A Decomposition Approach to Revisions in Aggregated Time Series Estimates

Duncan Elliott[1], Craig H. McLaren[1], Julian Chow[1]
[1]Office for National Statistics, Cardiff Road, Newport NP10 8XG, United Kingdom

Abstract
Users of official statistics are often concerned with the source and impact of revisions in time series estimates. In official statistics it can be difficult to assess the impact and source of revisions where seasonally adjusted estimates are derived indirectly through the aggregation of seasonally adjusted component series. We investigate how higher level aggregates are impacted by revisions from directly seasonally adjusted time series by using a decomposition framework. The implications of this approach for interpretation of time series estimates and the production process in official statistics are also discussed.

Key Words: revisions, seasonal adjustment, aggregation, time series, decomposition

1. Introduction

Time series published in official statistics, in particular, seasonally adjusted and trend estimates are subject to revisions. We identify three principal causes of revisions; change(s) to past data point(s) of the unadjusted time series, additional time point(s) in the unadjusted series, and a change to one or more of the parameters used to derive the seasonally adjusted or trend estimates.

In official statistics the cause and source of revisions is often not readily identifiable, as more than one of the principal causes occur simultaneously. Further complications in official statistics include the nature of data sets where estimates are part of an aggregation structure with revisions to lower level series flowing through to the aggregated series.

Cook (2004) notes the importance of revisions to official statistics and there has been a considerable amount of work done on the reporting and analysis of revisions. Some of this literature is discussed by Elliott *et al* (2007), but some recent work includes Menrhoff (2008) on quantifying the sources of revisions, and further efforts on modelling revisions such as Jacobs and van Norden (2008) and Cunningham *et al* (2008). The literature on revisions might be broadly classified as either describing or using revisions.

Elliott *et al* (2007) explore revisions due to an additional time point in an unadjusted time series utilising a decomposition of growth rates approach to describe revisions. In this current work we use the same framework to look at revisions due to a change in a past time point of the unadjusted time series within a simple aggregation structure. The aim is to provide users and producers of time series data with detailed information on the sources of revisions with a clear nomenclature of revisions that can be implemented in practice.

Mainwaring and Skipper (2007) provide a method for explaining the sources of revisions that involves a certain level of subjectivity concerning quantification of both revisions due to an annual seasonal adjustment review and revisions caused by seasonal adjustment updates due to later data. Using the approach described in Elliott *et al* (2007) for decomposing revisions due to an additional time point can reduce the subjectivity involved in the Mainwaring and Skipper (2007) approach. By addressing how revisions flow through an aggregation structure when a past time point in the unadjusted time series is revised, holding other sources of revision constant we contribute to a further step in providing detailed information on the sources of revision.

The next section describes the framework for revisions to seasonally adjusted estimates in an aggregation structure, whilst section three presents and discusses results of some simulations and section four concludes with areas for further work.

2. Revisions in an Aggregation Structure

2.1 Aggregation Structure

Any unadjusted time series ($V_{t|\tau}$ which is an estimate at time t given data up to and including time τ), can be described as a function f of component series

$$V_{t|\tau} = f(C_{V,t|\tau}, S_{V,t|\tau}, I_{V,t|\tau}, \ldots)$$

where

$C_{V,t|\tau}$ is the trend component of the series $V_{t|\tau}$
$S_{V,t|\tau}$ is the seasonal component of the series $V_{t|\tau}$
$I_{V,t|\tau}$ is the irregular component of the series $V_{t|\tau}$

for

$t = 1, 2, \ldots, \tau, \tau + 1, \tau + 2, \ldots, \tau + k$

and where k denotes a k-step ahead forecast (k = 1, 2, ...)

