Seasonal adjustment of CPIs during the COVID-19 pandemic and beyond

By Tom Arend, Jarmila Botev, Emmanuelle Guidetti, Annabelle Mourougane and Minsu Park [1]

Abstract

This paper examines the presence of seasonality in CPI in 36 OECD economies that provide monthly CPI data and reviews the properties of standard methods, namely X-13 and TRAMO-SEATS, in performing the adjustment. Evidence from statistical tests points to the presence of seasonality in headline CPI and its components, with stronger seasonality in some components, such as clothing and footwear. There are also indications of changes in seasonal pattern from 1980 to 2022, but it is not systematic across countries. Simulations suggest that differences between the two methods are small when applied to CPI in OECD countries in normal times. Differences between the direct (adjusting all-item CPI and components independently) and the indirect (aggregating the seasonally adjusted components) approaches are also minimal, limiting the need for a reconciliation method.

The paper further investigates whether large shocks, such as the COVID-19 pandemic, affect seasonality and how seasonal adjustment methods can accommodate them. Although large shocks should in theory affect seasonal adjustment, there is no strong evidence of a change in CPI seasonal patterns following the COVID-19 pandemic. This issue needs however to be revisited once the effects of the shock, including its impact on inflation have fully dissipated. The reason for this is that in times of a large shock, the type of outlier, that is usually added to account for the unusual variability, matters only after the impact of the shock has dissipated.

The extent of revisions implied by the seasonal adjustment should be among the criteria for choosing a seasonal adjustment method, as CPI is often used in indexation and legal documents. The paper provides a summary of how communication is handled by selected OECD countries and provides a list of best practices that can be drawn upon by a National Statistical Office aiming to publish seasonally adjusted CPI.

---

[1] Statistics and Smart Data Directorate, OECD. The paper benefited from comments from Benoît Arnaud, Christophe André, Anne-Sophie Fraisse, Pierre-Alain Pionnier, Paul Schreyer (all OECD) and Nick Johnston (IAE) and participants of seminar on times series analysis organised by the OECD-Eurostat-INSEE and TSACE on 14-15 December 2023.

Introduction

1. Adjusting Consumer Price Indices (CPIs) for seasonality is important to monitor recent underlying developments in inflation and to inform policymakers with clear and unambiguous signals. Adjusted series allow for comparability across months without the influence of seasonal fluctuations and are particularly useful in the conduct of monetary policy.

2. For a long time, CPIs were not “officially” seasonally adjusted. NSOs often communicated through year-on-year changes in the CPI, which are subject to a base effect. However, the practice has been changing with more and more National Statistical Offices (NSOs) now publishing seasonally adjusted CPIs. The most common approach is to use a filter and moving averages of historical data to estimate the seasonal pattern. CPIs are subsequently adjusted by removing the within-year seasonal movements from the time series. Additivity constraints between all-items CPI and its components also need to be accounted for using either indirect adjustment (aggregating the seasonally adjusted components) or direct adjustment (adjusting all-item CPI and components independently) with a reconciliation method. The use of different aggregation methods to calculate higher-level indices in the CPI also exacerbates the difficulties of adjusting CPI for seasonality in an international context.

3. The COVID-19 crisis has compounded the traditional difficulties in adjusting CPI for seasonal factors, as it triggered extreme volatility, altered seasonality test statistics and distorted the observed seasonal patterns. Several mitigation steps have been employed in NSOs to address these issues including identifying price indexes whose movements were affected by the pandemic, estimating time series models to quantify these effects and removing pandemic-related price movements from the data before estimating seasonal patterns. Beyond the pandemic, some of these steps can also be applied in case of extreme non-seasonal events such as natural disasters or wars which can also distort the underlying seasonal pattern of an index.

4. The paper reviews the main challenges in adjusting CPIs in the aftermath of the COVID-19 crisis and beyond and discusses the pros and cons of the most common methods. It relies extensively on selected NSO experience and simulations and quantifies differences across methods in a cross-country setting.

5. The main insights from the paper are as follows:

- There is evidence of seasonality in the raw CPI data, headline and components, but it is more marked for some specific categories, such as “Clothing and Footwear”, and to a lesser extent “Recreation and Culture, Education and Restaurants”.
- The seasonal pattern is found to have changed in more than half of the countries covered since 1980, with most of the time an increase in the magnitude of the seasonal fluctuations. But no discernible change was observed for the remaining smaller half.
- Although large shocks should in theory affect seasonal adjustment, there is no strong evidence of a change in CPI seasonal patterns following the COVID-19 pandemic. This issue needs however to be revisited once the effects of the shock, including its impact on inflation have fully dissipated. Indeed, in times of a large shock, the type of outlier, that is usually added to account for the unusual variability, matters only after the impact of the shock has dissipated.
- Most NSOs which publish seasonally adjusted CPIs use X-12/X-13 ARIMA and a direct approach. Many re-estimate the model each time the adjustment is performed.
- Simulations suggest that differences between X-13 and TRAMO-SEATS in adjusting CPI for seasonality are small in OECD countries in normal times. Differences between the direct and the indirect approaches are also minimal, limiting the need for a reconciliation method. The combined test for seasonality no longer detects the presence of seasonality after adjustment. TRAMO-

SEATS is also found to detect seasonality more often than X-13 when the sample size is long, and less often when the sample size is short.

- Publishing seasonally adjusted CPI faces two major difficulties: revisions may lower the credibility of the measure, at a time when official measures point to lower inflation than those perceived by households. Revisions may also be problematic when CPI is used for indexation or legal contracts. Simulation suggests that those revisions are likely to be moderate, on average. Looking at an extreme case of revisions, they appear to be larger for TRAMO-SEATS than X-13 ARIMA, although the size of the revisions is overall small. However, X-13 tends to detect more outliers as the length of the series increases, which is likely to lead to more revisions.
- Publication should nonetheless be accompanied by good communication, including alerting users not to use seasonally adjusted CPI for indexation purposes or in legal contracts.

6. The paper is organised as follows. The next section provides a descriptive analysis and presents tests to detect seasonality in CPI and its components in 36 OECD countries that provide monthly data. Section 3 describes the methods most commonly used by NSOs. Section 4 compares the most common approaches using simulations in normal times and in Section 5 in the event of a large shock. Section 6 examines the potential for more experimental approaches. Section 7 discusses what type of communication and guidance should be provided to users of adjusted CPI series. The last section concludes.

1. Is there seasonality in consumer prices?

7. The presence and importance of seasonality in the CPI has been an open question for years. Although CPI was traditionally not adjusted for seasonality by most statistical offices, there is evidence that such seasonality exists in some cases.

