An Empirical Comparison of Methods for Benchmarking Seasonally Adjusted Series to Annual Totals

Catherine C. Harvill Hood
HENDYPLAN Luxembourg, 41 Avenue de la Gare, L-1611 Luxembourg, catherine.hood@hendyplan.com

Keywords: Benchmark, seasonal adjustment

1. Summary

For benchmarking monthly and quarterly series to annual series and to the Economic Census every five years, the U.S. Census Bureau uses an iterative, nonlinear method known as the Causey-Trager method. However, the Census Bureau's X−12−ARIMA seasonal adjustment program uses a modified Denton procedure to benchmark the seasonally adjusted series to the annual totals of the unadjusted series. Some users have requested a different benchmark method in X−12−ARIMA.

Statistics Canada has proposed several different benchmark methods, including a regression procedure, to replace the method for benchmarking in X−12−ARIMA. Using a sample of U.S. time series, this paper investigates the properties of the benchmarks from the current procedure in X−12−ARIMA, the new methods proposed by Statistics Canada, and the Causey-Trager method for benchmarking the seasonally adjusted series to the annual totals of the raw data. The objective of this study is two-fold: 1) to look at some of the properties of the various benchmarking procedures under consideration for benchmarking seasonally adjusted series, and 2) to look at possible settings for the regression procedure from Statistics Canada.

There were very consistent results with smooth benchmark factors and small discrepancies in the month-to-month changes with both the Causey-Trager method and some settings of the Regression method. The Causey-Trager method gave results that were more consistent for every month though there is still a problem with the distortion of the month-to-month percent changes at the beginning and ending of the year. The Regression method gave results with smaller revisions when new data was added.

2. Background

Benchmarking procedures can be used to solve a variety of problems. There are many examples where monthly and quarterly series are benchmarked to annual series. Benchmarking procedures can also be used to adjust regional totals to country totals. Also, the U.S. Census Bureau used benchmarking procedures when changing from one industrial classification code another, benchmarking the past values of the new series to the past annual totals from the previous series.

Some users of seasonally adjusted data prefer that the annual totals for the seasonally adjusted series to match the annual totals for the original, not-seasonally-adjusted series. While this practice is common in many agencies, it is not common at the Census Bureau.

The annual totals of the unadjusted series and the seasonally adjusted series will be equal only when the series is adjusted additively, the seasonal pattern is fixed from one year to the next, and there are no trading day adjustments. The U.S. Census Bureau, in the documentation for X−12−ARIMA, recommends against benchmarking for seasonally adjusted series: "Forcing the seasonal adjustment totals to be the same as the original series annual totals can degrade the quality of the seasonal adjustment, especially when the seasonal pattern is undergoing change. It is not natural if trading day adjustment is performed because the aggregate trading day effect over a year is variable and moderately different from zero." (U.S. Census Bureau, 2002, p. 161) However, there are users who need the totals to match, and for these users, X−12−ARIMA does have the capability to force the annual totals of the not-seasonally-adjusted series and the adjusted series to match.

Unfortunately, forcing a series of numbers to sum to a certain total does not provide a unique solution. There must be some characteristic of the original, unbenchmarked series that we can use as a constraint. For this study, I ran the software with the constraint that the month-to-month percent changes are preserved as closely as possible. Therefore, the goal will be to produce a benchmarked time series where the percent changes between the months are preserved as much as possible while the annual total is changed to a new, desired annual total.

3. Description of Programs

All programs are written in FORTRAN. I used already existing interfaces or developed interfaces in Visual Basic to run the programs and simplify use.

3.1. Modified Denton

The method currently in X−12−ARIMA to force the annual totals of the original series and the seasonally adjusted series to match is the method inherited from

---

This report includes research done at the U.S. Census Bureau and at HENDYPLAN, SA. The views in this paper are those of the author and not necessarily those of the U.S. Census Bureau or of HENDYPLAN, SA.

X–11–ARIMA, due to Huot (1975) and Cholette (1978) and based on earlier work by Denton (1971). To ask for the seasonally adjusted series benchmarked to the original series annual totals, the command is "force=totals" in the x11 spec. The details of the procedure are in Ladiray and Quenneville (2001).

