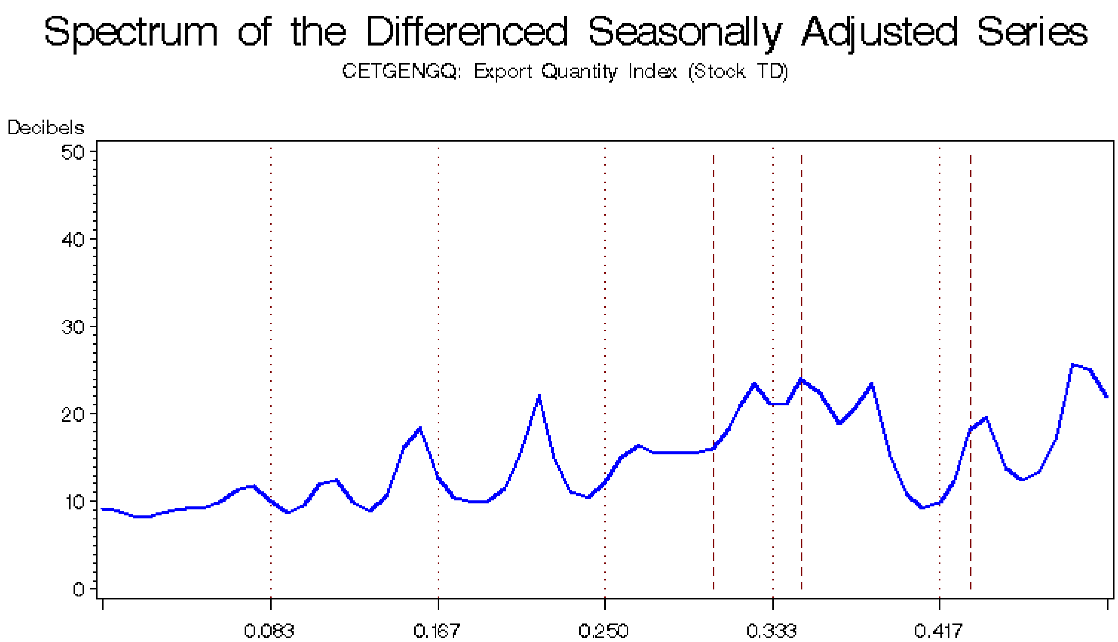


Figure 7. Spectrum of the Seasonally-adjusted Series from X-12-ARIMA For CETGENGQ with Stock Trading Day Variables



Once we realized we needed some kind of trading day regressor, we used the forecast error history diagnostics, discussed in the modeling diagnostics section, to decide between possible regressors.

Seasonal Adjustment History Graphs

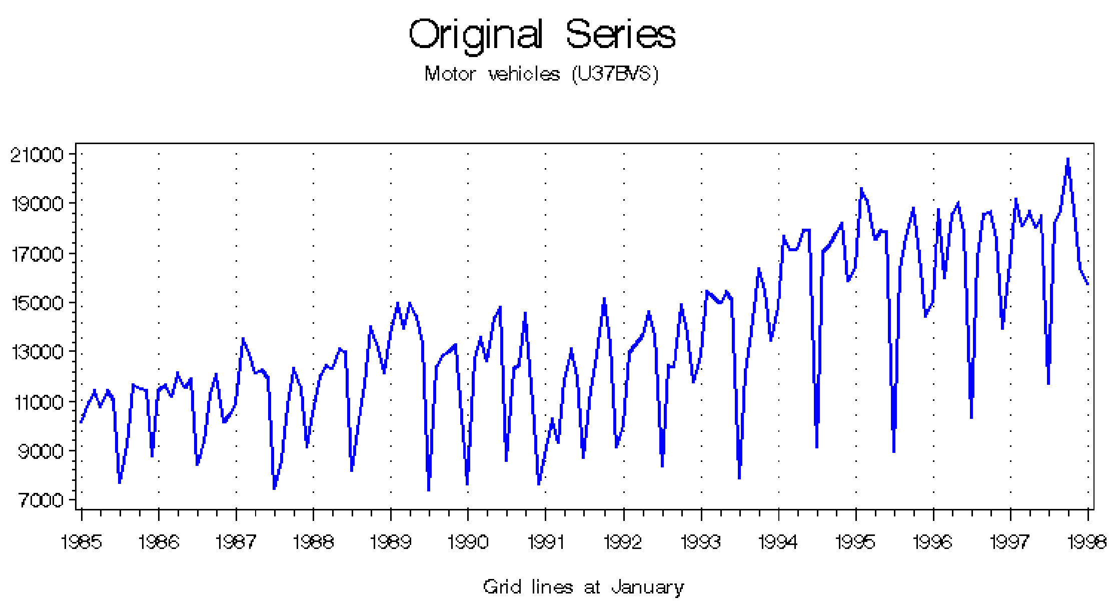
Another desirable property of the seasonally adjusted series are small revisions. Revisions are the change or percent change from the initial estimate to the final estimate for any given point.

