SEASONAL ADJUSTMENT USING THE X12 PROCEDURE

Tammy Jackson and Michael Leonard
SAS Institute, Inc.

Introduction

The U.S. Census Bureau has developed a new seasonal adjustment/decomposition algorithm called X-12-ARIMA that greatly enhances the old X-11 algorithm. The X-12-ARIMA method modifies the X-11 variant of Census Method II by J. Shiskin A.H. Young and J.C. Musgrave of February 1967 and the X-11-ARIMA program based on the methodological research developed by Estela Bee Dagum, Chief of the Seasonal Adjustment and Time Series Staff of Statistics Canada, September 1979. The X12 procedure is a new addition to SAS/ETS software that implements the X-12-ARIMA algorithm developed by the U.S. Census Bureau (Census X12). With the help of employees of the Census Bureau, SAS employees have incorporated the Census X12 algorithm into the SAS System. The X12 procedure was experimentally introduced in Release 8.0, and after careful testing it was introduced for production in Release 8.1. It has since been enhanced for Release 8.2.

There have been numerous papers on the X-12-ARIMA algorithm. This paper provides a brief summary of the algorithm with references for the interested reader. It also summarizes the benefits of using the SAS System for Census X-12 seasonal adjustment/decomposition, briefly describes how to use the X12 procedure, and provides examples that compare the Census X-12 program to the X12 procedure. More details of the X12 procedure can be found in the *SAS/ETS Users Guide, Release 8.1*.

The X12 Procedure Summary

The X12 procedure seasonally adjusts monthly or quarterly time series. The procedure makes additive or multiplicative adjustments and creates an output data set containing the adjusted time series and intermediate calculations.

The X-12-ARIMA program combines the capabilities of the X-11 program (Shiskin, Young, and Musgrave 1967), the X-11-ARIMA/88 program (Dagum 1988), and introduces some new features (Findley et al. 1988). Thus, the X-12-ARIMA program contains methods developed by both the U.S. Census Bureau and Statistics Canada. The four major components of the X-12-ARIMA program are regARIMA modeling, model diagnostics, seasonal adjustment using enhanced X-11 methodology, and post-adjustment diagnostics. Statistics Canada's X-11 method fits an ARIMA model to the original series, then uses the model forecast and extends the original series. This extended series is then seasonally adjusted by the standard X-11 seasonal adjustment method. The extension of the series improves the estimation of the seasonal factors and reduces revisions to the seasonally adjusted series as new data become available.

Seasonal adjustment of a series is based on the assumption that seasonal fluctuations can be measured in the original series (O_t, t = 1,..., n) and separated from the trend cycle, trading-day, and irregular fluctuations. The seasonal component of this time series, S_t, is defined as the intrayear variation that is repeated constantly or in an evolving fashion from year to year. The trend cycle component, C_t, measures variation due to the long-term trend, the business cycle, and other long-term cyclical factors. The trading-day component, D_t, is the variation attributed to the composition of the calendar. The irregular component, I_t, is the residual variation. Many economic time series are related in a multiplicative fashion ($O_t=S_tC_tD_tI_t$) and others are related in an additive fashion ($O_t=S_t + C_t + D_t + I_t$). A seasonally adjusted time series, C_tI_t or $C_t + I_t$, consists of only the trend cycle and irregular components.

Summary of Usage

The X12 syntax contains the following statements:

```
PROC X12 options;
        BY variables;
        ID variables;
        TRANSFORM options;
        ESTIMATE;
        IDENTIFY options;
        REGRESSION options;
        ARIMA options;
        X11 options;
        FORECAST options;
        VAR variables;
        OUTPUT options;
RUN;
```

The PROC X12 statements perform basically the same function as the Census Bureau's X-12-ARIMA specs. *Specs* or specifications are used in X-12-ARIMA to control the computations and output. The PROC X12 statement performs some of the same functions as the Series spec in the Census Bureau's X-12-ARIMA software. The TRANSFORM, ESTIMATE, IDENTIFY, REGRESSION, ARIMA, X11, and FORECAST statements are designed to perform the same functions as the corresponding X-12-ARIMA *specs*, although full compatibility is not yet available.

