Experiences with User-Defined Regressors for Seasonal Adjustment

Catherine C. Harvill Hood[1] and Roxanne Feldpausch[1]
[1]Catherine Hood Consulting, Auburntown TN 37016

Abstract
The U.S. Census Bureau's X-12-ARIMA has built-in regressors for some U.S. holidays and several types of trading day effects. However, there are many other effects that are not built-in. User-defined regressors are useful to estimate moving holidays not included in X-12-ARIMA, custom Easter effects, and custom trading day effects.

Unfortunately, many users do not take the time to create their own regressors. The Census Bureau provides a program called GENHOL to help users generate their own holiday effects; however, this program can be daunting to users who prefer point-and-click software interfaces. We have written a user-friendly program that assists in defining regressors by including many built-in holidays, a flexible number of days before and after the holiday, and different kinds of one-variable trading day effects. We demonstrate the value of seasonal adjustments that include custom holiday and trading day regressors.

Key Words: X-12-ARIMA, trading day effects, moving holiday effects

1. Background

Trading day effects and holiday effects are an important part of seasonal adjustment. The trading day effect is the effect that comes from the weekday composition of the month. Holiday effects are the effects from moving holidays. Holidays that are at the same time every year are easily estimated, but holidays that move between months, such as Easter or Chinese New Year, require special consideration. A successful seasonal adjustment often depends on successful estimation of these effects.

X-12-ARIMA estimates both trading day effects and moving holiday effects with regressors. For monthly or quarterly series, there are several types of trading day regressors, but two of the regressors that are the most common for flow (noninventory) series are a six-coefficient effect and a one-coefficient effect. The six-coefficient flow trading day effect estimates a separate regression coefficient for six days of the week and an implied coefficient for Sunday. The one-coefficient flow trading day effect treats the weekdays (Monday through Friday) as all the same and estimates an implied coefficient for the weekend.

The holidays available in X-12-ARIMA include Easter (as celebrated by the Western Christian Church), U.S. Labor Day, and U.S. Thanksgiving. The Easter effect assumes that a certain number of days before Easter are different in some way. The U.S. Labor Day effect is similar. The U.S. Thanksgiving effect assumes that a certain number of days

after Thanksgiving are different because Thanksgiving is the beginning of the Christmas shopping season in the U.S.

If the user is interested in other regression effects that are not built-in, X-12-ARIMA provides commands within the regression spec for "user-defined" regressors. We give examples of the code used in X-12-ARIMA and the regressors used for some user-defined effects later in the paper. For more information on the specific options in X-12-ARIMA, please see the *X-12-ARIMA Reference Manual* (U.S. Census Bureau, 2010).

Setting up the commands and the external data file with the regressors can be a bit tedious. The Census Bureau provides a program external to X-12-ARIMA to help set up the regressors. We wanted a more flexible program that would set up the regression effects we needed easily and quickly.

2. External Software for User-defined Regressors

2.1 The GENHOL Program
The Census Bureau provides a program called GENHOL that will create holiday regressors. It is quite flexible and can create regression effects in intervals before and after holidays. It can also center the regressors using either their overall means or their calendar-month means. It creates both the regression file and a file with the X-12-ARIMA commands that can be pasted into the X-12-ARIMA specification file. The dates of Easter are on the Census website, but for other holidays, the user has to supply the dates for the holidays.

The commands for the GENHOL program are similar to the commands for X-12-ARIMA. There are spec names with open and closed brackets and certain arguments that the user can specify. When the input text files are set, the program is run from a command prompt. For details on the syntax required for running GENHOL, please see Monsell (2009).

Though the learning curve for this program can be quite short for users familiar with the command prompt and with X-12-ARIMA, the program can have a steeper learning curve for users unfamiliar with programming and running command-line programs.

2.2 The EffectGen Excel Program
Our goal in writing the EffectGen program was to provide the same capabilities as the GENHOL program but with a point-and-click interface. We also wanted to add additional capabilities related to additional holidays and to different trading day variables. Like the GENHOL program, the EffectGen program writes the regressors and the commands for X-12-ARIMA, but instead of writing them to separate text files, writes them to new worksheets in the Excel file.

For Easter effects, we wanted to be able to generate effects similar to the effects already in the GENHOL program. Figure 1 below shows the interface for the Easter-related holidays. It is simple to specify the starting and ending dates and to specify the length of the interval before and after the holiday. The "name of the effect" tells the program what labels to use for the regressors in X-12-ARIMA. Besides an Easter effect, we have also added the ability to add regressors for the holidays related to Easter, such as Good Friday,

Easter Monday, Ash Wednesday, and Whit Monday. The user can choose between holidays using the drop-down box on the screen.

Figure 1: The screen to generate effects for Easter and holidays related to Easter

Another useful feature of the EffectGen program is the ability to generate one-coefficient trading day effects that are different from the built-in effects of X-12-ARIMA. As shown in Figure 2 below, the user can select the days of the week to group together, the starting and ending dates, and the name of the effect.

Figure 2: The screen to generate one-coefficient trading day effects

3. Results

To demonstrate the value of the user-defined regressors, we modeled some series using X-12-ARIMA's built-in regressors and our newly created user-defined regressors. We looked at model diagnostics to verify that the new regressors were helping the model.

3.1 Custom Easter and Easter-related Holiday Effects

To investigate alternative Easter regressors and Easter-related holidays, we looked at the French Industrial Production series from the book *Seasonal Adjustment with the X-11 Method* (Ladiray and Quenneville, 2001). European series demonstrate holidays related to Easter that are not commonly celebrated in the U.S.