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.

We then can calculate a mean and maximum absolute percent difference between the initial and final estimates for the seasonal adjustment and between the initial and final estimate of the percent change for the seasonal adjustment. But we can also look at graphs. Graphs of the initial and final estimates for the last years of the series allow us to look at individual months to see how various models and options affect the revisions to the final seasonally adjusted series and the month-to-month change in the series.

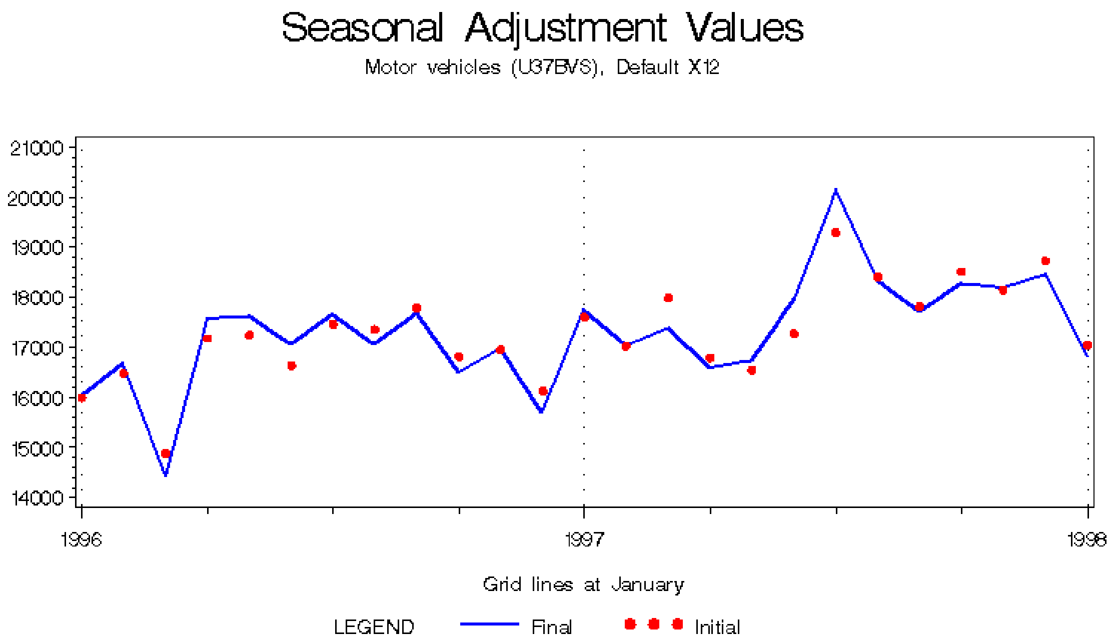
Let's look at an example — Motor Vehicle Value of Shipments. Notice in the graph of the original series there are troughs in July that are more dominant after approximately 1992.