3.2. U.S. Census Bureau's Causey-Trager Method

Causey (1981) developed a numerical algorithm that was later revised by Trager (1982), showing that there was an iterative solution to the problem of minimizing the changes to the month-to-month changes subject to a set of constraints (the benchmarks). The procedure minimizes the changes iteratively using steepest descent. Though the first solutions were not published, their notes are available as appendixes in a research report by Bozik and Otto (1988). This method is referred to generally as the Causey-Trager method.

The numerical algorithm was eventually programmed into FORTRAN, and the U.S. Census Bureau still supports and maintains the FORTRAN code. The Census Bureau has used the program for production of benchmarked numbers since the early 1980s. The program was not designed for benchmarking the seasonally adjusted series to annual totals from the original series since the Census Bureau does not do this type of benchmark.

3.3 Statistics Canada's Regression Method

Based on earlier work by Cholette and Dagum (1994), Statistics Canada has developed a regression-based benchmarking model. There are two main parameters for the user to set in the procedure: 1) λ, a parameter that relates to the weights in the regression equation, and 2) ρ, the value of the AR(1) parameter (set between 0 and 1). A special case of this method is the Denton method when ρ =1. Details of the procedure are in Quenneville, Cholette, Huot, Chiu, and Di Fonzo (2004).

For this research, I ran the program with λ values of 0, 0.5, and 1, as suggested in the documentation. With all three λ values, I ran ρ values of 0.8, 0.85, 0.9 and 0.95, based on the documentation which suggests values between 0.8 and 1.

3.4. Statistics Canada's Henderson-Filter Smoothing Method

This procedure uses Henderson Filters to smooth the gap between the annual totals of the original series and the seasonally adjusted series. The procedure iterates until the totals are within a set limit. The length of the Henderson filter can be fixed, or the procedure will find the best length automatically. For more information, see Quenneville, Huot, and Chiu (2001).

4. Methods

In this paper, I have assumed that the year is from January to December. All the procedures described below have options to allow you to use another month as the starting month, if you wish.

For the initial study, I used 30 time series from the U.S. Census Bureau including Import/Export series, Retail Sales series, and Construction series. For the second part of the study, I dropped the Henderson method, limited the number of parameters tested for the Regression method, and ran an additional 150 series from the U.S. Census Bureau. I used X–12–ARIMA to seasonally adjust all the series to obtain the same unbenchmarked series for every method.

I compared the levels of the benchmarked series and the benchmark factors for the various methods. For each method, I also compared the month-to-month changes of the benchmarked series to the unbenchmarked series. A common complaint against many benchmark methods is that month-to-month changes are distorted (particularly December to January for monthly series), so I also looked at month-to-month changes by month. I also looked at graphs both with Excel and with R.

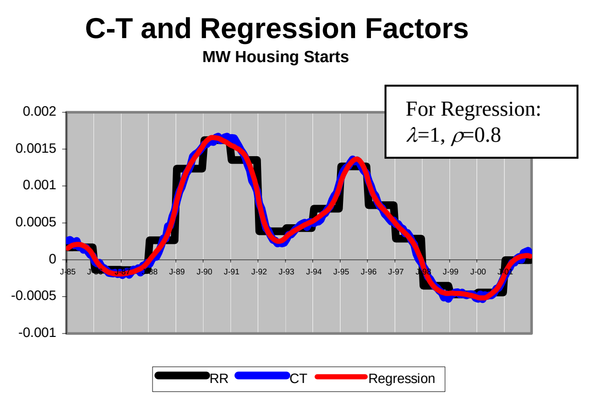
5. Results

5.1 Results from Initial Study (30 Series)

We found we had the best results for Regression and Causey-Trager methods. The Henderson method was very computer intensive and did not always give an answer. For our purposes, we found that the best settings for the regression method were with λ=1 and ρ > 0.8.

The graphs showed the similarity of the methods, with differences between the various factors often on the level of rounding error.

Figure 1. Example of Causey-Trager and Regression Benchmark Factors



Along with the benchmark factors in Figure 1, I show also the ratio between the annual total for the original series and the seasonally adjusted series, shown as the step function in black (RR). This ratio shows by how much the series would need to be changed and is a good guide for looking at the benchmark factors.