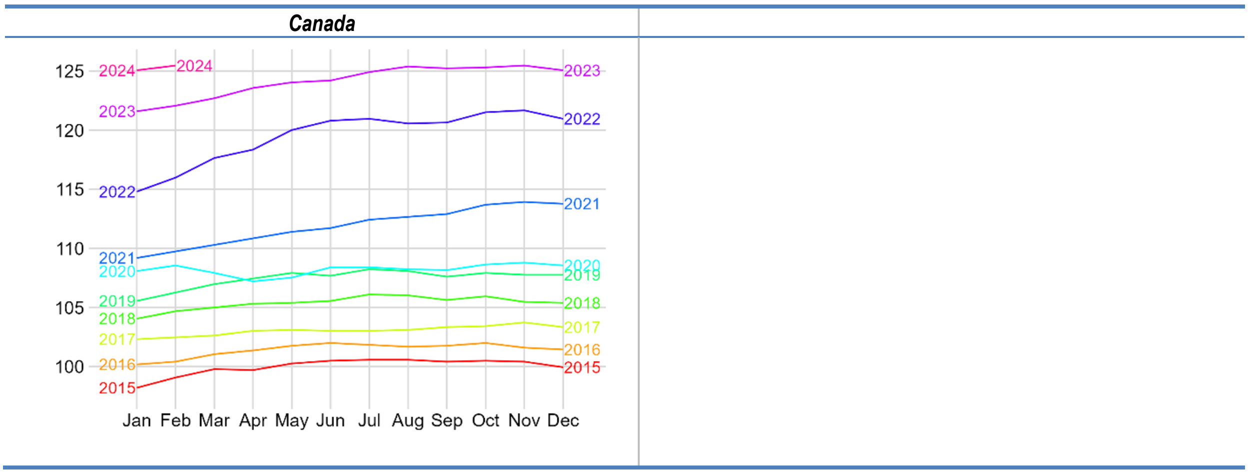
Seasonality of CPI appears to be item-specific and essentially idiosyncratic in nature, making its identification challenging

8. The existence of seasonality in prices has been a topic of research for years. Riley (1961) reported that the consumer price index from the US Bureau of Labor Statistics contains seasonality coming from both the demand side and the supply side due to climatic effects, changes in agricultural and industrial productions. Baxter (1999) showed that prices are subject to seasonality due to climatic conditions, taxes levied and price increases in certain sectors (e.g. transport) at a specific date in the year. In Europe, the ECB found the presence of seasonality in headline HICP, the foods and goods components but not in industrial goods and energy (ECB, 2000).

9. More recent evidence suggests headline CPI exhibits since 2015 some seasonal patterns only in about two-thirds of OECD countries, including Switzerland, Germany, Spain, France, Greece, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Slovenia, the United States (see Figure 1 for G7 countries, and Annex 1 for the rest of the OECD countries). By contrast, no such regularity was observed in Japan, where direction of changes from March to April could vary from one year to another (e.g. 2016 and 2019). In many countries seasonal patterns, when they exist, do not appear to be very pronounced.

