Methodology: Seasonal Adjustment of the Transportation Services Index Time Series Data

The data underlying the Transportation Services Index (TSI) are seasonally adjusted by the Bureau of Transportation Statistics (BTS) using the methodology discussed here. This seasonal adjustment enables consistent comparisons of data between time periods.

Origin

To seasonally adjust transportation services time-series data, BTS adopted X12-ARIMA, Release 0.2, which was created by the U.S. Department of Commerce, U.S. Census Bureau. X12-ARIMA grew out of X11-ARIMA, which in turn originated from Census X-11 software developed at the Census Bureau in the 1950s and 1960s. The basic approach of X-11 is undoubtedly the most widely used statistical method for seasonal adjustment.[1]

Isolating Seasonality

The basic model of this approach is to decompose the time series into three components: trend (including cyclic phenomena), seasonality, and irregular. By a series of iterative steps, the seasonality component is eventually isolated and removed from the original data series. In applying this methodology to the transportation services time-series data, we found that each element of the TSI—rail (passenger and freight), pipeline (petroleum and natural gas), transit, waterborne, trucking, and aviation (passenger and freight)—displays strong seasonality patterns. However, in some series the seasonality was less pronounced than in others. For example, transit seasonality must be isolated from a background of considerable fluctuations or noise. When there is a great deal of fluctuation in the data, as in this instance, the seasonality component is much smaller relative to the other two components (trend and irregular).

Seasonal, Trading-Day, and Holiday Effects

Seasonal

The seasonal effect in a time series is any effect that is reasonably stable in terms of annual timing, direction, and magnitude. Seasonal adjustment is the process of estimating and removing the seasonal effects from a time series. Because the seasonal effects can disguise important features of economic series such as direction, turning points, and consistency between other economic indicators, seasonal adjustment can also be thought of as focused noise reduction (Hood, 2009; Ashley, 2001).

Trading-Day

In addition to seasonal effects, monthly time series that are totals of daily economic activities are frequently influenced by the weekday composition of the month (Findley et al, 1998). Trading-day effects reflect the number of days in the month and the number of times each day of the week occurs in the month, which can affect the monthly totals of output services. Recurring effects associated with weekday composition in monthly (or quarterly) economic time series are called trading-day effects.

Holiday

Certain kinds of transportation services and their associated time series, such as aviation (passenger and freight), are affected significantly by holidays. Effects from holidays, such as Christmas, that always occur on the same date of a month each year are seasonal components of a time series. Effects associated with holidays that are not always on the same date of a month, such as Labor Day, Thanksgiving, and Easter, are called moving holiday effects.

Both trading day and moving holiday effects are persistent and predictable calendar-related effects. These calendar effects also need to be removed from the seasonally adjusted series. Not surprisingly, trading-day and moving holiday effects influence many of the time series used as components in the TSI.

Outlier Adjustments

In all of the transportation services time series it was necessary to make adjustments to the defaults offered in the software. Each of the series had too many extreme observations and outliers, when using the default lower sigma limit of 1.5 to estimate the seasonal component. Because the basic tool of both the X-11 and X-12 ARIMA are moving averages, which are linear operators, they respond dramatically to the presence of extreme observations (Ladiray and Quenneville, 2001). When doing seasonal adjustment, too many outliers may interfere with estimation of the seasonal component and make seasonal adjustment unstable. We had as our objective to ensure that no more than one out of eight data observations were of the extreme variety. The criterion of one out of eight comes from best practices in the field. For the most part we met that objective.

Two Kinds of Outliers: Extreme Values and Worst Outliers

Extreme Values

We limited the number of extreme values by increasing the lower sigma limit to 1.8. This is justified with the transportation services time series data because the 10 transportation time-series data sets are strongly impacted by the vicissitudes of weather patterns as well as a host of unknown variables, making transportation time-series data inherently variable. Our experience has shown that nearly all time series need a similar adjustment to reduce extreme values.