We asked the EffectGen program for regressors for Easter, Ash Wednesday, and Whit Monday. As a starting place, we looked at effects for eight days before Easter, four days before Ash Wednesday, and two days before Whit Monday. Easter falls in either March or April. Ash Wednesday is 40 days before Easter, so it falls in either February or March. Whit Monday is 50 days after Easter, so it falls in either May or June. The output from the program gives us both the commands we need to run the regressors in X-12-ARIMA (shown in Figure 3 below) and the text file that contains the regressors (named HolFIP.txt for this example and shown in Figure 4).

```
regression{
  variables=(td AO1987.Jan LS1992.Nov)
  user=(Easter Ash Whit)
  file='HolFIP.txt'
  format='datevalue'
  start=1986.1
  usertype=holiday
}
```

Figure 3: Example X-12-ARIMA code for the French Industrial Production series that includes user-defined regressors for Easter, Ash Wednesday, and Whit Monday

```
1986  1   0          0          0
1986  2   0          0.0857     0
1986  3   0.5929    -0.0857     0
1986  4  -0.5929     0          0
1986  5   0          0          0.3571
1986  6   0          0         -0.3571
1986  7   0          0          0
.
.
.
```

Figure 4: The first rows of the file HolFIP.txt that contains the regressors for Easter, Ash Wednesday, and Whit Monday

To evaluate the results, we looked at the significance of the individual regression effects and compared the AIC and other likelihood-based tests. The regression results are in Table 1. They show us that the Easter effect is significant (a fact already shown in Ladiray and Quenneville, 2001); however, the Ash Wednesday effect is also significant. The Whit Monday effect is not as strong and is not significant.

Table 1: Regression Results for the User-defined Holidays

Variable	Parameter Estimate	Standard Error	t-value
Easter	-3.4745	0.70373	-4.94
Ash	-1.9490	0.83689	-2.33
Whit	-1.0780	0.65474	-1.65

We can also use likelihood statistics to see if the Ash Wednesday holiday is useful or not. We ran the series with only the Easter holiday, and then ran the series again with both Easter and Ash Wednesday holidays. For likelihood statistics we are, of course, looking for smaller values. All of the statistics show a preference for the model with the Ash Wednesday regressor.

Table 2: Likelihood Statistics for Models with and without the User-defined Ash Wednesday Holiday

Likelihood Statistics	With Easter Only	With both Easter and Ash Wednesday
AIC	402.9333	399.4635
AICC	407.2666	404.5237
Hannan Quinn	416.5257	414.1014
BIC	436.5379	435.6530

X-12-ARIMA also provides a Chi-square test for the significance of groups of regressors which would allow us to test the holidays as a group. For series that are long enough, it would also be possible to use the history diagnostics in X-12-ARIMA to look at AIC differences and forecast error differences over time. For more information on model diagnostics, please see Findley et al (1998, particularly Section 5) and Findley and Hood (2000).

3.2 One-coefficient Trading Day Effects
We used customized one-coefficient trading day regressors quite effectively with some proprietary sales data from a company with only five years of data from their stores in Europe. The company expected a trading day effect with higher sales on Friday and Monday. However, the built-in six-coefficient regressors were not significant, and the built-in one-coefficient regressor was not appropriate.

As a proxy for these series for the purposes of this paper, we will demonstrate the one-coefficient trading day regressors on a U.S. retail sales series from the Census Bureau website, shortened to begin in 2006. Most of the retail sales series have very similar trading day patterns with more sales on Friday and Saturday.

The regression results, found in Table 3 below, show us that the trading day effect is significant. The likelihood statistics, found in Table 4, show a preference for the trading day regressors also.

Table 3: Regression Results for the One-coefficient Trading Day Effect

Variable	Parameter Estimate	Standard Error	t-value
Fri-Sat TD	0.0211	0.00571	3.70

Table 4: Likelihood Statistics for Models with and without the Trading Day Effect

Likelihood Statistics	Without Trading Day	With Trading Day
AIC	738.90	733.42
AICC	739.76	734.72
Hannan Quinn	741.90	737.16
BIC	746.71	743.17

4. Conclusion

User-defined regressors in X-12-ARIMA are useful to provide the best possible seasonal adjustment for many series. The software we have developed helps us use these regressions easily and quickly.

We have planned several additions to the software, including:
- A Chinese New Year component that would be very similar to the Easter regressor,
- Country-specific Ramadan-related holidays, and
- Working day adjustments to treat holidays like a weekend day. For example, a holiday like the U.S. Memorial Day would not be a typical Monday for the purpose of trading day estimation.

We have also found that we can adjust for holidays in series that are collected in periods of four weeks or five weeks instead of on a monthly basis. For example, a holiday like the U.S. Independence day on July 4 has the same effect every year in a monthly series, but the July 4 effect can change from year to year in a series where the first day of the period always falls on a Sunday, for example. Therefore, we also hope to find the best way to implement this kind of regressor into the EffectGen software.

MS-Excel is a trademark of Microsoft Corporation.

References

Findley, D.F. and C.C.H. Hood (2000). "X-12-ARIMA and Its Application to Some Italian Indicator Series." In *Seasonal Adjustment Procedures – Experiences and Perspectives*. Rome, Italy: Istituto Nazionale di Statistica. 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). In *Journal of Business and Economic Statistics*. Alexandria, VA: American Statistical Association. 16: 127-176.

Ladiray, D. and B. Quenneville (2001). *Seasonal Adjustment with the X-11 Method*. Springer-Verlag: New York, NY.

Monsell, B.C. (2009). "GENHOL--A Utility that Generates User-defined Moving Holiday Regressors for X-12-ARIMA." Washington DC: U.S. Census Bureau, U.S. Department of Commerce.

U.S. Census Bureau (2010). *X-12-ARIMA Reference Manual, Version 0.3*. Washington DC: U.S. Census Bureau, U.S. Department of Commerce.