There were some methods that had smoother benchmark factors, which led to a smoothing of the differences between the unbenchmarked and benchmarked series. It was generally true that the Regression method with settings of λ = 1 were closer to the Causey-Trager benchmarks than with other settings. For other values of λ, the Regression benchmarks were closer to the other methods. The benchmarks from the Causey-Trager method and the Regression method with λ=1 gave very consistent, smooth results for all the series.

Other than graphs, the main diagnostic reviewed was the average absolute difference of the benchmarked series from the unbenchmarked series. In no case was the difference significant from zero.

Table 1. Average Absolute Differences (AAD) of the Month-to-Month Percent Changes for All Methods

Method	AAD
Modified Denton	0.00074
Causey-Trager	0.00042
Henderson Filter	0.00054
Regression	
λ=1, ρ=0.8	0.00038
λ=1, ρ=0.95	0.00037
λ=0.5, ρ=0.8	0.00038
λ=0.5, ρ=0.95	0.00042
λ=0, ρ=0.8	0.00100
λ=0, ρ=0.95	0.00078

Based on the results with only 30 series, and because of the problems with the Henderson method, for the next set of series, I focused attention on the Causey-Trager and Regression methods for the rest of the study, though I still ran the Modified Denton method for comparison purposes. With the overall results for the methods being so similar, I also wanted to focus on AADs by month and on the beginning and ending of the series.

5.2 Results for the AAD by Month

For all 180 series, I ran the modified Denton, C-T, and Regression method with λ=1, ρ=0.95. A particular issue with the Denton method is a possible distortion of the month-to-month percent changes at the breaks between the years, in our case December to January. To investigate I looked at the ADDs by month.

For all methods, the ADDs were larger for November through February than for March through October. Average AADs for the groups of months are given in Table 2 below. There does appear to be less of a difference between the months for the Causey-Trager method.

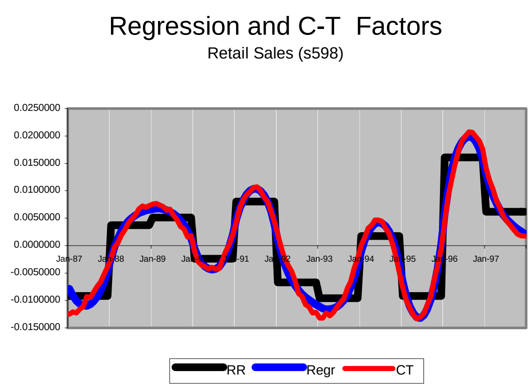
Table 2. Average Absolute Differences for Month-to-Month Percent Changes, Grouped by Months

	Nov to Feb	Mar to Oct
Regression	0.000642	0.000278
C-T	0.000599	0.000350
Mod. Denton	0.001443	0.000919

5.3 Results for the Beginning and Ending of the Series

For some series, graphs showed a difference between the Regression method and the other methods at the beginning and ending of the series. To investigate the ending of the series and possible revisions, we ran the benchmark programs for the series without the final year and then again for the full series. In this way, we could look at the revisions that could occur at the end of the series as new data is added.

Figure 2. Example Factors with Differences at the Beginning of the Series



For most methods, including the Denton method and the Causey-Trager method, the last benchmark factor for the last complete year is applied to the new data at the end of the series, assuming that the best forecast for next year's benchmark factor is the last year's factor. This might not be an unreasonable assumption in some uses of benchmark methods. For example, in a case where you have trouble estimating the number of new businesses for the monthly survey, you might have certain expectations on the next benchmark factor. However, Quenneville (2004) proposes that the best forecast of the benchmark for seasonally adjusted data is zero.

Even for series with small factors at the end of the series, there are smaller revisions for the regression method, as seen in Figure 3 below. Figure 4 shows a series with larger benchmark factors at the end of the series. The Causey-Trager method is shown in green and the Regression method in blue with the full-series factors shown by smooth lines.