We assume a simple aggregation structure,

$$Z_{t|\tau} = \alpha X_{t|\tau} + (1-\alpha) Y_{t|\tau}$$

where,

$Z_{t|\tau}$ is an unadjusted aggregate time series of length τ
$X_{t|\tau}$ is an unadjusted component time series of $Z_{t|\tau}$ of length τ
$Y_{t|\tau}$ is an unadjusted component time series of $Z_{t|\tau}$ of length τ
$0 \leqslant \alpha \leqslant 1$

We assume a multiplicative decomposition for the component series of $Z_{t|\tau}$, $X_{t|\tau}$ and $Y_{t|\tau}$ such that

$$X_{t|\tau} = C_{X,t|\tau} S_{X,t|\tau} I_{X,t|\tau}$$

$$Y_{t|\tau} = C_{Y,t|\tau} S_{Y,t|\tau} I_{Y,t|\tau}$$

$X_{t|\tau}$ and $Y_{t|\tau}$ may represent either a single series or an aggregate time series, each consisting of any number of component time series. For the purposes of this research we treat each as individual time series where both are directly seasonally adjusted, as described below. However, in interpreting the results discussed in section three, $Y_{t|\tau}$ can be considered as representing either one or multiple time series.

2.2 Seasonal Adjustment

Seasonal adjustment involves the estimation and removal of the seasonal component $S_{V,t|\tau}$ from an unadjusted series $V_{t|\tau}$. Direct seasonal adjustment for a multiplicative decomposition of $Z_{t|\tau}$ is given as

$$SA(Z_{t|\tau}) = \frac{Z_{t|\tau}}{\hat{S}_{Z,t|\tau}}$$

Indirect seasonal adjustment of $Z_{t|\tau}$ is given as

$$SAind(Z_{t|\tau}) = \alpha\left(\frac{X_{t|\tau}}{\hat{S}_{X,t|\tau}}\right) + (1-\alpha)\left(\frac{Y_{t|\tau}}{\hat{S}_{Y,t|\tau}}\right)$$

2.3 Revisions

Revisions as a proportion of the unadjusted time series $X_{t|\tau}$ can happen at any time point t and are denoted by the time series $R_{t|\tau}$,

$$R_{t|\tau} = \left(\frac{X^*_{t|\tau}}{X_{t|\tau}}\right)$$

where

$X^*_{t|\tau}$ is the revised version of the unadjusted time series $X_{t|\tau}$

The revisions to $X_{t|\tau}$ cause revisions to $Z_{t|\tau}$ of

$$Z_{t|\tau} + (R_{t|\tau} - 1)\alpha X_{t|\tau} = \alpha X_{t|\tau} R_{t|\tau} + (1-\alpha) Y_{t|\tau}$$

Revision of the unadjusted time series $X_{t|\tau}$ will cause revisions to the directly seasonally adjusted estimate of the revised aggregate series of

$$SA(Z_{t|\tau} + (R_{t|\tau} - 1)\alpha X_{t|\tau}) = \frac{Z_{t|\tau} + (R_{t|\tau} - 1)\alpha X_{t|\tau}}{\hat{S}^*_{Z,t|\tau}}$$

where

$\hat{S}^*_{Z,t|\tau}$ is the revised estimate of the seasonal component for the aggregate series $Z_{t|\tau} + (R_{t|\tau} - 1)\alpha X_{t|\tau}$

Whilst the indirect seasonally adjusted estimate of the revised aggregate series is given by

$$SAind(Z_{t|\tau} + (R_{t|\tau} - 1)\alpha X_{t|\tau}) = \frac{\alpha X_{t|\tau} R_{t|\tau}}{\hat{S}^*_{X,t|\tau}} + (1-\alpha)\frac{Y_{t|\tau}}{\hat{S}_{Y,t|\tau}}$$

where

$\hat{S}^*_{X,t|\tau}$ is the revised estimate of the seasonal component for the component series $X_{t|\tau} \cdot R_{t|\tau}$