Figure 8. Graph of Original Series
Motor Vehicle Value of Shipments



We can look at the initial and final seasonally adjusted series for the end of the series. The final estimates are graphed in a line and the initial estimates as dots. Notice some large discrepancies between the estimates for a default run of X-12-ARIMA, particularly in June 1996 and March 1997.

Figure 9. Graph of Initial and Final Estimates of the Seasonally Adjusted Series Motor Vehicle Value of Shipments, Default X-12

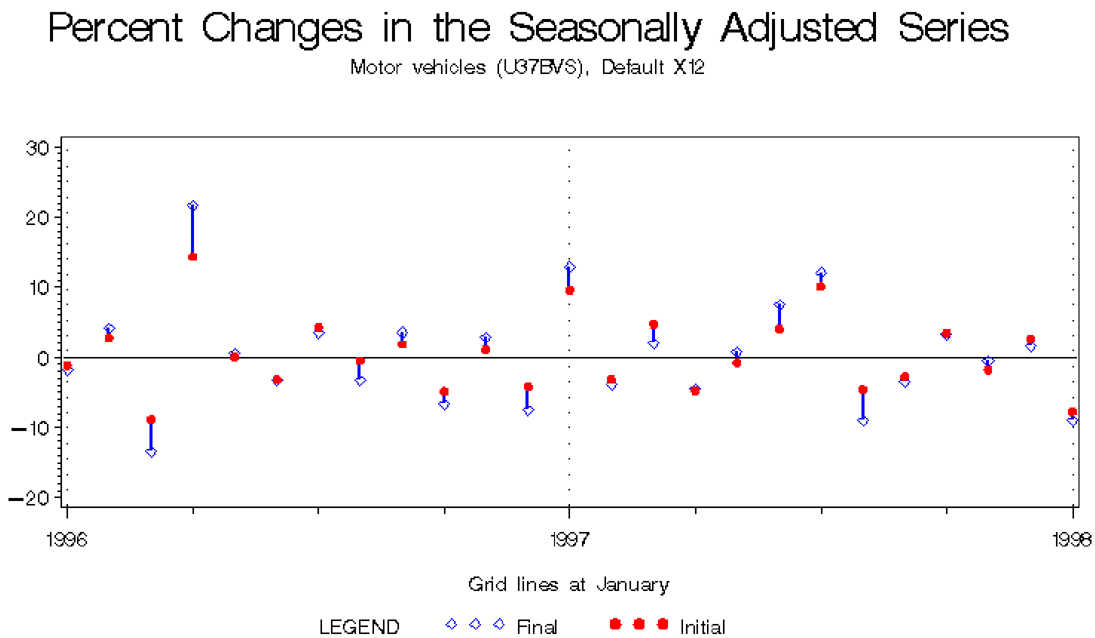


We can also look at graphs of the month-to-month (or quarter-to-quarter) percent changes. We can compare the initial estimate of the percent change to the final percent change. We are looking for differences between the initial and final estimates. It can be especially troublesome when the initial and final estimates have a different sign.

It is useful to see the initial and final percent changes as well as the difference between the two. For example, suppose the difference in the percent changes is 2%. If the percent change for the initial estimate is 3% and that of the final is 1%, then a difference of 2% is not very good. But if the percent change for the initial and final estimates are 20% and 18%, then a 2% difference is not dramatic.

