Statistics New Zealand
TATAURANGA AOTEAROA

Comparing revisions in time series data

A report on seasonally adjusted versus trend series

Ryan Buchanan and Conrad MacCormick

Crown copyright ©
This work is licensed under the Creative Commons Attribution 3.0 New Zealand licence. You are free to copy, distribute, and adapt the work, as long as you attribute the work to Statistics NZ and abide by the other licence terms. Please note you may not use any departmental or governmental emblem, logo, or coat of arms in any way that infringes any provision of the Flags, Emblems, and Names Protection Act 1981. Use the wording 'Statistics New Zealand' in your attribution, not the Statistics NZ logo.

Disclaimer
This paper represents the views of the authors. It does not necessarily represent the views of Statistics NZ and does not imply commitment by Statistics NZ to adopt any findings, methodologies, or recommendations. Any data analysis was carried out under the security and confidentiality provisions of the Statistics Act 1977.

Liability
While all care and diligence has been used in processing, analysing, and extracting data and information in this publication, Statistics New Zealand gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

Citation
Buchanan, R, and MacCormick, C (2014). *Comparing revisions in time series data: A report on seasonally adjusted versus trend series.* Available from www.stats.govt.nz.

ISBN 978-0-478-40877-5 (online)

Published in July 2014 by
Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

Contact
Statistics New Zealand Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4610
www.stats.govt.nz

Contents

List of tables and figures

List of tables

Figure

1 Purpose and summary

Purpose

In this paper, we describe our investigation into whether using the trend rather than the seasonally adjusted series could improve the stability of the figures in Statistics New Zealand publications.

Summary

We compared revisions in time series data to investigate if using the trend series could result in greater stability in our released figures.

Here is a summary of our methodology and results.

- We analysed 15 Statistics NZ time series from the business and social areas, using absolute total revisions and Levene’s test for equality of variances.
- We calculated and analysed the absolute total revisions for the seasonally adjusted and trend estimators using various statistical tests, including Levene’s test.
- We used Levene’s test to identify if there is a statistically significant difference between the variances of the seasonally adjusted and trend series in terms of total revisions.
- Levene’s test for equality of variances plus other tests confirmed that the difference between the variances of the seasonally adjusted and trend series was statistically significant for all 15 time series.
- The absolute total revision graphs showed that the trend series had larger revisions than the seasonally adjusted series in all 15 series we analysed.
- The plots of the seasonally adjusted revisions and trend revisions showed the trend series to be less stable than the seasonally adjusted series. This was more pronounced towards the recent end of the series due to the end-point problem.
- The seasonally adjusted series showed consistent significantly smaller variance than the trend cycle series.
- We found the seasonally adjusted series consistently outperformed the trend series in the relative size of revisions.
- For greater stability in revisions in Statistics NZ’s official estimates, we recommend leading with the seasonally adjusted series.

2 Introduction

At Statistics NZ, we are currently investigating the option of leading our media releases with either the trend or the seasonally adjusted series as the official estimate for our time series outputs. 'Lead' refers to which type of series (eg seasonally adjusted; trend) we emphasise when we report on estimates in our media and information releases. The lead series is the focus of the main headlines or titles, and the main pieces of analysis. It is important that we select a single series to lead with for the sake of both consistency and impartiality (ie we don't want to be seen to choose the series that paints the best picture).In recent times, attention has been drawn to the volatility of the seasonally adjusted series of some outputs.

We decided to investigate the benefits of using the trend series to gain more stability in our released figures.

3 Data preparation

An initial stage was preparing the input datasets for use in the tests.

We analysed the following social and economic series.

- Balance of Payments
  - The current account
- National Accounts
  - Gross Domestic Product (quarterly)
  - OECD aggregates of industries: Wholesale trade, Retail trade and accommodation, Communications
  - OECD aggregates of industries: Local and central government, Education and training, Health care and social assistance
  - Information Media and Telecommunications
  - Taxes on products: Import duties
- Trade imports and exports – Overseas Merchandise Trade
  - Intermediate goods
  - Value of imports (total)
- Household Labour Force Survey
  - Total employed
  - Total unemployed
  - Total not in the labour force
  - Total part-time employed
  - Total employment rate
  - Total unemployment rate
  - Total labour force participation rate

We tested these series to make sure they had stable seasonality at the 0.1 percent level. This check helped to ensure our tests would be valid. These series contain quarterly and monthly data.