Worst Outliers

Worst outliers were detected by the X-12 program using a default critical value. An example of an extreme outlier that strongly impacted many of the time-series data is the terrorist attacks of September 11, 2001. The impact of 9/11 was particularly powerful on aviation passenger and freight outputs. A few extreme outliers found by the program series are noted in the data summaries that follow.

Worst outliers—whether found by the X-12 program or other sources—are adjusted out of the data series when estimating the seasonal component so that they do not affect the estimate. These outliers and extreme values are included with the irregular component. Because the seasonally adjusted series is the trend and irregular components together, all outliers and extreme values are included in the seasonally adjusted series.

Modeling the Time Series: Multiplicative or Additive?

A series is modeled in one of two forms: multiplicative or additive. In practice, when the peaks and troughs remain constant as the trend increases or decreases, the additive model is used. When the peaks and troughs expand as the trend increases or decreases, the multiplicative model will fit the time series better. In the July 2009 revision of seasonality adjustment for the TSI, we determined the best model using the Census Bureau's X12-ARIMA monthly seasonal adjustment method. Seasonal adjustments for rail freight intermodal, natural gas, and petroleum had used the multiplicative model. An additive model was chosen for the remaining seven series: aviation freight, aviation passenger, trucking, transit, rail freight carloads, rail passenger, and waterborne.

Moving Seasonality

Adjustments were made to the seasonally adjusted data based on several criteria developed to assess the basic model, which are given as measures M1 through M11. These criteria involve the way the trend line is estimated, the amount and type of variation of the irregular component, and the stability of the model, especially indications of moving seasonality. The latter occurs when the high and low months and undulations in the data are shifting, thereby introducing additional variability. Although moving seasonality often presents when stable seasonality exists in a time series, it should cause concern only when it becomes so prominent that it makes the detection of the stable seasonality difficult or even impossible.

Moving seasonality is present in two TSI time series: transit and aviation passenger.

Meeting Criterion

Failing a single criterion by no means implies that the model is not good, just as passing all the criteria does not ensure a good adjustment. These measures only serve as aids to guide the adjustment. A criterion is met if it is less than 1.0; failure means that the measure was equal to or greater than 1.0. Gradually, with experience, we have learned to put the most emphasis on one measure, M7, which determines if stable seasonality is sufficiently greater than moving seasonality. A failure in the other M measures can indicate a weak trend line, high autocorrelation shifting seasonality, or other problem. Although these other commonly computed measures may fail for some of the TSI series, they do not act as show-stoppers. By using M7 as the dominant criterion, all the series used in the TSI display strong seasonality components.

Standardizing Each Series

The nearly dozen time series used to create the TSI had varying amounts of historical data. Some series began as early as 1973. Earlier data may be useful for historical purposes, but are no help in seasonally adjusting data in recent years. Why? The answer is that the X12-ARIMA program doesn't allow the earlier data to carry much weight for the years at the end of the series. For this reason, with the exception of aviation passenger, all time series used in the current (2009) TSI begin at January 1998. This is more than sufficient time to obtain a good seasonal adjustment

Waterborne

Comparing 12-month periods in the series, it is clear that January and February are the low points. March through August, October, and November are the peak months, with July (slightly) the highest point. No substantial amount of moving seasonality was detected. Extreme outliers include December 2000, September 2005, and March 2008. March 2008 was detected as a level shift, which decreased all observations from that time point onward by about 17 percent in the seasonally adjusted series.

An additive model was used to fit the series. No significant trading day and moving holiday effects were detected in this data series.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points for use in the final model.

Rail

Freight. Two time series constitute rail freight: carloads and intermodals. For both time series, we find a double peak in August and October with October as the larger of the two. February is reliably the low month. The moving holiday effects were not significant in either series with the exception of Easter in intermodals. We haven't seen extreme outliers in carloads since January 1998. For intermodals, October 2002, June 2004, and February and August 2006 were extreme outliers as was December 2008. Trading-day effects were a significant factor in both carloads and in intermodals in the June 2007 update. Rail freight is very seasonal; the adjustment in both series reveals clear trend lines and stable seasonality.