3. Revisions

3.1 Simulation Design

We simulated n revision series such that

$$R_{i,t|\tau} = \begin{cases} 1, & t \neq \tau - 1 \\ 1 + r_i, & t = \tau - 1 \end{cases}$$

where

$$r_i \sim N(\mu, \sigma^2)$$

for

$t \leq \tau$
$i = 1, 2, ..., n$

The n revision series were used to generate n revised versions of the component time series $X_{t|\tau}$ to allow analysis of the impact of such revisions on the growth rates of the derived components of $Z_{t|\tau}$ when directly and/or indirectly seasonally adjusted. Whilst revisions can and do occur at time points other that $\tau - 1$, in practice this is one of the most common time points to be revised, and the motivation for exploring the impact of revisions to this time point. It should be noted that there is an implicit assumption; there is no revision due to the one-step ahead forecast error, as discussed in Elliott *et al* (2007), as $V_{\tau|\tau} = V_{\tau|\tau-1}$

The growth rate of any time series $V_{t|\tau}$ at $t = \tau - 1$ is given by

$$\Delta V_{\tau-1|\tau} = \frac{V_{\tau-1|\tau}}{V_{\tau-2|\tau}} - 1$$

Whilst the change to the growth rate of any time series $V_{t|\tau}$ that has been revised to $V^*_{t|\tau}$ is given by

$$\nabla V^*_{\tau-1|\tau} = \Delta V^*_{\tau-1|\tau} - \Delta V_{\tau-1|\tau}$$

In these simulations we have used the US Census Bureau's official release version of X-12-ARIMA for seasonal adjustment. The results discussed in section 3.2 are based on simulated revisions using monthly UK Retail Sales data where

$Z_{t|\tau}$ is Non-specialised stores
$X_{t|\tau}$ is Non-specialised food stores
$Y_{t|\tau}$ is Non-specialised non-food stores

for

t = January 1986, February 1986, ..., October 2007
τ = October 2007

The parameters for seasonal adjustment include: an ARIMA model chosen by the 'pickmdl' modelling regime of X-12-ARIMA; no regression variables as the data has been adjusted for other calendar related effects; and the trend and seasonal filters chosen by default. The magnitude of the revisions simulated for the one time point in the component time series does not change the order of the ARIMA model selected or length of seasonal or trend filters chosen by default for the directly seasonally adjusted component or aggregate series.

3.2 Results

Table 1 reports the median revision to the growth rate of components in the aggregate series relative to the revision to the growth rate of $X_{t|\tau}$ for $n = 100$, $\mu = 0$, $\sigma^2 = 0.01$and for different weights α.

If the aggregate series $Z_{t|\tau}$ is indirectly seasonally adjusted with $\alpha = 0.5$ then a 1% change to the growth rate $\Delta X_{\tau-1|\tau}$ due to a revision in the level $X_{\tau-1|\tau}$ will cause a revision of 0.39 to the growth rate of the aggregate seasonally adjusted series, driven predominantly by a 0.36 revision of the growth rate of the irregular component ($\Delta I_{Z,\,\tau-1|\tau}$).

In this instance the directly seasonally adjusted series shows a similar magnitude of relative revisions for the growth rates of its components (table 2) as the indirect seasonally adjusted series (table 1). However, the revisions to the trend component ($C_{Z,t|\tau}$) are marginally smaller for comparable values of α.

Table 1: Impact of revisions in an aggregate series indirectly seasonally adjusted at τ-1

Weight	Seasonally Adjusted (CI)	Trend (C)	Seasonal Factor (S)	Irregular (I)
$\alpha = 0.9$	0.65	0.05	0.27	0.60
$\alpha = 0.8$	0.59	0.04	0.25	0.54
$\alpha = 0.7$	0.52	0.04	0.23	0.48
$\alpha = 0.6$	0.46	0.03	0.21	0.42
$\alpha = 0.5$	0.39	0.03	0.18	0.36
$\alpha = 0.4$	0.31	0.02	0.15	0.29
$\alpha = 0.3$	0.24	0.02	0.12	0.22
$\alpha = 0.2$	0.16	0.01	0.08	0.15
$\alpha = 0.1$	0.08	0.01	0.04	0.08