To make sure the data we were analysing was homogenous (ie not influenced by the filters of the X-12 seasonal adjustment package), we shortened the start and end of the series. We excluded the first five years of the series to ensure the filter selection functioned properly in X-12 (ie the first five years uses non-standard filters). We excluded the last three years to obtain convergence (ie stability) near the end of the series. This ensures the real-world revisions in the datasets are analysed, and not revisions caused by the moving filters of X-12.

What we have left can be considered the ‘true’ final estimates. These ‘true’ estimates mean we can be confident in our revision estimates. For more information about how we removed the ends of the series, see the appendix.

We created graphs of the absolute total revisions over the period where the data can be considered homogenous. A total revision is the revision of each estimate from when we first published it to what the estimate is currently – when the series has stabilised.

4 Methods: Using Levene's test for equality of variances

Many of the distributions we tested were non-normal, so we decided to use Levene's test for equality of variances (or simply, Levene's test).

Levene's test is a non-parametric test that determines if there is a statistically significant difference between the variances of two distributions. This method can robustly deal with non-normal data. It is widely recommended in literature to use Levene's test instead of the Brown-Forsythe test when testing only two distributions (comparing seasonally adjusted and trend series in our case), as it has greater power and efficiency.

There are two approaches to computing Levene's test: the deviations from the group mean are either squared, or transformed to become absolute. We used the absolute approach in this investigation because the underlying data contained some large outliers, which become problematic when squared.

Levene's test is defined as:

Ho: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

Ha: $\sigma_i^2 \neq \sigma_j^2$ for at least one pair (i, j)

Test statistic: Given a variable Y with sample of size N divided into k subgroups, where N_i is the sample size of the ith subgroup, Levene's test statistic is defined as

$$W = \frac{(N-k)}{(k-1)} \frac{\sum_{i=1}^{k} N_i \, |(\bar{Z}_{i.} - \bar{Z}_{..})|}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} |(\bar{Z}_{ij} - \bar{Z}_{i.})|}$$

It is important to note that there are cases where Levene's test will not perform well. If the mean of the revision series is biased (ie where the final estimate is consistently under- or over-estimated) then this undesirable bias should be accounted for in the test statistic. In this case, a revised Levene's test statistic should be used. For more information and the Levene's test statistic, see the appendix.

The structure of the datasets that feed into Levene's test is detailed in the appendix.

5 Results: The trend series is generally less stable

Our investigation focused on which of the seasonally adjusted series and trend series produces the most stable revision series. The desired series will display, on average, smaller percentage-based revisions over the time series, and minimal outliers in its revisions.

The effect of the end-point problem caused by using symmetric filters (specifically, Henderson filters) in X-12 will be as small as possible in the preferred series. The end-point problem refers to trend estimates being more volatile (ie highly variable, or unpredictable) and slower to react to changes towards the end of a series, especially when approaching turning points. A turning point in a time series is where the series changes from heading in one direction to heading in the opposite direction (ie increasing to decreasing, or decreasing to increasing).

It is important to note that the results of this paper apply to the entirety of a seasonal time series. That is, including the first five years removed the beginning of the series and the final three years removed from the end.

Because the results of the 15 series all displayed similar characteristics, we'll use the intermediate goods (Intgoods) series as an example in the remainder of the results section.

We used the UNIVARIATE procedure (PROC UNIVARIATE) in SAS (statistical analysis software used at Statistics NZ) to test if the selected series followed a normal distribution. From these tests we determined that many of the series did not follow a normal distribution and contained high levels of skewness and kurtosis. Skewness refers to the symmetry of a particular distribution, while kurtosis describes the heaviness of a distribution's tail. A perfectly symmetrical distribution without any large values in the tail is preferred.