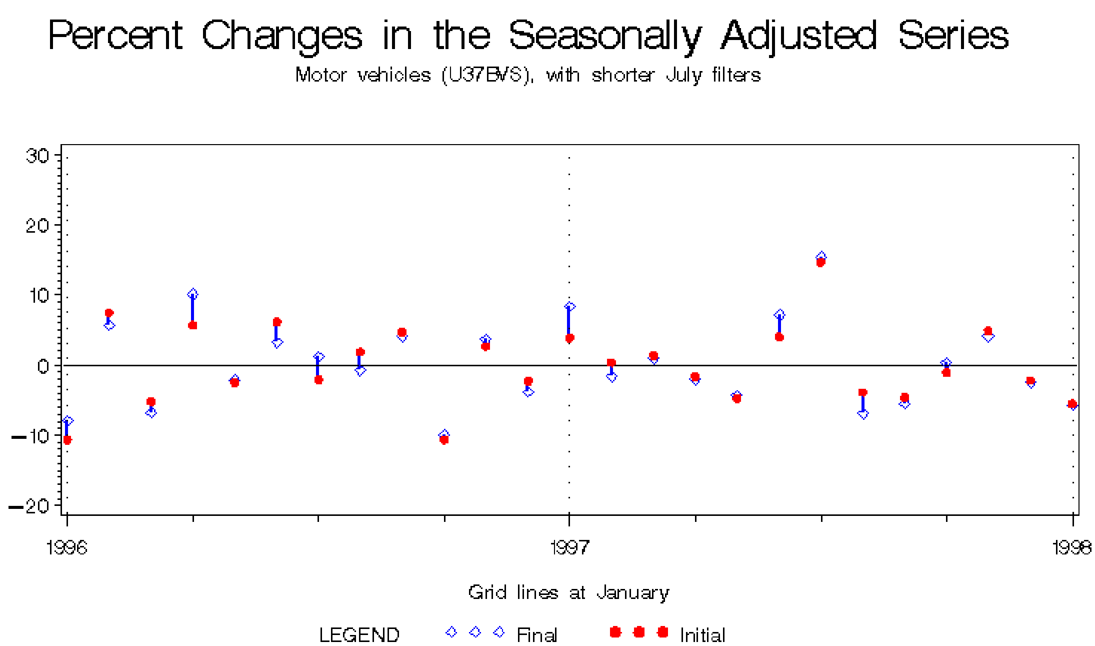
A vertical line shows the change between the two estimates. A diamond marks the final estimate. A circle marks the initial estimate.

Figure 10. Graph of Initial and Final Estimates of the Seasonally Adjusted Series Motor Vehicle Value of Shipments, Default X-12



Since the July values for this series have more variation than the other calendar months, it may be useful to use a shorter X-11 filter for July. To see if changing the filters gives us a better adjustment, we can compare the seasonal adjustment history graphs. The differences in the percent changes are smaller when we use a shorter filter for July.

Figure 11. Graph of Initial and Final Estimates of the Seasonally Adjusted Series Motor Vehicle Value of Shipments with shorter July filters



Graphical Modeling Diagnostics

Spectral Graphs

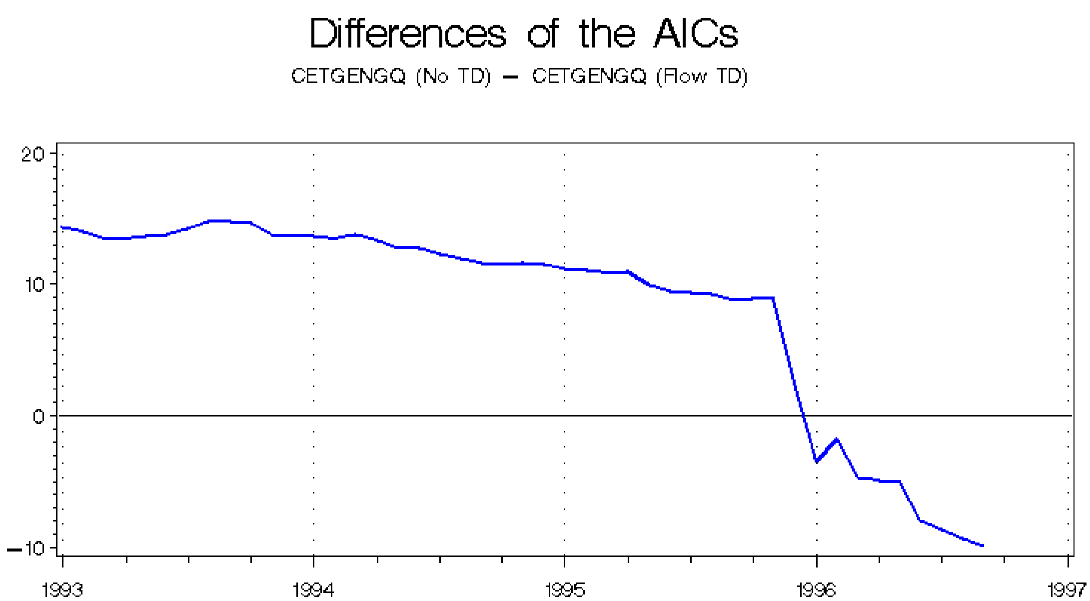
X-12-ARIMA and X-12-Graph provide an optional spectral diagnostic for the regARIMA model residuals. This is a useful tool to look for problems with the regARIMA model. Seasonal and trading day peaks may indicate a problem with the model. Common remedies include shortening the span of data for the seasonal adjustment or for the modeling, changing the type for trading day variable, and trying different ARIMA models.

