The seasonal adjustment methodology applied to economic statistics by the South African Reserve Bank

by S Knox and M Smal

Introduction

Economic statistics are often influenced by seasonal variations which impact on the activity being measured. The seasonal adjustment of economic statistics removes the seasonal component in the measured activity to render historically comparative observations for analysis. The seasonal adjustment process is expected to treat and eliminate patterns in an economic time series that are usually repeated annually and mask underlying movements in a time series. Seasonal adjustment should not be applied in instances where a time series does not display any recurring seasonal pattern.

This article defines seasonality in economic statistics and presents the seasonal adjustment methodology applied by the Economic Statistics Department (ESD) of the South African Reserve Bank (SARB). All statistics published in the statistical tables section of the SARB's *Quarterly Bulletin* which are seasonally adjusted by the SARB are adjusted according to this methodology.

Time series decomposition

A time series is a sequence of observations recorded at regular, consecutive time intervals or frequencies such as daily, monthly, quarterly or annually. Time series analysis is performed to describe trends in a time series as well as relationships or correlations between different time series, and also to predict or forecast the future values of a time series. For economic analysis purposes, a time series can be decomposed into its separate components. An economic time series Y_t can be divided into the following four components:

$$Y_t = T_t \, . \, C_t \, . \, S_t \, . \, I_t$$

a Trend component (T_t):

The long-term movement that reflects the general direction of the time series.

b Cyclical component (C_t):

A sequence of repeated, but non-periodic, smooth fluctuations around the long-term trend, characterised by alternating periods of expansion and contraction. The duration of these fluctuations depends on the nature of the time series but, on average, these cycles last between three and five years, and generally coincide with the business cycle.

c Seasonal component (S_t):

Seasonal fluctuations that are repeated on a regular basis within a calendar year, with the timing, direction and magnitude more or less stable and predictable.

d Irregular component (I_t):

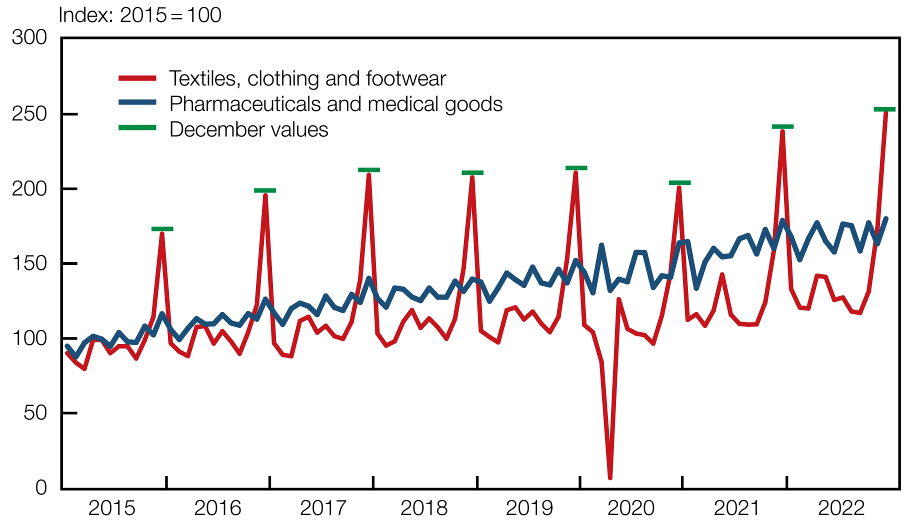
Irregular and unpredictable fluctuations not covered by the previous three components, thus representing the residual movement. This component is often referred to as 'white noise' and has a random character. Some irregularity is to be expected as no two periods will be exactly the same due to various factors. However, these movements should be relatively small and not dominant. Although the irregular component is sometimes mistakenly seen as outliers, the two are not the same; an irregular component is of a smaller magnitude than an outlier.

Seasonality and its causes

Seasonality refers to a pattern or movement in an economic times series that is consistently repeated at the same time (e.g. every month or quarter) and with more or less the same magnitude or intensity every calendar year. Some industrial sectors' economic activity is more seasonal than others', for instance in agriculture where weather and seasonal patterns affect the intensity of that industry's activities. Seasonality is influenced by factors other than the economy, and mostly stems from annual climate changes and conventions such as religious, social and public events which are repeated annually and have an impact on economic activity.