No significant moving seasonality was detected in either series. Both series receive a "pass" on the M7 diagnostic statistic.

Each of the rail time series was first seasonally adjusted and then combined into a rail output index. An additive model was used to fit carloads, while a multiplicative model was used to adjust intermodals time series.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points in either series (carloads and intermodal) for use in the model.

Passenger. Like freight data, rail passenger data are also well behaved with a clear trend and consistent and stable seasonality. Rail passenger adjusts similarly to rail freight-intermodals. Trading-day effects were not significant, but Easter week, Labor Day week, and Thanksgiving were determined significant factors in rail passenger. No significant moving seasonality was detected. The high points of the year are July and August; the low months were in the winter with February most often the lowest. The only extreme outlier found in the series was January 2003, which was detected as a level shift and which increased all observations from that time point onward by about 10 percent in the seasonally adjusted series. Like rail freight, rail passenger easily passed the diagnostic standard M7. Rail passenger data are best modeled with the additive model.

NOTE: For the November 2010 TSI, it was necessary to forecast one data point for use in the rail passenger model.

Pipeline

Petroleum. Petroleum and petroleum products monthly amounts show a distinct pattern of February as consistently the low month. May, July, August, October, December, and January are the peak months, with August slightly more often the annual high month. Significant effects for trading days were detected, but no significant moving holiday effects were presented. The number of outliers was reduced by adjusting the weighting as in the other series. In January 1985, a new upward trend occurred. When seasonality is removed, the monthly increase (i.e., from December 1984 to January 1985) was 37 percent from the previous month. This level is sustained over the following months, never to return to its previous level. This change in the trend was one reason why we start the series a little later at January 1990. Extreme outliers include March 1991, February 1992, October 1994, and February 2000. February occurs as a moderate and extreme outlier quite frequently, indicating that the model is compromised to some degree by the relatively extreme results in February.

The criterion for each of the diagnostic measures, M1 through M11, was easily met for the Petroleum series. No significant amount of moving seasonality was detected. Multiplicative models were used for both series of pipeline.

NOTE: For the November 2010 TSI, it was necessary to forecast one data point for use in the model.

Natural gas. The peaks and low months in natural gas complement the petroleum time series. January, February, March, and December are the high gas consumption months each year. Gas consumption is relatively low in May, June, September, and October. The final adjustment included a change in the weighting to manage the number of extreme values. Natural gas data series is not impacted by either trading days or moving holidays. Two extreme outliers were January 2006 and February 2007. The criterion for diagnostic measure M7 of pipeline gas was met. No significant moving seasonality was detected in this time series.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points for use in the model.

Trucking

October appears to be the most frequent peak month in the calendar year. Other high points are August, March, and June. The low points occur during the late fall and winter months from November through February; February and December are most often the two lowest months. The seasonality aspect is very strong and consistent throughout the decades. No evidence of significant moving seasonality was detected. Only one extreme outlier was found—January 2000, which was a level shift.

Currently, trading-day effects affect trucking. In the past, Easter week and Thanksgiving have been significant factors, but these holidays don't appear to have any effect now. The deseasonality conducted on the trucking data met the diagnostic measure M7. Since the November 2004 TSI, the model for trucking was changed to an additive model. The multiplicative model was more valid in the past because of the lower fluctuations of the series apparent during the earlier years. As the series has progressed over time, the fluctuations are becoming approximately constant as the trend increases.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points for use in the model.

Transit

An examination of the raw data going back to 1979 shows evidence of seasonality. The lowest month in ridership is February; the highest month is October. The pattern of seasonality is statistically significant, but the frequent shifts in the troughs and peaks make the seasonal adjustment a real challenge. Although the series is thus highly erratic, there is a little more stability in recent data. For this reason, we decided to work with just the last 13 years.

The day of the week plays a strong role, particularly Thursday through Sunday, with Sunday having an especially large impact. Surprisingly, most moving holiday effects were not worth hard-coding because the day of the week absorbed most of the moving holiday effects. The effect of Easter holiday was highly significant in the transit series. Significant moving seasonality (1 percent level) was detected in the series. The diagnostic measure M7 was met. Two extreme outliers were found in June: 1994 and 1998. The other two extreme outliers were December 1993 and April 1995. An additive model was used to fit transit time series data.