Table 2: Impact of revisions in an aggregate series directly seasonally adjusted at τ-1

Weight	Seasonally Adjusted (CI)	Trend (C)	Seasonal Factor (S)	Irregular (I)
$\alpha = 0.9$	0.65	0.03	0.28	0.62
$\alpha = 0.8$	0.59	0.03	0.25	0.55
$\alpha = 0.7$	0.53	0.03	0.23	0.49
$\alpha = 0.6$	0.46	0.03	0.20	0.43
$\alpha = 0.5$	0.40	0.03	0.17	0.37
$\alpha = 0.4$	0.33	0.02	0.14	0.30
$\alpha = 0.3$	0.25	0.01	0.11	0.24
$\alpha = 0.2$	0.17	0.01	0.07	0.16
$\alpha = 0.1$	0.09	0.01	0.04	0.08

Figures 1 and 2 show revisions to the month-on-month movements of the directly estimated trend and seasonally adjusted aggregate series at $\tau - 1$ respectively. These revisions are plotted against revisions to the level of $X_{t|\tau}$ where the revisions in this instance have been generated for n = 100, μ = -0.015 and σ^2 = 0.00158, the mean and variance of revisions to the UK Retail Sales Index.

Figure 1 shows that revisions to the growth rate of the trend in the directly seasonally adjusted estimate are proportional to the revision of the level of the component series $X_{t|\tau}$. The revision to the trend is also proportional to the value of α, as noted in the tables above. Despite some relatively large revisions to the level of $X_{t|\tau}$, the revisions to the growth of the trend in the seasonally adjusted series are relatively small.

For α less than 0.3 the revisions to growth of the trend component are approximately constant in proportion to the revision of the level of $X_{t|\tau}$, whilst for α greater than or equal to 0.3 there is some variation in this proportion. This variation in proportion is found at smaller values of absolute revisions of $X_{t|\tau}$ as the value of α increases.

A similar effect can also be found in figure 2 that shows revisions to the growth rate of the directly seasonally adjusted estimate of the aggregate series, albeit to a lesser extent. This is due to the compensating effect of revisions in the irregular and trend components.

The potential causes of this change in proportion are the identification of outliers as part of the X-11 algorithm and the impact of the revised forecast. In no case does the revision cause a change to the default selection of seasonal or trend filter length or the order of the ARIMA model.