To illustrate a difference in seasonality, two subcategories of retail trade sales in South Africa are compared in Figure 1. Although the nominal value of both indicators increases every December, that of textiles, clothing and footwear increases to a greater extent than that of pharmaceuticals and medical goods. This shows that the seasonal impact of the annual December holidays and Festive Season is much larger on the former category.

Figure 1 Nominal retail trade sales (not seasonally adjusted)



Source: Stats SA

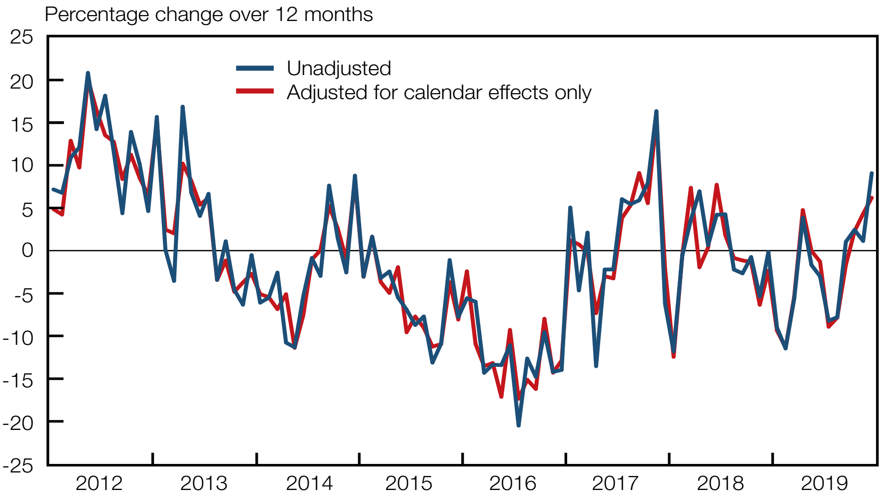
There are four major causes of seasonality in an economic time series:

a. seasons with colder or warmer weather impacting on, for example, the sales of heaters and the type of clothing apparel, or the planting and harvest times in the agricultural sector;
b. major institutional deadlines such as academic years or the end of a tax year;
c. expectations which result in an increase in certain activities occurring in a specific month of each year, such as the registration of new vehicles at the beginning of a new calendar year; and
d. the composition of the calendar, which includes the number of working days per month, Easter (which is a moving holiday period) and public holidays.

The first three causes are pure seasonal factors, while the fourth is referred to as calendar effects, which are treated separately and differently from distinct seasonal factors. Calendar effects are related to the number of working and trading days in a particular month or quarter, and include Easter and other public holidays which can have an impact on economic activity.

Seasonal adjustment is the process of smoothing data by means of statistical filters to isolate the seasonal effects with the purpose of identifying the underlying data patterns. Removing the seasonal and calendar effects assists in determining whether a time series value truly increased or decreased due to economic factors, or if it was just momentarily caused by an exogenous factor such as a flood or industrial action. Neglecting to adjust for calendar effects can lead to the incorrect interpretation of the movements in an economic time series. To circumvent seasonal adjustment, some economic analysts focus on the year-on-year growth rates in a time series. However, this approach ignores the possible calendar effects. To illustrate this point, the year-on-year rates of change in the unadjusted (actual) number of new passenger vehicle sales and the adjusted time series where only calendar effects are removed are shown in Figure 2. The latter is clearly different (exhibiting less volatility) than the unadjusted time series, emphasising the importance of removing calendar effects. *When year-on-year rates of change are calculated, the data should be adjusted for calendar effects but not for seasonal effects.*

Figure 2 Number of new passenger vehicle sales*



* Data only up to the end of 2019 in order to avoid COVID-19 lockdown distortions

Sources: NAAMSA and SARB

Outliers

Outliers are extreme values that can distort the seasonal adjustment of a time series as their magnitudes are larger than those of irregular movements. Outliers are either a single or a few consecutive abnormal data points in a time series which are caused by exogenous factors such as strikes, droughts, floods, pandemics and data processing errors. They can occur at different periods throughout the time series, and more than one type of outlier can be present in a specific time series. *Outliers should therefore be temporarily removed during the seasonal adjustment process and then added back to the seasonally adjusted series.* There are various types of outliers (see Table 1).