The online help, online documentation, and printed documentation describe the X12 procedure syntax in greater detail. The Census Bureau documentation *X-12-ARIMA Reference Manual* can also provide added insight about the functionality of these statements. Appendix A contains a cross-reference between the X12 procedure and the X-12-ARIMA syntax.

Summary of Benefits

The X12 procedure is seamlessly incorporated into the SAS system. As with other analytical tools provided by SAS, this incorporation provides the following benefits:

Data Storage
Data can be efficiently stored in SAS data sets or warehoused in SAS data warehouses. Once data is stored in the SAS System, the X12 procedure and other analytical procedures can be used to analyze the data.

Data Preparation
The SAS language (DATA Step) of Base SAS can be used to prepare generic data for analysis. The EXPAND procedure of SAS/ETS software can be used to prepare time series data for time series analysis, decomposition, adjustment, modeling, and forecasting.

Output Delivery System (ODS)
ODS allows the output of the SAS procedures to be directed to a variety of destinations. These destinations include HTML (Web pages), Listing (Output Window), Printer (Network Printer), Output (SAS Data Set), and others. ODS also allows the format of the output to be customized as desired. In particular, the output of the X12 procedure can be customized to create reports specific to the needs of the organization.

Graphics
SAS/GRAPH software is the information and presentation graphics component of the SAS System. High-quality graphics can be generated for time series data. In particular, seasonal decomposition/adjustment graphs can be created from the data sets created by the X12 procedure.

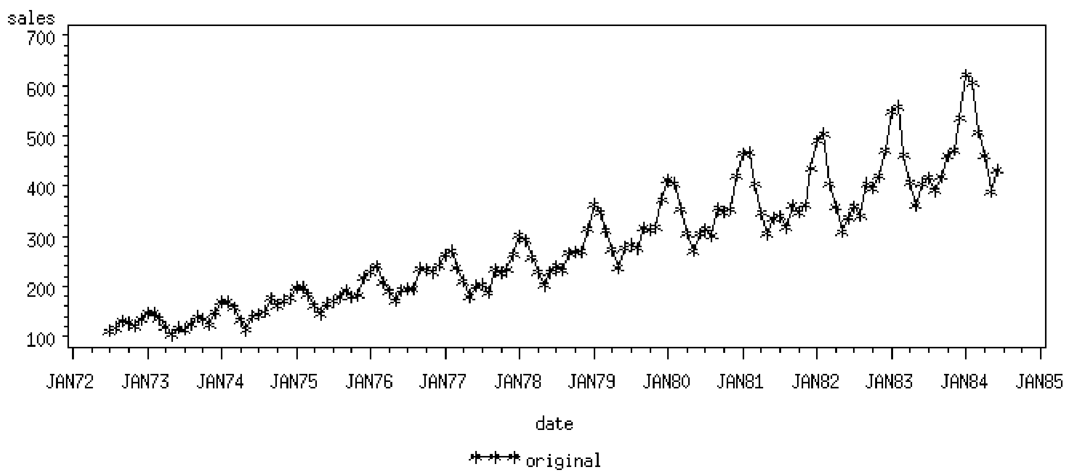
Application Development
SAS/AF (SCL based) or SAS/WebAF (Java based) applications can be custom-built for specific data analysis needs. In particular, applications for seasonal decomposition/adjustment using the X12 procedure and other analyses such as time series forecasting can be custom-built to address the specific needs of an organization.

Cross-Platform Compatibility
SAS programs and applications work on most major operating systems. SAS programs and applications developed on one platform can be used on other platforms

As shown, the SAS system provides many benefits for the seasonal decomposition/adjustment.

Examples of Usage

The following examples compare the syntax and output of the Census X-12 Spec File and the X12 procedure. Each of the following examples uses twelve years of monthly sales data (SALES). The sales data is plotted in the graph below.



Example 1

In this first example, the data is log transformed (POWER=0) and time series identification is specified. The IDENTIFY Spec in the Census X-12 program is compared to the IDENTIFY statement in the X12 procedure. As can be seen, the syntax is very similar. The IDENTIFY spec/statement determines the appropriate simple and seasonal differencing as well as tentatively identifying the ARMA(p,q)(P,Q)s orders.