Table 1
SAS PROC UNIVARIATE test results

Skewness	3.1887589
Kurtosis	14.8288539

Levene's test confirmed at the 95 percent confidence level that there is a statistically significant difference in the variance of the trend and seasonally adjusted series for each series.

Output of Levene's test confirmed that the seasonally adjusted series consistently outperformed the trend series across all 15 analysed series.

Table 2
Levene's test results (F and p values) – Intgoods

F value	7.28
p value	0.0079

Table 3
Levene's test results (mean and standard deviation) – Intgoods

Series	Mean	Std dev
Seasonally adjusted	2303.73642	44847.221
Trend	-9949.92416	117437.735

To verify that Levene's test was not affected by underlying bias in any of the series, we performed a one-way t-test on each series.

We found no statistically significant bias in any of the 15 series we analysed.

This result verified that the results of Levene's test were valid.

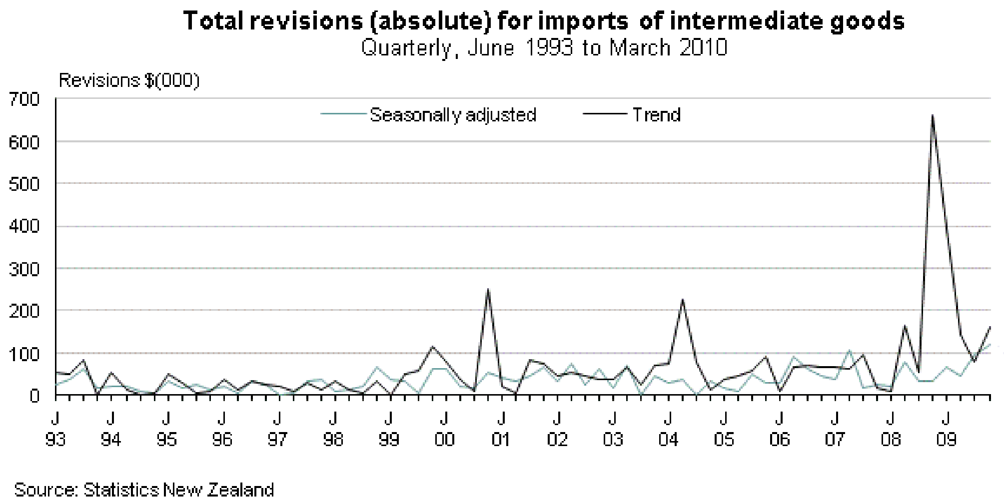
A modified Levene's test – for cases where there is a statistically significant bias – is detailed in the appendix.

Table 4
t-test results – Intgoods

Series	t value	p value
Seasonally adjusted	0.42	0.6732
Trend	-0.70	0.4872

These graphs showed that the trend series is generally less stable than the seasonally adjusted series. This instability can be seen by the trend series having larger absolute revisions on average, but also by the trend series containing more large spikes where the revisions are far greater than those of the seasonally adjusted series. Figure 1 shows the results for the intermediate goods series.

Figure 1



To get an overview of the difference between the seasonally adjusted and trend series, we normalised the total revisions for each of the 15 series and combined them to give one overall distribution for each of the seasonally adjusted and trend series.

We then tested this distribution for significance using a t test. This showed that the two distributions were significantly different and that over the two series, the seasonally adjusted had smaller overall variance than the trend.

Table 5
t-test results (t and p values) – Intgoods

Series	t value	p value
Pooled	-2.10	0.0355
Satterthwaite	-2.10	0.0355

Table 6
t-test results (mean and standard deviation) – Intgoods

Series	Mean	Std Dev
SA	-0.00013	0.0297
Trend	0.00315	0.0417

6 Which series do other countries and organisations lead with?

Country/Organisation	
Eurostat	Recommends leading with seasonally adjusted
OECD	Recommends leading with seasonally adjusted
USA	Leads with seasonally adjusted
UK	Leads with seasonally adjusted
Canada	Leads with seasonally adjusted
Australia	Leads with trend
Denmark	Leads with seasonally adjusted
Hungary	Leads with seasonally adjusted
Ireland	Leads with seasonally adjusted
Israel	Leads with seasonally adjusted

For further information, see 'References' section.