Forecast Error History Graphs

Besides computing seasonal adjustment estimates from a sequence of runs from truncated sets of data, the history diagnostic can also compute historical likelihood statistics and historical forecast errors from the regARIMA model estimation on a sequence of runs from truncated set of data (Findley, 1990). Then we can compare the likelihood and forecast errors over time.

Let's go back to the example from the Italian Export Index. The spectral graph has indicated a need for a trading day variable, even though the AIC test said that a model with no trading day had a lower AIC. We can see from the graph of the AIC differences that the choice between the two options is changing over time. At the end of the series, the AIC difference is negative, which means the flow trading day variable has a higher AIC value. But earlier in the series, the AIC preferred the flow trading day regressors to a model with no trading day regressors.

Figure 12. AIC History Graph
Italian Export Quantity Index (CETGENGQ)
No Trading Day Variable vs. Flow Trading Day Variable



We can also look at forecast performance for various models. Since we use the regARIMA models to forecast the series, it is a perfectly natural thing to look at forecast performance when deciding between competing regARIMA models.

We can use the forecast error history diagnostic to choose between ARIMA models and regression variables. They can be especially helpful when the AIC values are close. We can also use the forecast error history diagnostic to choose between an additive or multiplicative model, again, especially if the AICs are close.

The forecast error history diagnostic is also useful in situations where the AIC is not appropriate, for example, when you are looking at series of different lengths. This can happen when you are trying to decide between two ARIMA models with different differencing operators, or when you want to look at possibly shortening the series.

X-12-ARIMA's *history* spec computes the sums of squared forecast errors at specified forecast leads, or forecast leads of 1 and 12 by default. We then used X-12-Graph to produce graphs of the accumulating differences between the two models. In the graphs below, the direction of the accumulating differences is predominantly downward, especially at lead 12. Thus the forecast errors are persistently smaller for the first model, the regARIMA model with Stock TD. Therefore, we prefer Stock TD over Flow TD and no trading day adjustment.