NOTE: The most recent data available for transit is September 2010. To make up for the data gap, it was necessary to forecast two data points for the November 2010 TSI. Normally, Transit lags well behind the other series in up-to-date data. For example, in the December 2004 update, 8 years of transit data were revised.

Aviation

Freight. The seasonality pattern in aviation freight shows high points most often in October and low points in February. From early 1990 to late 2008, the general trend for tons carried was increasing. Aviation freight is strongly influenced by trading-day effects. The moving holidays of Easter and Thanksgiving were also found to be significant factors. A very significant outlier was September 2001, which is to be expected after the 9/11 catastrophe. The value for September 2001 dropped 12 percent from the previous month in the seasonally adjusted series. Other extreme outliers include October 2002, February 2004, and November 2008. February 2004 was detected as a level shift, which increased all observations from that time point onward by about 11 percent after adjusting for seasonality. Whereas the November 2008 was detected as another level shift, which decreased all observations from that time point onward by 12 percent in the seasonally adjusted series. Aviation freight easily passed the diagnostic standard M7. An additive model was used to model the Aviation Freight data series.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points for use in the model.

Passenger. July and August are the highest months for passenger data. February is consistently the low month. The seasonality of aviation passenger traffic is very evident. All three moving holiday Easter, Labor Day, and Thanksgiving had significant effects on aviation passenger series.

Like aviation freight, aviation passenger was impacted strongly by 9/11, where there was a 28.5 percent drop after seasonality was adjusted. In fact, much of 2001 was unusual with outliers from September through November, and where the October outlier was a level shift. The only other extreme outlier was December 2002. No other extreme outliers were uncovered. Aviation passenger data easily passed the diagnostic standard M7. An additive model was used to model the aviation passenger data series.

Moving seasonality was detected at the 1 percent level. Although it was also detected in the transit data series, it is never as consistent as in airline passenger.

NOTE: For the November 2010 TSI, it was not necessary to forecast any data points to determine the seasonal factors.

[1] For a good overview of the X-11 method, see D. Ladiray and B. Quenneville. *Seasonal Adjustment with the X-11 Method*. 2001. New York: Springer-Verlag.Ashley, J. D. 2001. Why Seasonal Adjustment – Draft. Washington, D.C.: Bureau of the Census. Available online at http://www.catherinechhood.net/WhySeasAdj.pdf.

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B. 1998. New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program. Washington, D.C.: Bureau of the Census. Available online at http://www.census.gov/ts/papers/jbes98.pdf.

Hood, C. 2009. Seasonal Adjustment and Time Series FAQ. Nashville, TN: Catherine Hood Consulting. Available online at http://www.catherinechhood.net/safaqmain.html.

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

Transportation Service Index Composition

This table gives basic information for the time series data used to compose the Transportation Services Index (TSI). Seven of the ten series have data to November 2010. The remaining three series had missing data that had to be forecasted (see column 3).

Mode	Moving seasonality present?	Data points forecasted	Model[1]	Trading days[2] and holidays[3]
Rail				
Passenger	No	1	A	No TD; all moving holidays
Rail Freight				
Carloads	No	None	A	TD; No holidays
Intermodals	No	None	M	No TD; Easter
Trucking	No	None	A	TD; No holidays
Waterborne	No	None	A	No TD; No holidays
Transit	Yes, at 1% level	2	A	TD; Easter
Aviation				
Freight	No	None	A	TD; Easter & Thanksgiving
Passenger	Yes, at 1% level	None	A	No TD; all moving holidays
Pipeline				
Natural gas	No	None	M	No TD; No holidays
Petroleum	No	1	M	TD; No holidays

1 M = Multiplicative; A = Additive

2 TD = Trading Days

3 (Moving) Holidays =Easter, Labor Day, & Thanksgiving. Memorial Day is not available.