7 Conclusion: The seasonally adjusted series is preferable

To minimise revisions near the end-point of a seasonal time series, we recommend leading with the seasonally adjusted series.

After analysing the previously mentioned series, we found that the results were consistently in favour of the seasonally adjusted series across all individual distributions and also at the total variance level. Therefore, at turning points and given the volatility of our series, the seasonally adjusted series outperformed the trend series in terms of revisions.

The trend series shows generally larger inconsistency in its revisions than the seasonally adjusted series close to the end-point. This is illustrated by large spikes that appear in the total absolute revisions graph.

References

Central Bureau of Statistics (Israel) (nd). Seasonal adjustment (PDF, 530kb). Accessed 7 February 2014, available from www.cbc.gov.il.

Central Statistics Office (CSO – Ireland) (2000). Comparing seasonally adjusted and trend series: An assessment. Accessed 7 February 2014, available from www.cso.ie.

Eurostat (2009). Eurostat methodologies and working papers: ESS guidelines on seasonal adjustment (PDF, 615kb). Accessed 7 February 2014, available from http://epp.eurostat.ec.europa.eu.

Linde, P (Statistics Denmark) (2005). Seasonal adjustment (PDF, 688kb). Accessed 7 February 2014, available from www.dst.dk.

McDonald-Johnson, KM, Monsell, B, Fescina, R, Feldspauch, R, Hood, CCH, & Wroblewski, M (2010). Seasonal adjustment diagnostics: Census Bureau guideline (PDF, 55kb). Accessed 7 February 2014, available from www.census.gov.

McLaren, CH, & Zhang, X(M) (2010). The importance of trend-cycle analysis for national statistics institutes (PDF, 399kb). *Estudios de Economia Aplicada*, *28*(3), 607–23. Accessed 7 February 2014, available from www.redalyc.org.

Hungarian Central Statistical Office (2007). Seasonal adjustment: Methods and practices version 3.1 (PDF, 2.47mb). Accessed 7 February 2014, available from http://epp.eurostat.ec.europa.eu.

OECD (2004). OECD Short-term Economic Statistics Expert Group draft metadata and reporting manual (PDF, 522kb). Accessed 7 February 2014, available from www.oecd.org.

ONS (nd). Methodology of the monthly Index of Services: Seasonal adjustment (PDF, 125kb). Accessed 7 February 2014, available from www.ons.gov.uk.

Rodnyanskiy, L (2010). *Guidelines for presentation and analysis of time series*. Unpublished document, Collection and Statistical Methodologies DocONE library (CSM-RES-03-02-08-06, DocID W193289). Available on request from Statistics NZ, Wellington.

Statistics Canada (2009). Statistics Canada quality guidelines (fifth edition): Seasonal adjustment and trend-cycle estimation. Accessed 7 February 2014, available from www.statscan.gc.ca.

Appendix 1: Methodology for measuring the size of revisions

Consider a time series vector $\mathbf{Z}^{P,T}$ with an annual period of P (4 quarters or 12 months) and with T time elements defining the time series $\mathbf{Z}$ running from time $t = 1 \text{ to T}$. The t th elements of $\mathbf{Z}^{P,T}$ is given by $z_t^{P,T}$. We will assume that the time series $\mathbf{Z}^{P,\mathrm{T}}$ has a homogenous generating function and has no structural changes over the time period $t = 1 \text{ to T}$.