Figure 13. Stock TD Versus No TD for CETGENGQ

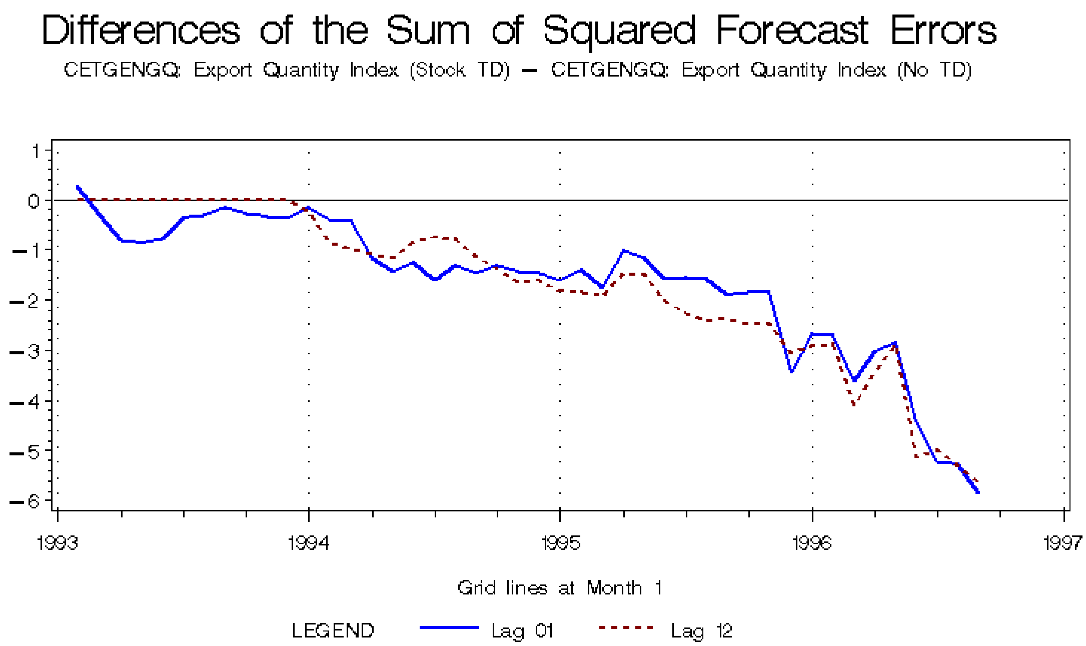
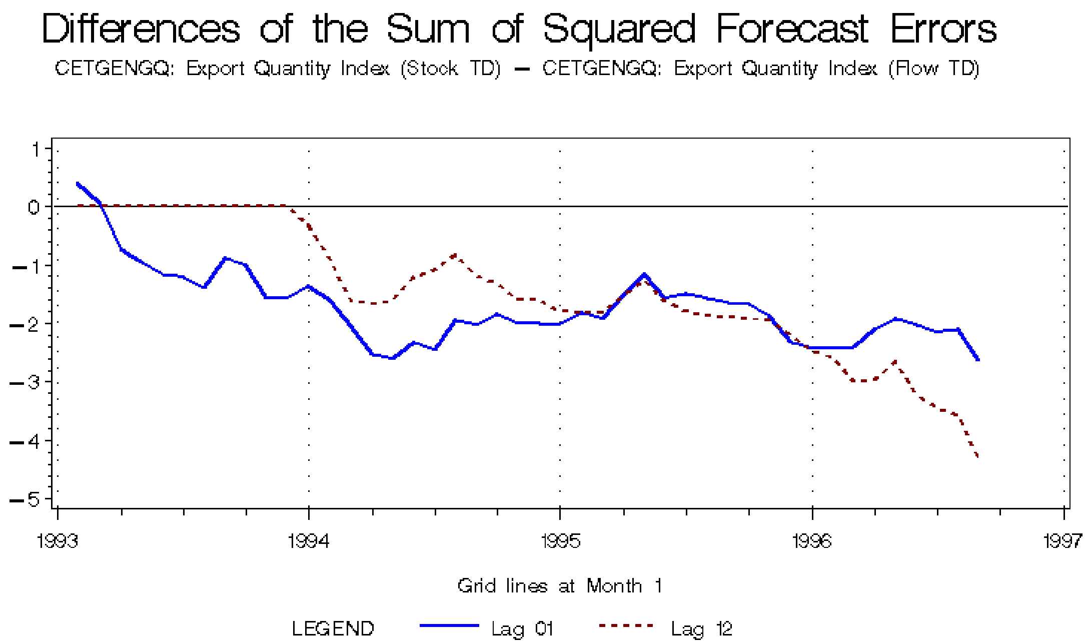


Figure 14. Stock TD Versus Flow TD for CETGENGQ



SAS INTERFACE TO X-12-ARIMA

This program was initially written to facilitate multiple X-12-ARIMA runs. The program also collects several useful diagnostics to facilitate comparisons from the multiple runs.