Table 1 Time series component classification of calendar effects and outliers

Description	Original series (Y_t)	Trend cycle (T_t / C_t)	Seasonality (S_t)	Irregular (I_t)	Seasonally adjusted series
Trading days*	X		X		
Easter*	X		X		
Other moving holidays*	X		X		
Additive outlier**	X			X	X
Transitory change**	X			X	X
Level shift**	X	X			X
Ramp**	X	X			X
Seasonal outlier/seasonal level shift**	X		X		
Re-allocation outlier**	X			X	X

* Calendar effects
** Outliers

The time series components under which each of the calendar effects and outliers is categorised are shown in Table 1. *Note that all outliers, except seasonal outliers and level shifts, should still be part of the seasonally adjusted time series.*

Seasonal adjustment software

A time series can be seasonally adjusted with various software packages, including EViews, JDemetra+[1], Python, R, SAS and Win-X. All of these packages follow the same basic methodology but differ in terms of options to select and refine the specifications and test results.

1 JDemetra+ is an open-source application available free of charge. It includes training manuals, guidelines and technical support from Eurostat. The latest version is JDemetra+ V2.2.4, available at https://ec.europa.eu/eurostat/cros/content/download_en.

The SARB uses JDemetra+, which was developed and released by Eurostat in 2012 to promote standardisation across European Union countries. This software consists of a collection of reusable and extensible Java components, and is regularly updated and improved by Eurostat with new functionalities.

JDemetra+ provides two methods of seasonal adjustment, namely X13 ARIMA-SEATS (X13) and TRAMO-SEATS. Initially, the X13 method was more widely used due to its better diagnostics and smaller revisions. However, recently TRAMO-SEATS has become more popular because it can accommodate a change in the seasonal pattern better than X13.

Statistical processes applied to seasonally adjust a time series

The first step in the seasonal adjustment process is to fit a RegArima (regression and autoregressive integrated moving average (ARIMA)) model to the time series, also known as the ‘pre-treatment’ of the time series. By applying a regression to the time series, outliers are identified and eliminated, while the series is also treated for calendar effects. The ARIMA function assists in obtaining estimated values for the pre- and post-sample periods, which are required to avoid the so-called ‘end-point problem’ encountered when trend and seasonal filters need to be fitted to the time series. The statistical tests for the ARIMA results are in the form of AIC (Akaike Information Criterion) and AICC (a second order estimate version of the AIC, which corrects for smaller sample sizes) values. The model with the lowest AIC/AICC value is preferred. During this process, the coefficients of the regression and ARIMA model are generated.

Following the ‘pre-treatment’ of the time series, an iterative process is applied to separate it into the different components (i.e. T_t, C_t, S_t and I_t). During this process, the time series that is obtained after the RegArima treatment is considered to be the ‘original’ time series.

Firstly, the trend and cyclical components are dealt with in the same step as they are both separated by applying moving averages to the original time series. In the first round, this is a centred 12-term moving average. In subsequent rounds, a Henderson filter of various lengths is used. These lengths are determined by the irregular-to-cyclical (I/C) ratio.

Following the removal of the trend and cyclical components, the seasonal component of the time series is isolated by using a seasonal filter which focuses on a specific month or quarter, spanning over the whole selected sample period. These can vary from a 3x3- to a 3x15-term filter. The more the unmodified seasonal components fluctuate from year to year, the broader the term filter will be. However, if this fluctuation is not a gradual shift but a relatively quick change, a narrower base is required to correctly record the actual seasonal influences, especially towards the latter part of the time series. Each month or quarter could be treated with its own filter specification. For instance, January could have a 3x5-term seasonal filter, while August could have a 3x9-term seasonal filter.

It is now possible to isolate the irregular component of the time series and adjust for it. This is done by calculating the standard deviation of the time series and multiplying it by 1.5 and 2.5 (the default sigma limits). If an original value of the time series is more than 2.5 standard deviations away from the mean, it is assigned the value of 0. If it falls between 1.5 and 2.5, the weight will vary between 0 and 1 depending on where it falls within the range. If the value is smaller than

1.5 times the standard deviation, the full observation is included. These sigma limits can be manually changed in JDemetra+ to make the specifications either more or less strict, and the weighting structure is then adjusted accordingly. In exceptional circumstances, judgement can be applied to deviate from the default limits to incorporate specialist subject-matter knowledge and experience of the specific time series.

Through this, the time series has now undergone its first round of filtering. The steps already described will be subject to three iterations. The first round of the iterative process uses fixed filters, and JDemetra+ generates the results presented in Tables B1 to B20.[2] The length of the filters used in the second and third rounds will be determined by various statistical tests. The second round's results are shown in Tables C1 to C20, while round three's results are indicated in Tables D1 to D18. After the third round of filtering, the seasonal factors have been finalised, and all the results will be available for evaluation.