Now let us apply a transformation F_τ^{SA} to the vector defined by the first τ elements in vector $\mathbf{Z}^{P,\mathrm{T}}$.

$$\mathbf{Y}_\tau^{SA,\mathrm{P,T}} = \mathrm{F}_\tau^{SA}\left(\mathbf{Z}_\tau^{P,T}\right) \quad (0.1)$$

There are T possible vectors $\mathbf{Y}_\tau^{SA,\mathrm{P,T}} \ \forall \ \tau\text{=1 to T}$ that can be produced by F_τ^{SA}. Within the vector $\mathbf{Y}_\tau^{SA,\mathrm{P,T}}$ there will be τ time elements running from time $t = 1 \text{ to } \tau$. The tth elements of $\mathbf{Y}_\tau^{SA,\mathrm{P,T}}$ is given by $y_{t,\tau}^{SA,P,T}$. In our case, the transformation function $\mathrm{F}_\tau^{SA.P}$ will produce the final seasonally adjusted series (D11) of the X-12-ARIMA seasonal adjustment program.

As we keep adding time points to the input vector $\mathbf{Z}^{P,\tau}$ we get revised seasonally adjusted vectors $\mathbf{Y}_\tau^{SA,\mathrm{P,T}}$.[1] As we add each new time series point to $\mathbf{Z}_\tau^{P,T}$ the estimates for $y_{t,\tau}^{SA,P,T}$ will constantly be revised.[2] Thus for each time point t, we will get a vector of revised estimates for $y_{t,\tau}^{SA,P,T}$. There will be T possible output vectors $\mathbf{Y}_\tau^{SA,\mathrm{P,T}}$. Adding one extra data point to $\mathbf{Z}_\tau^{P,T}$ produces the following revision at time point t.

$$r_{t,\tau}^{SA,P,T} = y_{t,\tau+1}^{SA,P,T} - y_{t,\tau}^{SA,P,T} \quad (0.2)$$

The total observed revision at time t is given by:

$$R_t^{SA,P,T} = \sum_{s=t}^{T-1} r_{t,\mathrm{s}}^{SA,P,T} = y_{t,\mathrm{T}}^{SA,P,T} - y_{t,t}^{SA,P,T} \quad (0.3)$$

At each time point t the total observed revision will be made up of (T-t-1) incremental revisions $r_{t,\mathrm{s}}^{SA,P,T}$. As one continues to add time points, the X-12 filters become fixed finite

---

[1] Later in this section we define and use an alternative transformation function F_τ^{TC} that produces the final trend-cycle series (D12) of the X-12-ARIMA seasonal adjustment program.

[2] Each new input data point $z_{t+1}^{P,T}$ that is added to the vector $\mathbf{Z}_\tau^{P,T}$ will affect the estimates of the seasonal component and trend component for previous periods. Each added data point can potentially cause revisions along the full length of the seasonally adjusted vector $\mathbf{Y}_\tau^{SA,\mathrm{P,T}}$.

lengths and are symmetric in time. The filters used to estimate $y_t^{SA,P,T}$ will become fixed, then symmetric revisions will stop[3]. Adding extra data points should no longer revise the data point at time t. This convergence should happen before three years of revisions. Thus, in practice, we should get convergence in estimates of total revisions within $P*3$ revisions.

For long time series, the nominal observed level of the time series can change by orders of magnitude. Thus the variance (or dispersion) of $R_t^{SA,P,T}$ is not homogeneous through time. Normalising $R_t^{SA,P,T}$ can significantly reduce the non-homogeneity of the total observed revision, $R_t^{SA,P,T}$. $S_t^{SA,P,T}$, the normalised total revision for transformation $\mathrm{F}^{SA.P}$, is given by:

$$S_t^{SA,P,T} = \frac{y_{t,\mathrm{T}}^{SA,P,T} - y_{t,t}^{SA,P,T}}{y_{t,\mathrm{T}}^{SA,P,T}} \qquad (0.4)$$

The X-12-ARIMA transformation process has time non-homogeneities in how it treats the time series. X-12-ARIMA requires years of data before the filter selection functions properly. Thus time series of less than 5*P elements use non-standard filters.