EXAMPLE 1	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(112 118 132 129 121 135 148 148 136 119 104 118 115 126 141 135 125 149 170 170 158 133 114 140 145 150 178 163 172 178 199 199 184 162 146 166 171 180 193 181 183 218 230 242 209 191 172 194 196 196 236 235 229 243 264 272 237 211 180 201 204 188 235 227 234 264 302 293 259 229 203 229 242 233 267 269 270 315 364 347 312 274 237 278 284 277 317 313 318 374 413 405 355 306 271 306 315 301 356 348 355 422 465 467 404 347 305 336 340 318 362 348 363 435 491 505 404 359 310 337 360 342 406 396 420 472 548 559 463 407 362 405 417 391 419 461 472 535 622 606 508 461 390 432)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; 112 118 132 129 121 135 148 148 136 119 104 118 115 126 141 135 125 149 170 170 158 133 114 140 145 150 178 163 172 178 199 199 184 162 146 166 171 180 193 181 183 218 230 242 209 191 172 194 196 196 236 235 229 243 264 272 237 211 180 201 204 188 235 227 234 264 302 293 259 229 203 229 242 233 267 269 270 315 364 347 312 274 237 278 284 277 317 313 318 374 413 405 355 306 271 306 315 301 356 348 355 422 465 467 404 347 305 336 340 318 362 348 363 435 491 505 404 359 310 337 360 342 406 396 420 472 548 559 463 407 362 405 417 391 419 461 472 535 622 606 508 461 390 432 run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
identify{diff=(0, 1) sdiff = (0, 1)}	identify diff=(0,1) sdiff=(0,1);
	run;

Example 2

Continuing from the first example, the ARIMA Spec in the Census X-12 program is compared to the ARIMA statement in the X12 procedure. As can be seen, the syntax is similar. The ARIMA spec/statement specifies the simple and seasonal differencing as well as the ARMA(p,q)(P,Q)s orders.

EXAMPLE 2	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
	run;

Example 3

Continuing from the second example, the X11 Spec in the Census X-12 program is compared to the X11 statement of the X12 procedure. The X11 spec/statement specifies X-11 decomposition.

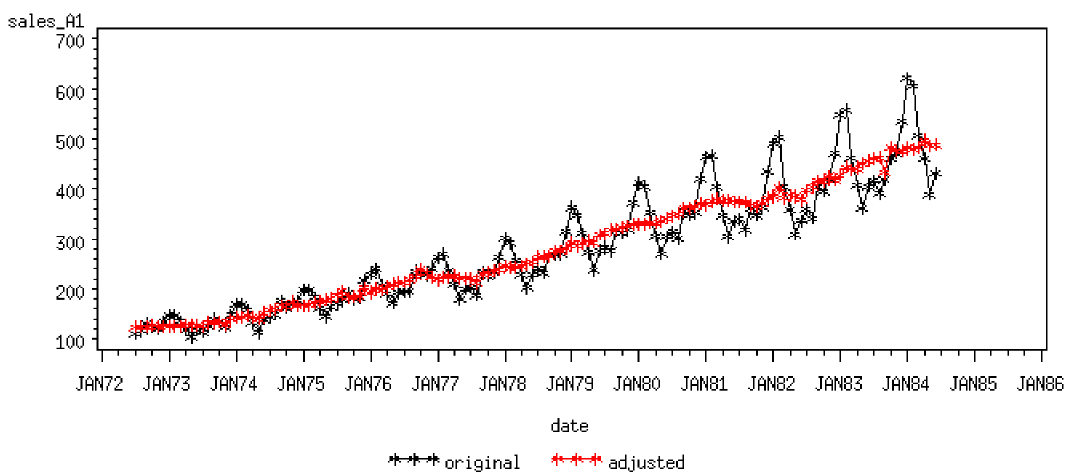
EXAMPLE 3	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	run;

Example 4

Example 3 has been expanded to include an output statement. SAS/GRAPH is used to plot the original and seasonally adjusted series contained in the dataset.

EXAMPLE 4	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	output out=out a1 d11;
	run;

	symbol1 i=join v='star';
	symbol2 i=join v='circle';
	legend1 label=none value=('original' 'adjusted');
	
	proc gplot data=out;
	plot sales_A1 * date = 1
	sales_D11 * date = 2 / overlay legend=legend1;
	run;
	quit;



Example 5

Here the results from Example 3 are directed to HTML files using the SAS Output Delivery System (ODS).