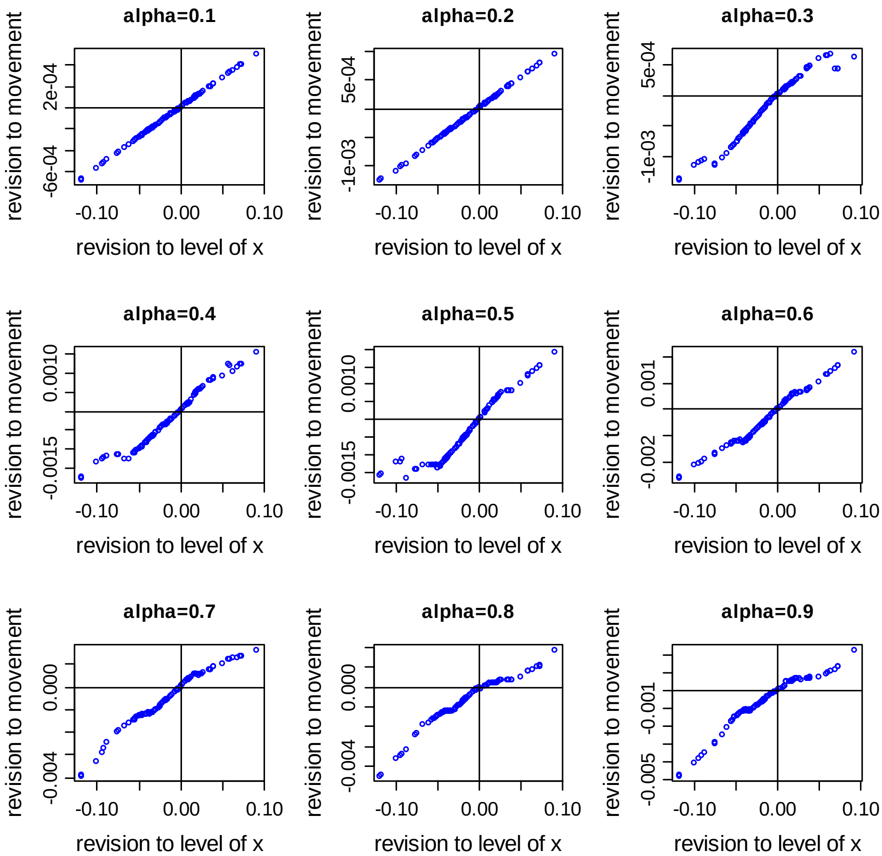

Figure 1: revision to month on month movement of trend at τ-1

4. Discussion

We show that if the revisions to a component series ($X_{t|\tau}$) are known, and we know the proportion that the component series represents in a higher level aggregate, we can identify what would be the approximate contribution to the revision of the growth rate of the components of the aggregate series ($C_{Z,t|\tau}$, $S_{Z,t|\tau}$, $I_{Z,t|\tau}$, …) whether directly or indirectly estimated. Such a system could be implemented as part of a revisions analysis tool set for decomposing the contributions to revisions.

One area for further research is to explore the variation in proportion of revisions to the growth rates of components in the aggregate series relative to the revisions in the level of $X_{t|\tau}$. The same variation in proportion is found in the revision to the growth rates of the components of the component series ($C_{X,t|\tau}$, $S_{X,t|\tau}$, $I_{X,t|\tau}$, …) for a given revision to the level of $X_{t|\tau}$ and is therefore not an issue of aggregation, but concerns the process of seasonal adjustment. The most likely causes are the automatic outlier identification in the X-11 algorithm and the re-estimation of parameters in the ARIMA model.

The use of the information presented here as part of a revisions analysis system would require a number of steps whilst holding other sources of revision unchanged. We have focussed on revision due to a change in the level of $X_{t|\tau}$ where the length of the series (τ) remains fixed, and where the length of seasonal and trend filters and the order of the ARIMA model used in the estimation of the components of series in an aggregation structure also remain fixed. In practice the length of the series increases at the same time as past points are revised and the trend and seasonal filters or order of an ARIMA model may also change. A revisions system would therefore look at the contribution and interaction between these sources of revision and could conceivably include a decomposition of the types of revision to the level of $X_{t|\tau}$.

Furthermore, as noted in Elliott et al (2007) if there are revisions in the level of component series $Xt|\tau$ and $Yt|\tau$ then the revisions seen in the aggregate series $Zt|\tau$ will depend upon the correlation of revisions in the component series. In the results above we have presented a simplified aggregation structure, treating $Yt|\tau$ as one series, but further complications could be made where $Yt|\tau$ is multiple series. A further simplifying assumption that could be relaxed concerns the value of α which we have treated as a constant. In practice α will often vary over time. Another area that has not been considered is the impact of revisions policies, where certain time points may be revised at certain points in time.