Figure 1. Seasonality in headline CPI in G7 countries

Index = 100 in 2015

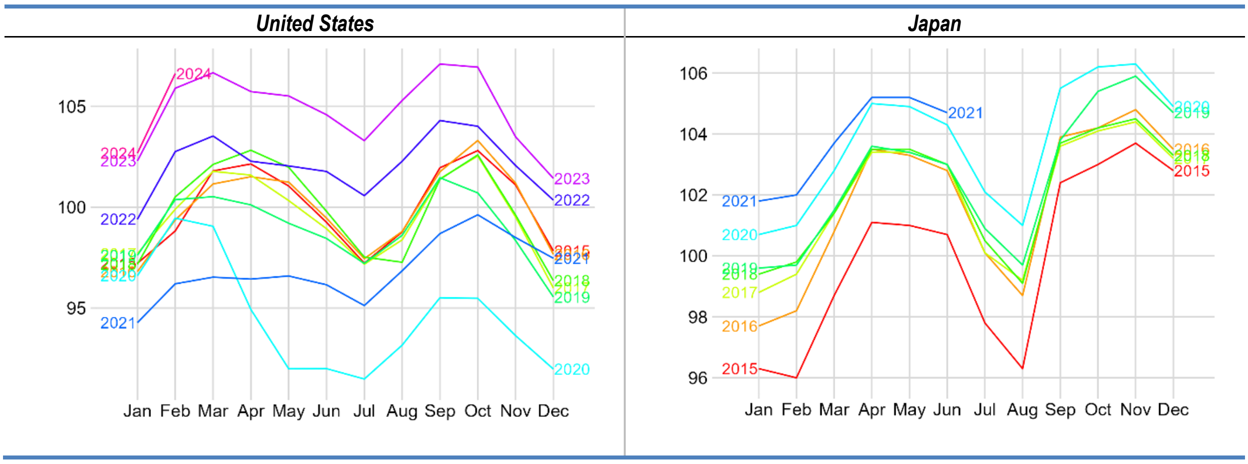


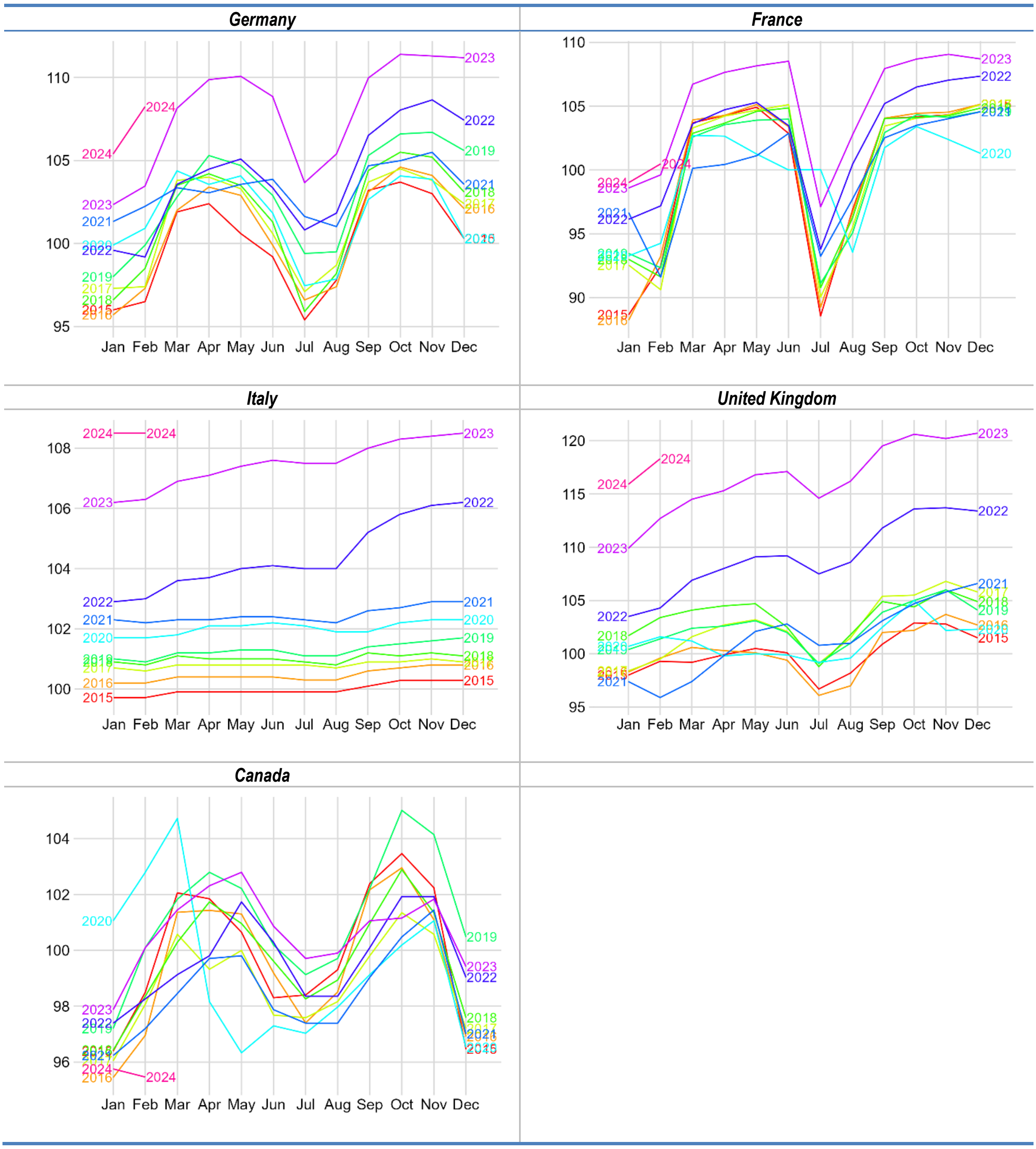
Note: Japan's CPI is based on COICOP 2018 classification, other countries are based on COICOP 1999.
Source: OECD CPI database.

10. However, some seasonal patterns can be observed at a more disaggregated level for specific categories. There is marked evidence of seasonality in "Clothing and Footwear", and to a lesser extent "Recreation and Culture, Education and Restaurants" (Figure 2). In the other categories no seasonal pattern is visible, with some isolated exceptions for some countries. Seasonality also appears to be more frequent in some countries such as Norway or France.

Figure 2. CPI "Clothing and footwear" in G7 economies

Index = 100 in 2015



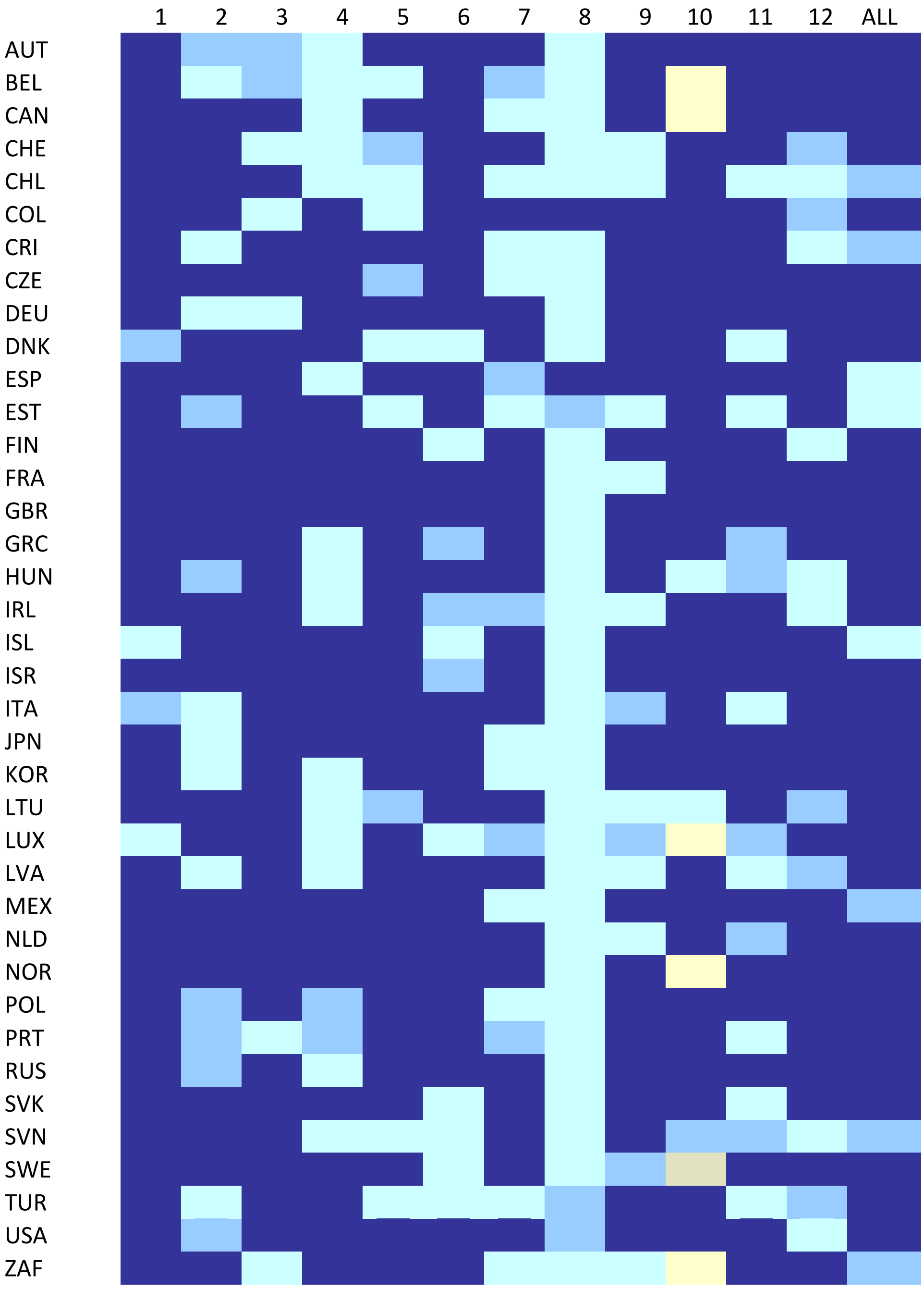


Note: Japan’ s CPI is based on COICOP 2018 classification, other countries are based on COICOP 1999.
Source: OECD CPI database.