The aforementioned process can be executed by either a multiplicative or an additive procedure, which will depend on the stationarity[3] of the time series. When a series is non-stationary, a multiplicative process is applied, while an additive process is applied for a stationary variance or to any time series that contains negative or zero values.

Test results

Various statistical tests are conducted to determine the quality of the seasonal adjustment results obtained. The most frequently used test results are listed below:

a Sliding span analysis:
One way to analyse the results of the seasonal adjustment process is to conduct a sliding span analysis for a time series that stretches over a long period of time. This is done over sequential time periods. These tests show how the seasonal factors change over time (moving seasonality) and how their month-to-month growth rates change.[4]

b Statistical tests of pre- and post-adjusted time series:
- auto-correlation at seasonal lags;
- Friedman (non-parametric);
- Kruskall-Wallis (non-parametric);
- spectral peaks;
- periodogram; and
- seasonal dummies.

More emphasis is placed on the results of the first two tests. If the pre-adjusted time series returns a 'YES' result and the post-adjusted time series a 'NO' result, there was definite seasonality in the original time series that has successfully been removed.

c Contribution of the seasonal component to total variation:
JDemetra+ produces an output table showing the relative contributions of the cyclical, seasonal, irregular, pre-adjustment (outliers) and calendar effect to the stationary portion of the variance in the original time series. The seasonal part should be the main contributor for the seasonality to have been successfully removed.

d M statistics (quality measures):
There are 11 M statistics generated that measure various aspects of the time series. The weighted average of these is called the Q-statistic, or quality measures. The value of the M statistics should be less than 1 for the results to be accepted. The M7 value, which tests the amount of moving seasonality relative to the amount of stable seasonality, is the most important of the individual test results. It also has the largest weight in the Q-statistic.

e Spectral analysis:
A graphical spectral analysis tool is also available in JDemetra+, with the spectral peaks indicating whether seasonality is present in the original time series.

2 The results of the seasonal adjustment process are presented in Tables B, C and D, as generated by JDemetra+.

3 Stationarity refers to the statistical properties of a time series. For a stationary time series, the mean, variance and auto-correlation structure do not change over time.

4 An adjustment process is seen as unstable if, for a specific month or quarter (e.g. every February or every first quarter), the maximum estimated value *minus* the minimum estimated value *divided by* the minimum value is larger than 0.03. These maximum and minimum values are used for all the time spans tested.

In most instances, all of the aforementioned test outcomes should indicate similar results on whether the adjustment is statistically acceptable or not. However, if there are contradicting test results, judgement and expert knowledge of the original time series will inform the final decision as to whether the results should be accepted or rejected.

Improving the test results

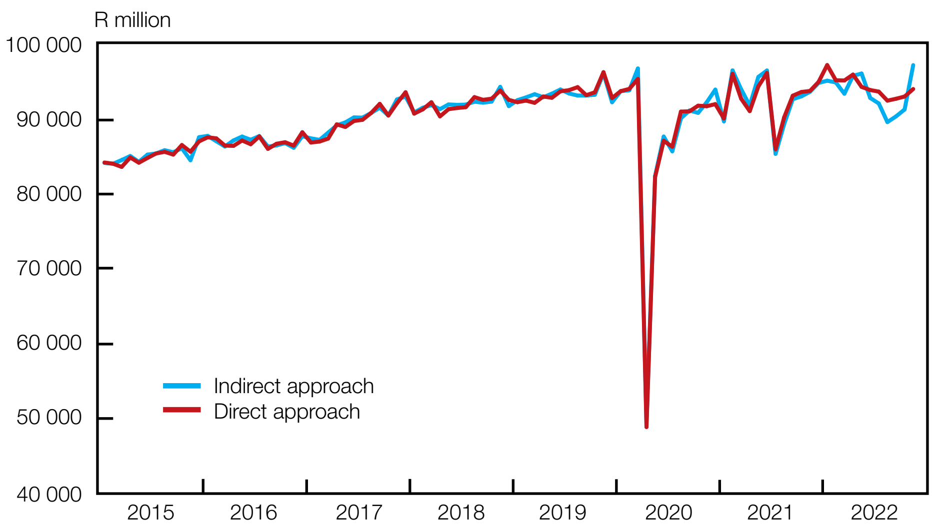
Various interventions are available in JDemetra+ to improve test results, including:

- changing the duration of the period that is subject to adjustment while considering structural breaks;
- changing the length of the Henderson filter;
- changing the length of the seasonal filters; and
- increasing or decreasing the sigma limits applied to the irregular observations.