EXAMPLE 5	
Census X-12 Spec File	**PROC X12 Code**
series{start=1972.07 data=(*...see datalines in example 1 ...*)}	data sales; input sales @@; date = intnx('month', '01jul72'd, _n_-1); format date monyy.; datalines; *...see datalines in example 1 ...* run;
	Ods html file="out.html" contents="out_index.html" frame="out_frame.html";
	proc x12 data=sales seasons=12 date=date;
	var sales;
transform{power=0}	transform power=0;
arima {model=(0,1,1) (0,1,1)}	arima model=((0,1,1) (0,1,1));
estimate { }	estimate;
x11{}	x11;
	run;
	ods html close;

Exact ARMA Maximum Likelihood Estimation					
For variable sales					
Parameter	Lag	Estimate	Standard Error	t Value	Pr > |t|
Nonseasonal MA	1	0.40181	0.07887	5.09	<.0001
Seasonal MA	12	0.55695	0.07626	7.30	<.0001

Table F 2: Summary Measures					
Table F 2.F: Relative Contribution of the components to the stationary portion of the variance in the original series					
For variable sales					
I	C	S	P	TD&H	Total
0.39	11.27	87.04	0.00	0.00	98.70

Table F 3: Monitoring and Quality Assessment Statistics			
All the measures below are in the range from 0 to 3 with an acceptance region from 0 to 1.			
For variable sales			
1.	The relative contribution of the irregular over three months span (from Table F 2.B).	M1=	0.038
2.	The relative contribution of the irregular component to the stationary portion of the variance (from Table F 2.F).	M2=	0.039
3.	The amount of month to month change in the irregular component as compared to the amount of month to month change in the trend-cycle (from Table F2.H).	M3=	0.000
4.	The amount of autocorrelation in the irregular as described by the average duration of run (Table F 2.D).	M4=	0.875
5.	The number of months it takes the change in the trend-cycle to surpass the amount of change in the irregular (from Table F 2.E).	M5=	0.268
6.	The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal (from Table F 2.H).	M6=	0.700
7.	The amount of moving seasonality present relative to the amount of stable seasonality (from Table F 2.I).	M7=	0.198
8.	The size of the fluctuations in the seasonal component throughout the whole series.	M8=	0.435
9.	The average linear movement in the seasonal component throughout the whole series.	M9=	0.352
10.	Same as 8, calculated for recent years only.	M10=	0.461
11.	Same as 9, calculated for recent years only.	M11=	0.414

****** ACCEPTED *** at the level 0.26***
****** Q (without M2) = 0.29 ACCEPTED.***

Conclusion

The X12 procedure of SAS/ETS software is an adaptation of the U.S. Bureau of the Census X-12-ARIMA Seasonal Adjustment program. The X12 procedure is fully incorporated into the SAS system. This incorporation permits the storage and the preparation of data for subsequent analysis and for the presentation of the analysis using high-quality graphics, customized reports, and applications.

Acknowledgments

SAS Institute, Inc. is thankful for the support of the U.S. Census Bureau for the assistance provided in the development of the X12 procedure. In particular, Brian Monsell and Catherine Hood have contributed greatly. This paper and its associated practical demonstration relied heavily on the contributions of Evan Anderson, Virginia Clark, and Mark Traccarella of SAS Institute Inc.

References

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis: Forecasting and Control*, Third Edition, Englewood Cliffs, NJ: Prentice Hall, Inc.

Dagum, E. B. (1988), *The X-11-ARIMA/88 Seasonal Adjustment Method: Foundations and User's Manual*, Ottawa: Statistics Canada.

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

Ladiray, D. and Quenneville B. (1999), "Understanding the X11 Method," Working Paper, Time Series Research and Analysis Centre, Statistics Canada.

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

SAS Institute Inc. (2000), *SAS/ETS User's Guide, Release 8.1,* Cary, NC: SAS Institute Inc.

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

U.S. Bureau of the Census (1999), *X-12-ARIMA Reference Manual*, U.S. Department of Commerce, Washington, DC, [ftp://ftp.census.gov/pub/ts/x12a/].

U.S. Bureau of the Census (1999), "X-12-ARIMA Seasonal Adjustment Program," [ftp://ftp.census.gov/pub/ts/x12a/].