Statistical tests point to the presence of seasonality in CPI and its components

11. JDemetra+'s combined test has detected the presence of seasonality in CPI and components in most of the 36 countries covered in the analysis (Table 1, see Annex A for a description of the test and Annex B for a list of COICOP categories).

Table 1. Combined test of CPI seasonality by COICOP categories



Note: Dark blue: means "passed the combined test and detected the presence of seasonality", blue means detected the presence of "possible seasonality", light blue means seasonality not detected. Yellow is used when the same value is used every month for a year for a subcategory of CPI, which makes it impossible to detect seasonality in those series. See Annex C for the list of COICOP categories.
Source: Authors' calculations using the OECD CPI database.

12. The test points to the presence of seasonality in headline CPI in all OECD countries which publish monthly indicators, but Island, Estonia and Spain. In Chile, Mexico, Costa Rica, Slovenia and South Africa, only weak signals of seasonality could be found.

13. At the disaggregated levels, the picture is mixed. Seasonality is detected for most categories, but the number of countries which do not exhibit seasonality in prices is generally higher than for headline CPI. Two categories stand out. There is no sign of seasonality in the information and telecommunication category in most countries, and in the housing, water, electricity, gas and other fuels category in about half of the countries.

Seasonal fluctuations can vary over time

14. Lis and Porqueddu (2018) noticed that seasonal fluctuations in the euro area core HICP (i.e. excluding food and energy) have become more pronounced over time. The seasonality impact on clothing prices has become substantially larger since 2001. Part of this reflects methodological changes, such as enhanced price collection, improvements in methods for compiling price changes in winter and summer clothing and, since 2011, the introduction of the EU regulation on the treatment of seasonal products which harmonises the treatment across countries (Box 1). This may not be reflected in national CPIs, though.

Box 1. Seasonal products in CPI

CPI measures the change of price of a representative basket of goods and services of constant quality purchased by households. To provide meaningful comparison over time, the products and services in the basket should be available for a reasonably long period, the items typically should not change from month to month, nor should there be a significant change in quality, or if there is, it should be accounted for.

The inclusion of seasonal products in the CPI can thus pose challenges for compilation and introduce seasonal fluctuations into the CPI. Their supply and/or demand regularly varies over the year, usually due to climate, traditions or institutional arrangements (e.g. fresh fruits and vegetables, seasonal clothing, water, electricity and fuels). They are not uniform across countries, or even within countries. *Weakly seasonal products* are available throughout the year, but their availability and prices fluctuate significantly over time. This can introduce fluctuations or “noise” in the resulting CPI, but usually does not require special treatment in its compilation. *Strongly seasonal products* are only available part of the year, when in season. Consequently, it is not possible to compute price relatives for these products for periods when their prices are unavailable, and a special treatment of these items in the CPI compilation is required. The Manual on CPI Concepts and Methods (2020) recommends one of two methods to include strongly seasonal items into the CPI (ILO et al., 2020):

i. **fixed-weight approach** that assigns the same fixed weight in the basket to the product in all months of the year and imputes or carries forward its prices in the out-of-season months, e.g. based on the last observed price or a “typical” price. In this case the monthly CPI variations are affected by the chosen method of imputation.

   This approach is used by all G7 countries, many European countries (e.g. Austria, Norway, Greece, Switzerland), and for the HICP in EU countries since the introduction of the regulation that harmonises the treatment of seasonal product in 2011.

ii. **seasonal-weight approach** where the weights for the product are fixed for in-season periods and set to zero in out-of-season periods. Consequently, the respective CPI elementary aggregate’s weights are redistributed to the other available items. In this case, the resulting CPI is distorted by including not only price changes, but also quantity changes between periods. This approach is used e.g. by Finland.

However, there is no ideal approach to treating seasonal products in the CPI compilation, and none of the recommended methods eliminates seasonal patterns in the resulting CPI. Alternatives such as the maximum overlap approach, which only uses items whose prices are available over two adjacent periods, presents similar limitations and additional drawbacks.

An example of the impact of different approaches to treatment of seasonal products

In 2019 Germany switched its HICP approach for package holidays for data from 2015 onwards from a seasonal-weight approach to fixed weights with out-of-season prices imputed using price changes of other trips. In addition, as of 2015, winter and summer holidays were treated in an integrated manner, rather than separately. In both samples, package holidays exhibit a strong seasonal pattern. However, the pattern became more pronounced with the methodological change from 2015 onwards and led to a revision of about 0.2-0.3 percentage point in the 2015 year-on-year inflation rates for services, core HICP, as well as headline HICP for Germany (Eigelspelgler, 2019).

Source:
Eigelspelgler,(2019), "A new method for the package holiday price index in Germany and its impact on HICP inflation rates", In: Economic Bulletin, Issue 2/19, European Central Bank; ILO, IMF, OECD, EU, UN, WB (2020), "Consumer price index manual: concepts and methods", Washington, DC: International Monetary Fund.

15. Looking at a larger set of OECD countries, there is evidence of an increase in the magnitude of the CPI seasonal patterns over time, in the period 1980 to 2022 in a number of European countries, the United States and Canada (Figure 3, Table 2). However, in a handful of countries, the magnitude of those fluctuations has diminished, for instance in Japan. No clear pattern or no change is discernible in other countries, in particular in many Eastern European countries and Nordic countries, but also in Latin American and Asian economies.

16. There is no strong evidence that seasonal fluctuations changed at the start of the pandemic (see Figure 4 for an illustration in the G7 countries). In a number of European countries, the start of the war in Ukraine and the ensuing effect on energy prices, have had a marked effect on the irregular component of seasonality.