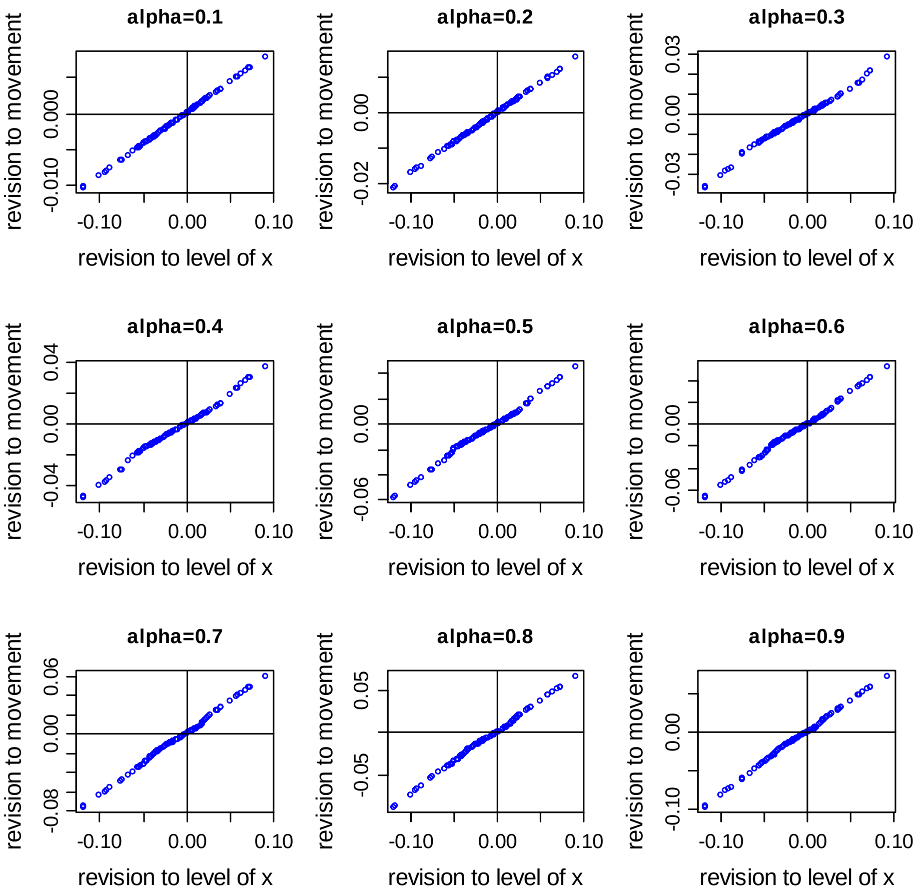

Figure 2: revision to month on month movement of seasonally adjusted series at τ-1

Acknowledgements

The analysis and views presented in this paper are those of the authors and do not necessarily represent those of the Office for National Statistics.

References

Cook, L. (2004) 'Revisions to statistics: their role in measuring economic progress' *Economic Trends*, (February 2004), pp. 36-43

Cunningham, A., J. Eklundy, C. Jeffery,G. Kapetanios, and V. Labhard (2008) 'A State Space Approach To Extracting The Signal From Uncertain Data' *National Statistics Methodology Advisory Committee,14th meeting, 13 May 2008 http://www.statistics.gov.uk/methods_quality/nsmac_fourteenth_meeting.asp*

Elliott, D., C.H. McLaren, and X. Zhang, (2007) 'Quantifying and Measuring Revisions in Time-Series Estimates' *Proceedings of the Survey Research Section*, American Statistical Association, Joint Statistics Meeting, Salt Lake City, July 29 to August 2, 2007.

Jacobs, J. P.A.M. and S. van Norden (2008) 'Modelling Data Revisions: Measurement Error and Dynamics of "True" Values', 5th Colloquium on Modern Tools for Business Cycle Analysis, 29th September – 1st October 2008.

Mainwaring, H. and H. Skipper (2007) 'GDP(O) revisions analysis system: overview and indicative results' *Economic and Labour Market Review*, Vol. 1, No. 10, pp. 36-42

Mehrhoff, J. (2008) 'Sources of Revisions of Seasonally Adjusted Real Time Data', 5th Colloquium on Modern Tools for Business Cycle Analysis, 29th September – 1st October 2008.