U.S. Bureau of the Census (1999), *X-12-ARIMA Quick Reference for Unix*, U.S. Department of Commerce, Washington, DC, [ftp://ftp.census.gov/pub/ts/x12a/].

Appendix A – Cross Reference of PROC X12 and X-12-ARIMA Syntax

SAS (V. 8.2) STATEMENT	SAS OPTION	DESCRIPTION	CENSUS SPEC	CENSUS ARGUMENT
PROC X12		Mostly data specifications	series{}	
	DATA=	Should specify the input data set		data=
	DATE=	Date variable name		none equivalent
	START=	Date of 1st observation		start=
	SPAN=	(monyy,monyy) or ('yyQq','yyQq')		span=
	SEASONS=	4 for quarterly, 12 for monthly data		period=
	INTERVAL=	QTR or MONTH		period=
	NOPRINT	Suppress all printing	All specs	print=none
TRANSFORM		Transform or prior adjust series	transform{}	
	POWER=	Box-Cox power transformation parameter		power=
	FUNCTION=	Transformation specified by name: NONE, LOG, SQRT, INVERSE, LOGISTIC, AUTO		function=
IDENTIFY		Used to identify the ARIMA portion of the model using seasonal and nonseasonal differencing	identify{}	
	DIFF=	Orders of nonseasonal differencing		diff=
	SDIFF=	Orders of seasonal differencing		sdiff=
REGRESSION		reg information for regARIMA model	regression{}	
	PREDEFINED=	List of predefined regression variables: CONSTANT, LOM, LOMSTOCK, LOQ, LPYEAR, SEASONAL, TD, TDNOLPYEAR, TD1COEF, TD1NOLPYEAR		variables=
ARIMA		ARIMA modeling information	arima{}	
	MODEL= ((p d q)(P D Q)s)	Specify an ARIMA model (p d q)(P D Q)s using Box-Jenkins notation (if s is omitted, s=seasons)		model=
ESTIMATE		Estimates the regARIMA model specified by the regression and arima statements	estimate{}	
X11		Seasonal adjustment info	x11{}	
	MODE=	MULT, ADD, LOGADD, PSUEDOADD		mode=
	SEASONALMA=	Seasonal moving average used to estimate seasonal factors: S3X1, S3X3, S3X 5, S3X9, S3X15, STABLE, X11DEFAULT, MSR		seasonalma=
	TRENDMA=	Value for Henderson moving average		trendma=
	OUTFORECAST	Appends forecasts to tables A6, A8, A16, B1, D10, and D16		appendfcst=yes

FORECAST		Control forecast options	forecast{}	
	LEAD=	The number of periods ahead to forecast		maxlead=
VAR		SAS standard statement to specify the time series variables to be adjusted/forecast		
BY		SAS standard statement to specify variables used in By-Group processing	none equivalent	
ID		SAS standard statement to specify variables used for identification purposes only		
OUTPUT		Information for output datasets for time series		
	out=	SAS-data-set name		
	A1	Original series	series{}	save=(span)
	A6	regARIMA trading day component	regression{}	save=(tradingday)
	A8	regARIMA combined outlier component	regression{}	save=(outlier)
	B1	Prior adjusted or original series	x11{}	save=(adjoriginal)
	C17	Final weight for irregular components	x11{}	save=(irrwt)
	D8	Final unmodified S-I rations	x11{}	save=(unmodsi)
	D9	Final replacement values for extreme S-I rations	x11{}	save=(replacsi)
	D10	Final seasonal factors	x11{}	save=(seasonal)
	D10D	Final seasonal difference	x11{}	save=(seasonaldiff)
	D11	Final seasonally adjusted series	x11{}	save=(seasadj)
	D12	Final trend cycle	x11{}	save=(trend)
	D13	Final irregular series	x11{}	save=(irregular)
	D16	Combined adjustment factors	x11{}	save=(adjustfac)
	D16B	Final adjustment differences	x11{}	save=(adjdiff)
	D18	Combined calendar adjustment factors	x11{}	save=(calendar)
	E5	Percent changes in original series	x11{}	save=(origchanges)
	E6	Percent changes in final seasonally adjusted series	x11{}	save=(sachanges)
	E7	Differences in final trend cycle	x11{}	save=(trendchanges)
	MV1	Original series adjusted for missing value regressors	series{}	save=(missingvaladj)

Missing values are automatically imputed.