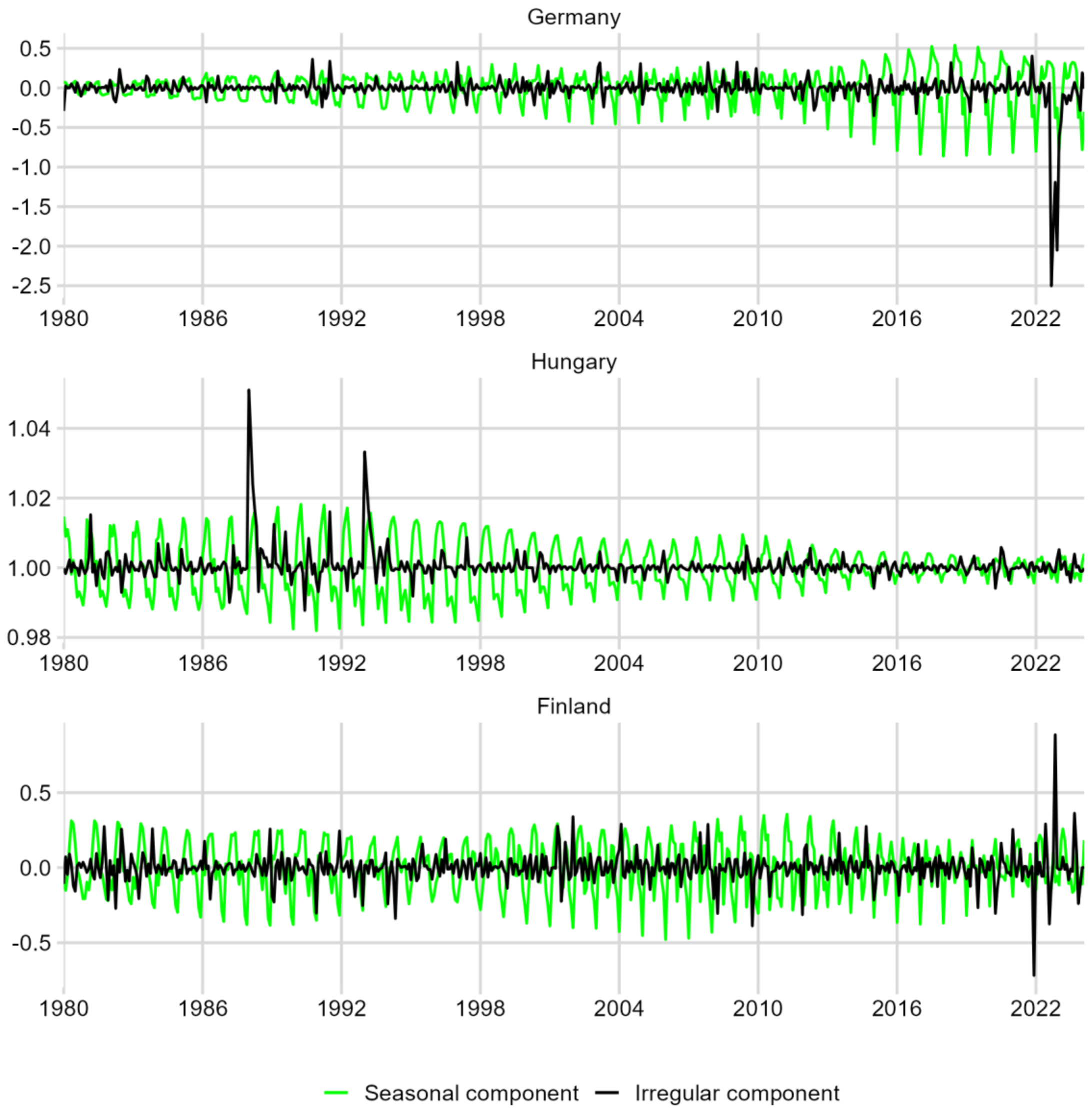
Table 2. Change in seasonal patterns from 1980 to 2022

	Increase	Decrease	No change or no clear pattern
Europe	BEL, DNK, EST, FRA, GER, IRL, ITA, LUX, NLD, NOR, ESP, CHE, SVK, SWE, GBR	GRC, HUN, PRT	CZE, ISL, FIN, LAT, LTH, SVL, POL, TUR
Asia-Oceania	KOR	JPN,	AUS, NZL,
America	USA, CAN	COL	CHL, CRI, MEX
Africa-Middle East			ISR, ZAF

Source: OECD CPI database and OECD calculations.

Figure 3. Changes in seasonal patterns from 1980 to 2022

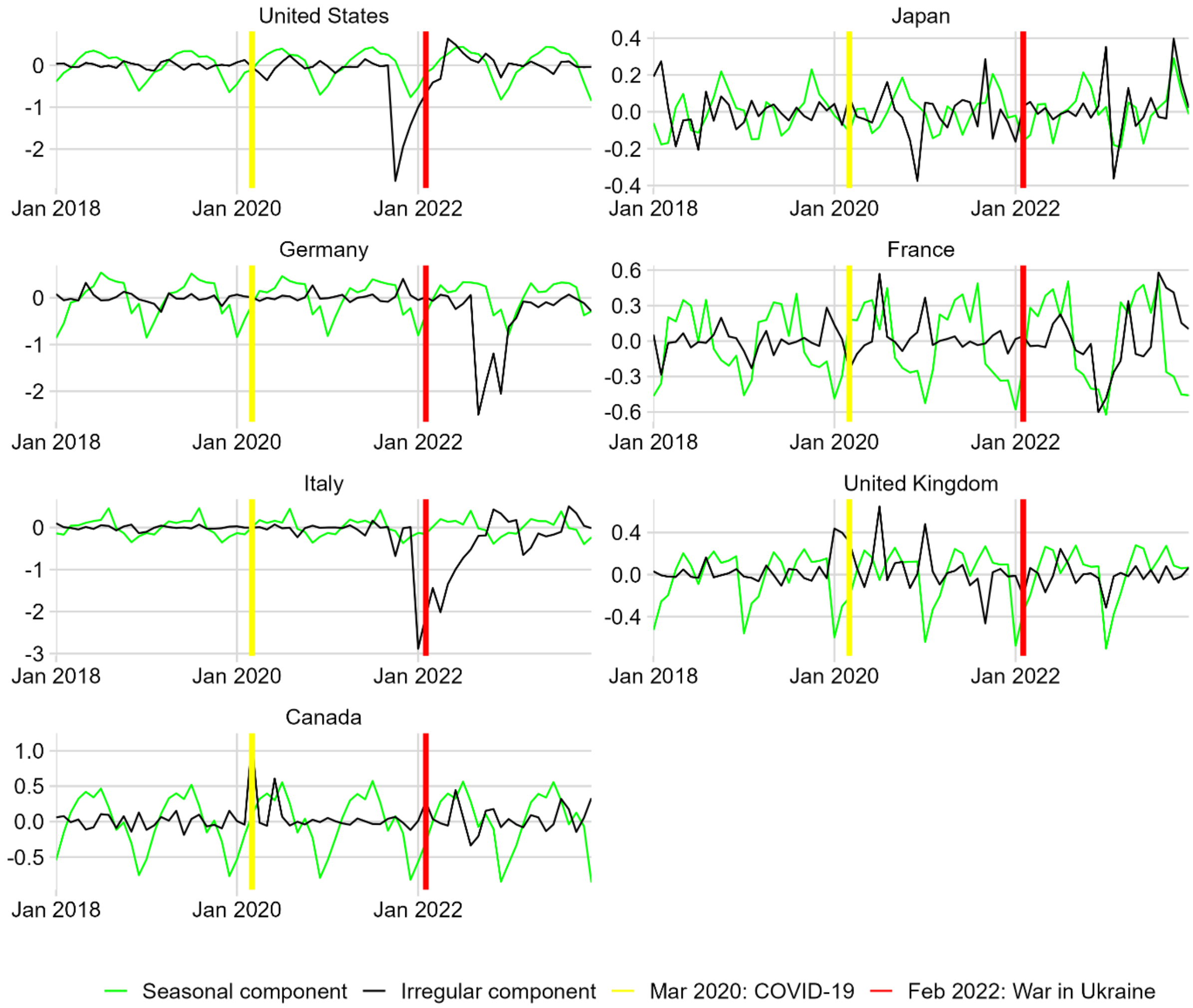
Multiplicative factors applied to the seasonal decomposition