To remove this inhomogeneity, we must drop the first 5*P elements from (0.4). As mentioned previously, we should get convergence in estimates of total revisions only after $P*3$ revisions occur. This represents another form of inhomogeneity, which can be eliminated by dropping the last three years or 3*P elements in (0.4). Thus, our normalised and standardised revision vector $\mathbf{S}^{SA,P,T}$ will contain the T-8*P observations defined by:

$$\left\{\mathbf{S}^{SA,P,T} \mid S_t^{SA,P,T} \forall t = \left(5*P+1\right),\ldots,\left(T-3*P\right)\right\} \qquad (0.5)$$

The vector $\mathbf{S}^{SA,P,T}$ defines the set of standardised revisions for vector $\mathbf{Y}^{SA,\mathrm{P},\mathrm{T}}$ and we can consider the observed dispersion in $\mathbf{S}^{SA,P,T}$ a proxy measure for the size of the revisions caused by transformation F^{SA} .

Now let us define a second transformation F^{TC} that produces the final trend-cycle series (D12) of the X-12-ARIMA seasonal adjustment program. Then the revision vector for this transformation will be $\mathbf{S}^{TC,P,T}$. If we wish to establish which of the two transformation minimizes the revisions, then the objective becomes a statistical hypothesis test comparing the size of the dispersions for the two revision sets, $\mathbf{S}^{SA,P,T}$ and $\mathbf{S}^{TC,P,T}$. The size distribution of the observed revision datasets shows a heavy-tailed distribution. Thus a robust non-parametric test is required, such as Levene's test for the equality of variances.

The standard Levene's test would remove the mean from both $\mathbf{S}^{SA,P,T}$ and $\mathbf{S}^{TC,P,T}$. The absolute value would then be taken, which would convert the statistics from a measure of dispersion into a measure of location. Then a non-parametric t-test equivalent would be applied. Unfortunately, in the case of revisions, the mean of the revisions represents an undesirable bias or consistent over- or under-estimation of the final estimate. So for revisions, we do not want to subtract the mean. The true measure of dispersion should

[3] The treatment of extremes, trading day, and the length of filters chosen are based on the values of the full time series. Thus in theory, convergence to zero revision sizes may never occur.

include the undesirable bias. Let us define the mean of the absolute deviations of $\mathbf{S}^{SA,P,T}$ and $\mathbf{S}^{TC,P,T}$ as:

$$\dot{\bar{\mathbf{S}}}^{SA,P,T} = \frac{\sum_{t=5P+1}^{T-3P} \left| S_t^{SA,P,T} \right|}{(T-8P)} \quad (0.6)$$

$$\dot{\bar{\mathbf{S}}}^{TC,P,T} = \frac{\sum_{t=5P+1}^{T-3P} \left| S_t^{TC,P,T} \right|}{(T-8P)} \quad (0.7)$$

The revised Levene's test statistic then becomes:

$$W^{SA/TC} = \frac{(T-8\mathrm{P}-1)(T-8P)\left(\dot{\bar{\mathbf{S}}}^{SA,P,T} - \dot{\bar{\mathbf{S}}}^{TC,P,T}\right)^2}{\sum_{t=5P+1}^{T-3P}\left(\left(\left|S_t^{SA,P,T}\right| - \dot{\bar{\mathbf{S}}}^{SA,P,T}\right)^2 + \left(\left|S_t^{TC,P,T}\right| - \dot{\bar{\mathbf{S}}}^{TC,P,T}\right)^2\right)} \quad (0.8)$$

The significance of $W^{SA/TC}$ is tested against $F\left(\alpha, 1, (T-8P)\right)$, where F is a quantile of the F-test distribution with $\left(1, (T-8P)\right)$ degrees of freedom, and α is the chosen level of significance. The statistic $W^{A/B}$ is our standardised test statistic for comparing the revisions of transformation F^A versus transformation F^B. The average size of the improvement (or enhancement) in the revisions from using transformation F^A versus transformation F^B will be given by:

$$\Delta\dot{\bar{\mathbf{S}}}^{A/B} = \dot{\bar{\mathbf{S}}}^{A,P,T} - \dot{\bar{\mathbf{S}}}^{B,P,T} \quad (0.9)$$