Figure 3. Example Revisions for Causey-Trager and Regression Benchmark Factors

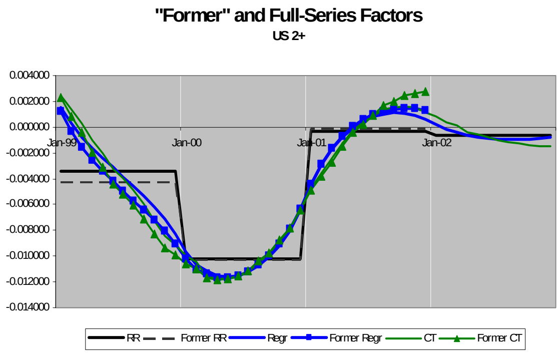
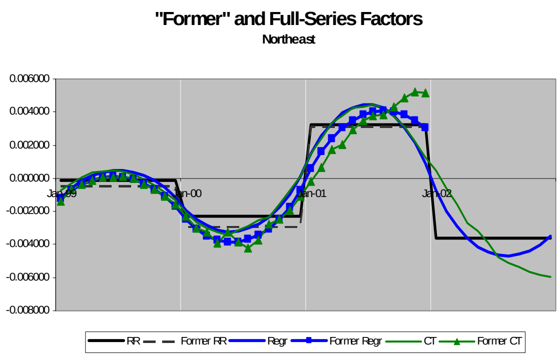


Figure 4. Example Revisions for Series with Larger Benchmark Factors



For series where there was a difference in the revisions, the revisions were smaller for the Regression method. The Regression method is more stable than the other methods at the beginning and ending of the series.

6. Conclusions

The Causey-Trager method and the Regression method with $\lambda = 1$ gave very consistent results with smooth benchmark factors and small discrepancies from the unbenchmarked seasonally adjusted series. The Causey-Trager method gave results that were more consistent at the break between year, but there is still a problem with month-to-month percent changes being distorted at the beginning and ending of the year, for all methods. Regression method gave results at the beginning and ending of the series that may be more appropriate for benchmarking seasonally adjusted series with smaller revisions when new data was added.

In the future I plan to look at possible modifications to the Causey-Trager method to adjust the factors at the beginning and ending of the series and to investigate generalized methods for aggregate series.

7. Acknowledgements

The author wishes to thank especially Benoit Quenneville and others at Statistics Canada for sharing their benchmarking software. Thanks also go to the support from the U.S. Census Bureau: Brian Monsell for his help with the Causey-Trager FORTRAN code; Nash Monsour and David Findley for their support in the benchmark research; and Nash, Brian, Kathy McDonald-Johnson, Roxanne Feldpausch, and Ayonda Dent for their suggestions and comments on the paper.

8. References

Bozik, J.E. and M.C. Otto (1988), "Benchmarking: Evaluating Methods that Preserve Month-to-Month Changes." Statistical Research Report Series 88/07, U.S. Census Bureau: Washington, DC.

Causey, B. and M.L. Trager (1981), "Derivation of Solution to the Benchmarking Problem: Trend Revision." Unpublished research notes, U.S. Census Bureau.

Cholette, P.A. (1978), "A Comparison and Assessment of Various Adjustment Methods of Sub-Annual series to Yearly Benchmarks." Ottawa, Canada: Statistics Canada, Methodology Branch, Seasonal Adjustment and Time Series Staff Working Paper.

Denton, F.T. (1971), "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization." *Journal of the American Statistical Association, 82*, 99-102.

Huot, G. (1975), "Quadratic Minimization Adjustment of monthly or Quarterly Series to Annual Totals." Ottawa, Canada: Statistics Canada, Methodology Branch, Seasonal Adjustment and Time Series Staff Working Paper.

Ladiray, D. and Quenneville, B. (2001), *Seasonal Adjustment with the X-11 Method.* New York, NY: Springer-Verlag, Lecture Notes in Statistics 158.

Trager, M.L. (1982), "Derivation of Solution to the Benchmarking Problem: Relative Revision." Unpublished research notes, U.S. Census Bureau.

Quenneville, B., G. Huot, and K. Chiu (2001), "Benchmarking of D11 Using the Henderson Filter to Smooth the Discrepancies Between the Annual Totals of A1 and D11."

Quenneville, B., P. Cholette, G. Huot, K. Chiu, and T. Di Fonzo (2004), "Adjustment of Seasonally Adjusted Series to Annual Totals."

U.S. Census Bureau (2002), *X-12-ARIMA Reference Manual, Version 0.2.10.* Washington, DC: Time Series Staff, Statistical Research Division, U.S. Census Bureau.