Note: The green line depicts the seasonal component and the black line the irregular component of the headline CPI series.
Source: Authors' calculations using the OECD CPI database.

Figure 4. Seasonal patterns in headline CPI at the start of the pandemic

Multiplicative factors applied to the seasonal decomposition



Note: The yellow line in January gives a reference point but does not necessarily coincide with the start of the COVID-19 pandemic which was country specific.
Source: Authors' calculations using the OECD CPI database.

2. How do NSOs and Central Banks adjust CPI for seasonality?

17. At the time of writing seven OECD countries published seasonally adjusted CPI: Australia, Canada, France, Germany, Japan, Norway and the United States (Table 3). The ECB publishes seasonally adjusted Harmonised Index of Consumer Prices (HICP) for European countries. X-12 ARIMA is the most commonly used method (see Annex B for a short description). The United States and Mexico employ a very similar method based on X-13 ARIMA. Germany uses a specific adjustment relying on regressions, BV.4, which allows to separate calendar effects and extreme values from the residual.

18. Most countries make the adjustment on headline CPI (direct method) rather than on the components (indirect method). Australia, the United States and the ECB are the exceptions to this rule. According to the Eurostat Handbook on Seasonal Adjustment (2018), the direct method should be preferred in case where all the components exhibit similar seasonality patterns. Visual inspections

presented in the previous section suggests some categories exhibit very specific pattern, but their weights in CPI is limited. "Clothing and Footwear" make up about 4% and "Recreation and Culture" 9% of total on average in G7 economies in 2022. Illustrative simulations presented below seek to quantify the differences between the two approaches in OECD countries.

19. About two-thirds of the countries use a concurrent adjustment, whereby the seasonal models are re-estimated each time a new observation becomes available, and one third of the countries a current adjustment method where the models are not re-estimated systematically.

Table 3. Methods employed by NSOs and Central Banks to seasonally adjust CPIs

	Method	Direct or indirect	Special adjustment for the COVID crisis
Australia	X-12-ARIMA, with a *concurrent adjustment* implemented when a new quarter is available to estimate seasonal factors for the current and previous quarters and derive the combined adjustment factors for the previous quarter and for the same quarter in the preceding year.	Indirect	yes
Canada	X-12-ARIMA with a *current adjustment* where each month, the previous month's seasonally adjusted index is subject to revision. The seasonally adjusted values for the last three years are revised with every January data release. At the same time, the models used to obtain seasonally adjusted data and their parameters are reviewed and updated when necessary.	Direct	
France	X-12-ARIMA with a *current adjustment*. Seasonal factors are calculated at the beginning of the year (for the current year and for the past).	Direct	yes
Germany	BV4.1 which decomposes the series into a trend-cycle, a seasonal, a calendar-effect, and a residual component, the last of which may include a few extreme values. The first mentioned two components are estimated by moving filter applications derived from approximating functions by a regression approach. The appropriateness of the filters is judged and controlled by their transforms into the frequency domain. A *concurrent adjustment* is used.	Direct	
Japan	X-12-ARIMA with a *current adjustment*.	Direct	
Mexico	X-13-ARIMA-SEATS with a *concurrent adjustment* implemented when a new data is available. The model for seasonally adjusting the series is typically revised by Banco de México once a year.	Direct	
Norway	X-12-ARIMA, with concurrent adjustment	Direct	
United States	X-13-ARIMA-SEATS with a *current adjustment*: CPI index series are adjusted using the multiplicative model. Seasonal factors are updated annually. Each year in February, BLS recalculates and publishes revised seasonally adjusted indexes for the previous five years. Seasonally adjusted indexes become final in the last and 5th year of revision. Each January, the seasonal pattern of index series is re-evaluated and status can move from seasonally adjusted to not adjusted, or vice versa.	Indirect	yes
ECB	X-12 ARIMA with a current adjustment. The adjustment is based on a multiplicative Airline model (i.e. ARIMA (0,1,1) model). The overall HICP index is the aggregation of four seasonally adjusted components (unprocessed food, processed food, non-energy industrial goods and services) and one non-adjusted component (energy).	Indirect	yes

Note: Concurrent adjustment means that every time a new observation is available, the model, filters, outliers, regression variables and transformation type are re-identified and the corresponding coefficients and factors are re-estimated, while current adjustment strategy means that the ARIMA model, outliers and other regression variables are not re-identified and the values of the associated coefficients are fixed. The transformation type also remains unchanged. (see https://jdemetradocumentation.github.io/JDemetra-documentation/pages/case-studies/revision-ao.html).

Source: Authors' compilation using NSO and Central Bank websites.

3. How well do standard seasonal adjustment methods perform in normal times?

20. There exists a breadth of seasonal adjustment procedures, with newer methods being developed (see Annex B). As of yet clear indications on which of these methods produce the best outcomes when applied to consumer prices in OECD countries is lacking. The objective of this section is to examine the performance of the two most common procedures, X-13 and TRAMO-SEATS in adjusting monthly CPI, in a variety of simulated and observational settings. Data spans from January 1990 to December 2022 for 36 countries, while data for Japan and Costa Rica are considered up to July 2021 and February 2022 respectively, when they switched their classification scheme to COICOP 2018. Tests are run using the R implementation of the JDemetra+ software.