PLANS FOR THE FUTURE

A. Graphs for sliding spans diagnostic
B. Report options in the interface

CONCLUSION

With the SAS System, we have been able to develop a set of very useful programs for graphical seasonal adjustment diagnostics. X-12-Graph is easy to use and has a wide range of options for different types of graphs and options for those graphs.

ACKNOWLEDGMENTS

David Findley and Brian Monsell provided valuable suggestions for the design of the graphs, the layout of the program, and the documentation. Thanks also go to Kathleen McDonald-Johnson for her help in testing and design. Many thanks to everyone who gave us suggestions.

REFERENCES

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.

Cleveland, W.S. and I. Terpenning (1982), "Graphical Methods for Seasonal Adjustment," *Journal of the American Statistical Association*, 77: 52-72.

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

Findley, D.F. (1990), "Making Difficult Model Comparisons Graphically," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Alexandria, VA, 622-627.

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.

Hood, C.C. (1999a), "X-12-Graph: A SAS/GRAPH® Program for X-12-ARIMA Output, User's Guide for X-12-Graph Interactive for PC/Windows, Version 1.2," U.S. Census Bureau: Washington, DC.

Hood, C.C. (1999b), "X-12-Graph: A SAS/GRAPH® Program for X-12-ARIMA Output, User's Guide for X-12-Graph Batch, Version 1.1," U.S. Census Bureau: Washington, DC.

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. Bureau of the Census (1998), *X-12-ARIMA Reference Manual, Final Version 0.1*, Washington, DC: U.S. Bureau of the Census.

APPENDIX

X-12-Graph is available by anonymous FTP from the Census Bureau's site (ftp.census.gov). All X-12-Graph distribution files contain the User's Guide in both PostScript and Adobe PDF format. You will find detailed instructions for installation and use of the program in the User's Guide.

All versions require Final X-12-ARIMA Version 0.1 or later. X-12-Graph Interactive requires SAS Version 6.11 or later.

The names and locations of the distribution files are listed below. They can be downloaded with an FTP client or WWW brower.

X-12-Graph for IBM® Compatible PCs

In your internet browser, go to
http://www.census.gov/pub/ts/x12a/final/pc

Available files:

x12gintr.exe: A compressed self-extracting file containing eight files needed to install and run X-12-Graph Interactive: initx12g.sas, main.sc2, help.sc2, templt.sc2, namelist.sd2, indx.sd2, x12gpc.ps, and x12gpc.pdf. The last two files contain the User's Guide, which includes instructions for installation.

x12gbat.exe: A compressed self-extracting file containing six files needed to install and run X-12-Graph Batch: x12gmac.sas, templt.sc2, namelist.sd2, indx.sd2, x12gbat.ps, and x12gbat.pdf. The last two files contain the User's Guide, which includes instructions for installation.

how2down: Instructions for downloading Final X-12-ARIMA

X-12-Graph for UNIX Workstations

In your internet browser, go to
http://www.census.gov/pub/ts/x12a/final/unix

Available files:

x12gintr.tar.Z: A compressed tar file containing eight files needed to install and run X-12-Graph Interactive: initx12g.sas, main.sct01, help.sct01, templt.sct01, namelist.ssd01, indx.ssd01, x12gunix.ps, and x12gunix.pdf.

x12gbat.tar.Z: A compressed self-extracting file containing six files needed to install and run X-12-Graph Batch: x12gmac.sas, templt.sct01, namelist.ssd01, indx.ssd01, x12gbat.ps, and x12gbat.pdf.

how2down: Instructions for downloading Final X-12-ARIMA

For more information on X-12-Graph or the SAS interface to X-12-ARIMA, please contact Catherine Hood by email at chood@census.gov or by telephone at (301) 457-4985.