Direct versus indirect seasonal adjustment

When the components of an economic time series are summed to a total (e.g. total retail trade sales and its subcomponents), the question arises whether the total should be adjusted independently (directly) or if it would be preferable to sum the various subcomponents' seasonally adjusted time series to obtain the seasonally adjusted total (indirectly). The decision depends on various factors, including the structure of the time series as well as the seasonal patterns of the individual time series. If they all have a similar pattern, the indirect approach would be preferable. However, the total should still be tested for the presence of residual seasonality and, if present, the total should rather be directly adjusted. As shown in Figure 3, the results can differ significantly between the direct and indirect methods.

Figure 3 Seasonally adjusted real retail trade sales



Sources: Stats SA and SARB

Annual totals

Due to possible calendar effects, the annual totals of the original time series and the sum of the monthly or quarterly seasonally adjusted time series might differ somewhat. Despite this difference, the outcome is still considered statistically sound. However, in some statistical areas (such as the national accounts and the balance of payments), it is imperative for the two time series versions to have the same annual totals as they form part of an accounting framework which is required to balance. In such instances, the time series is adjusted in such a manner that its values add up to the same annual total as the original time series. This is accomplished by activating a function which is referred to as benchmarking[5] on JDemetra+. However, this adjustment can also be executed manually.

5 This is not the same as the benchmarking of economic statistics discussed under the 'Revision policy' section below.

Revision policy

A seasonally adjusted time series should be revised either because of amendments to the original unadjusted data or because of a better estimation of the data due to new or additional information which has become available. Data-driven revisions can include methodological changes, the addition of new data, benchmarking, error corrections and annual data revisions. Methodological changes can include definition or classification changes, refinement, harmonisation and re-weighting.

To enhance the credibility and usefulness of economic statistics, it is important to have a clear revision practice and policy. A balance should be found between having the best possible seasonally adjusted statistics and avoiding frequent and/or insignificant data revisions. When and how often to revise a seasonally adjusted time series depends mostly on the nature of the revision. If there are major revisions to an original time series, the entire time series should be readjusted and revised, for example if subcomponents were re-weighted or if a sample change was introduced.

The SARB adheres to the following seasonal adjustment revision guidelines:

a When, and how regularly, to revise:

- for the addition of a new data point – once a year;
- for the revision of the underlying data – rerun the model in JDemetra+ if more than three data points are affected, otherwise once a year;
- for an annual sample change – when it is implemented;
- for a weight change – when it is implemented; and
- for benchmarking – when it is implemented.

b How far back are seasonal factors revised?

This depends mostly on the nature of the revision:

- for an annual update – not more than three years;
- for an underlying revision update – not more than three years;
- for a sample change – from the inception of the new sample;
- for re-weighting – from the inception of the new weights; and
- for benchmarking – from where the newly benchmarked unadjusted time series was revised.

Various revision options are available in JDemetra+, as listed in Table 2. The appropriate choice depends on the user or institution as well as the reason for, and frequency of, the revision. The SARB prefers to re-estimate the regression coefficients, in other words the partial concurrent adjustment. With this revision option, the model, filters, outliers and calendar regressors stay the same, and only the parameters and seasonal factors are re-estimated.

Table 2: Revision options in JDemetra+

Option	Meaning
Current adjustment (AO* approach)	Fix the model (including its parameters); handle any new observation as an AO*
Fixed model	Re-estimation, with parameters fixed
Estimated regression coefficients	Re-estimation of the regression coefficients
+ ARIMA parameters	and re-estimation of the parameters of the ARIMA model
+ last outliers	and re-estimation of the outliers of the last year only
+ all outliers	and re-estimation of all the outliers
+ ARIMA model	and re-estimation of the ARIMA model
Concurrent adjustment	The reference specification is used

* Additive outlier

Conclusion

Seasonally adjusted time series estimates greatly enhance the analysis of economic statistics by providing historically comparative and meaningful observations to inform policy decisions. The SARB adheres to international best practice in its seasonal adjustment process, supported by the use of the JDemetra+ software package and accompanying methodological guidelines developed by Eurostat, as applied by the European Statistical System, the Deutsche Bundesbank, Eurostat and various other international statistical offices.