21. The analysis is restricted to X-13, rather than its conceptually similar precursor versions, X-11 and X-12, as the former improves on early versions in important ways (US Census Bureau, 2023). In particular, X-13 includes not only the enhanced X-11 seasonal adjustment procedure but also the capability to generate ARIMA model-based seasonal adjustment using a version of the SEATS procedure. Newer, more complicated approaches such as STL, CAMPLET and CiSSA only provide marginal improvements over either the X-13 or TRAMO-SEATS procedures (Cleveland et al., 1990; Abeln and Jacobs, 2015; Bógalo, 2021). Only when computational resources are not a consideration and user input should be limited – circumstances that are unlikely to consistently arise in practice – do more involved and flexible procedures like STR or STL outperform the industry-standard (Cleveland et al., 1990; Dokumentov, A. and R. J. Hyndman, 2015; Ollech, 2018). The new STAHL method developed by Quantcube is also not considered (Daniel, Haller and Bellone, forthcoming). Although this method would prevent revisions in the seasonally adjusted CPI, the algorithm is not open source.

Simulation

22. The fundamental problem of testing and comparing the adjustments based on different approaches lies in the non-observable nature of the seasonal, trend-cycle and irregular components. The different algorithms identify seasonal patterns but harbour a certain level of uncertainty that makes knowing the true scope of the seasonal pattern and the trend-cycle impossible (Manski, 2015).

23. To have an objective benchmark against which to compare alternative adjustment methods, benchmark series are generated using a simulation-based approach commonly used in the literature (Ollech and Webel, 2020; Bógalo et al., 2021; Tiller and Evans, 2014; Durbin and Koopman, 2012). Simulated data that resemble the CPI are generated using a statistical process and calibrated to fit the characteristics of Consumer Price Indices in OECD countries. The performance of the most common default implementations of X-13 and TRAMO-SEATS are assessed against this benchmark. Detailed results are presented in Annex C. The seasonal adjustment is performed mechanically, first using the most common default applications embedded in the JDemetra+ software package and subsequently exploring other options. In all the tests, model performance is assessed using the mean absolute error (MAE), the root mean squared error (RMSE) and the mean absolute percentage error (MAPE).

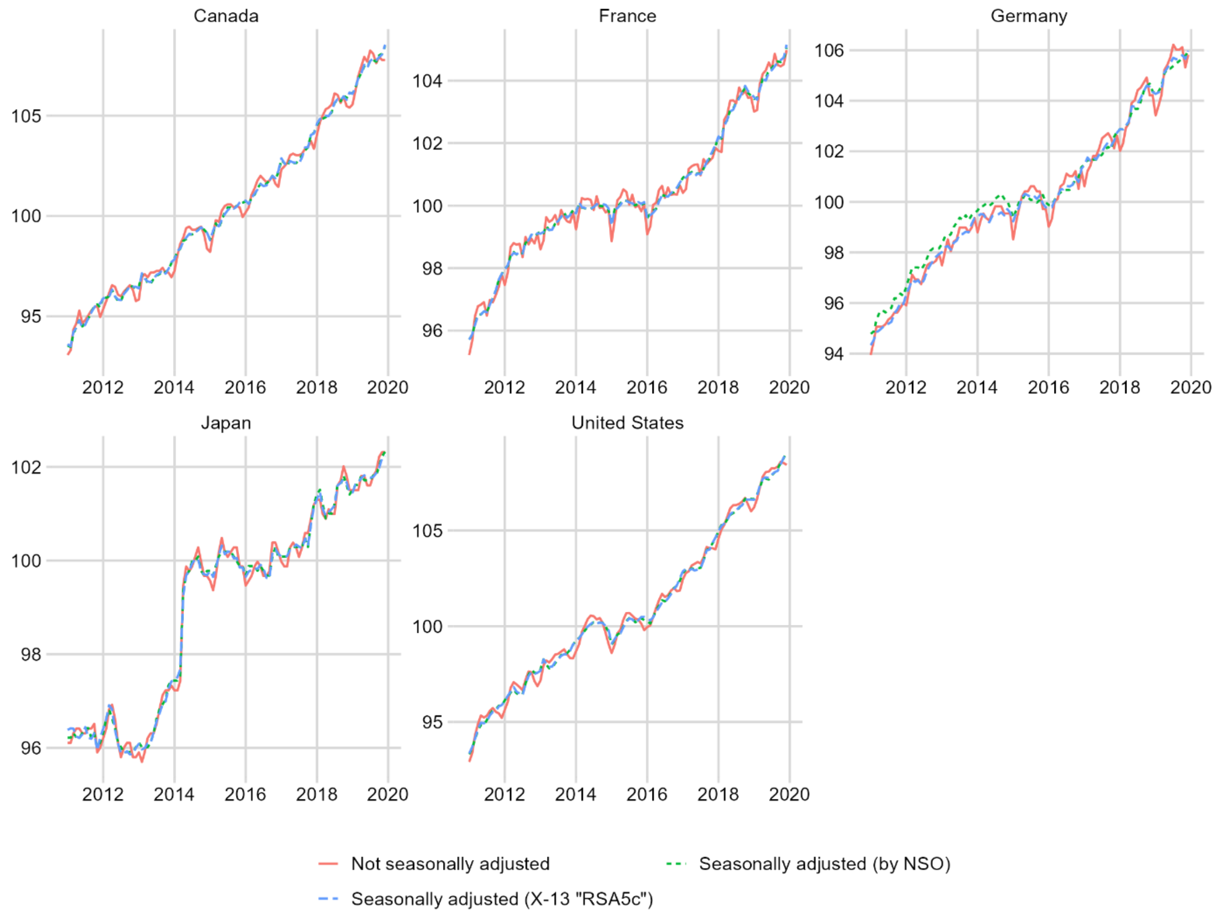
24. Overall, even though TRAMO-SEATS marginally outperforms X-13, there seems to be little actual difference in performance between the two methods. The small MAPE values (less than 1 percentage point) indicate that both methods are well suited to perform seasonal adjustment. In terms of calendar effects, explicitly accounting for the existence of an Easter effect can improve the performance of seasonal adjustment procedures, if Easter is present in the data. It should be noted that TRAMO-SEATS offers more options to deal with potential Easter effects in CPI series, which explains why it marginally outperforms X-13. Tweaking the default specifications with other user-defined inputs does not markedly improve the quality of the seasonal adjustment.

Observational approach

25. Simulations have the advantage of providing a general and objective test of the performance of seasonal adjustment methods. By design, they cannot, however, take into account the national specificities of observed time series. In practice, NSOs often tweak seasonal adjustment procedures to account for national specificities, for example, by providing user-defined regressors for specific national holidays.

Figure 5. Seasonal adjustment using X-13 versus seasonal adjustment performed by NSOs

Index = 100 in 2015



Note: Red is the non-seasonally adjusted time series. The dashed blue line represents the NSO adjusted time series and the green line represents our calculations using JDemetra+. Seasonal adjustment is performed on the entirety of the time series (1980-2022). To improve visibility only a subsection of the time period (2012-2020) is shown here. The main message would not be changed if a longer time period was shown.
Source: Authors' calculations using the OECD CPI database.

26. To quantify the improvements that specific national specifications engender compared to the broader cross-national specifications, the X-13 and TRAMO_SEAT default specifications are applied to the actual CPI for those countries that also publish seasonally adjusted data (Australia, Canada, France,

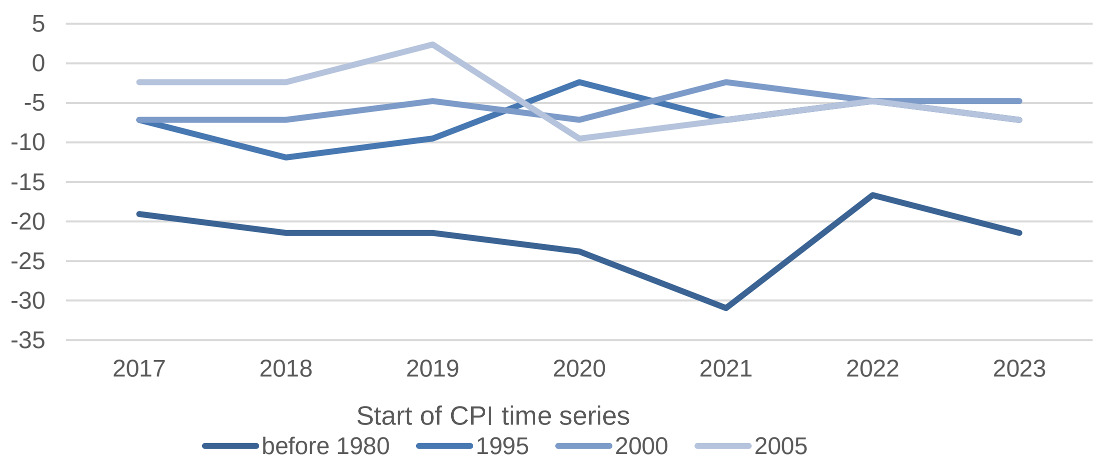
Germany, Japan and the United States).[2] These calculated seasonally adjusted values are then compared to the seasonally adjusted series published by NSOs.

27. The difference between the default procedures embedded in JDemetra+ and the seasonally adjusted data provided by the NSOs is exceedingly small (Figure 5). The correlation between NSO seasonally adjusted time series and OECD seasonal adjustment (TRAMO-SEATS) adjusted time series is unity or very close to unity (Table AC 16). The only exception is Germany where the correlation is lower, reflecting the use of a different procedure called BV.4 in this country.

28. Where differences between X-13 and TRAMO-SEATS appear, however, is in their ability to detect seasonality for different time series lengths. When taking all CPI observations into account, X-13 detects seasonality in 2 to 3% fewer cases than TRAMO-SEATS. By contrast, when time series are shorter, TRAMO-SEATS finds seasonality on average in 2% less cases (Figure 6). This confirms findings by Webel (2016) that X-13 underperforms TRAMO-SEATS when longer time series (more than 12 years) are considered but the former procedure outperforms TRAMO-SEATS for shorter time series.

Figure 6. Comparison of seasonality detection between X-13 and TRAMO-SEATS over different time series lengths

Difference in detected seasonality, per cent of all-item country CPIs



Note: Negative values mean that TRAMO-SEATS performed better at detecting seasonality in headline CPI. To increase visibility, only the period from 2017 to 2023 is shown here, but the calculations are conducted on the full extent of the time series.
Source: Authors' calculations using the OECD's CPI database.

Indirect versus direct seasonal adjustment

29. This section investigates the extent to which applying the adjustment directly or indirectly makes a substantive difference in the final output. In the direct approach, headline CPI and components are

[2] Although the Bank of Mexico performs seasonal adjustment for purposes of conducting monetary policy, these seasonally adjusted CPI series are not made public and thus not included in the analysis here. Norway is not included in simulations, as the seasonally adjusted CPI by Statistics Norway was not available in the OECD CPI database at the time of writing.

seasonally adjusted directly. In the indirect approach, the seasonally adjusted estimate for a time series is derived by combining the estimates for two or more directly adjusted component series (Eurostat, 2015).

30. According to the Eurostat Handbook on Seasonal Adjustment (2018), the direct approach is preferred for transparency and accuracy, especially when the component series show similar seasonal patterns. The indirect approach is preferred when components series show seasonal patterns differing in a significant way, including when certain components do not present any seasonality. This distinction is relevant for CPI, which is often broken down into subcomponents to better understand underlying price trends. As seen in Section 2, there are potentially significant differences in the seasonal patterns of CPI subcomponent series, even though the two categories that display a strong seasonal pattern have a relatively small weight in the headline CPI.

31. Yet past evidence also suggests that adjustment using the indirect methods may not be sufficient to fully remove seasonality. Biehl and Judd (1993) found that some seasonality remained in the headline CPI adjusted by the US Bureau of Labor using the indirect method. They put forward two possible reasons for this residual seasonality: a possible change in seasonality over time and difficulties to estimate seasonal patterns accurately when components are very volatile. They concluded that a direct seasonal adjustment would be better suited. This is in line with Bryan and Cecchetti (1993) who also recommend to directly apply seasonally adjustment to the higher-level CPI series. More recently, Peneva (2014) found evidence of residual seasonality in the US CPI, even though it is small relative to the variance of quarterly inflation movements.

32. The calculation is performed for 36 OECD countries for the period 2010 to 2021, as the CPI subcomponents are only available from 2009 onwards for the United States. Countries use different aggregation methods to calculate higher-level indices in the CPI, which complicates a cross-national comparison, and exacerbates the difficulties of adjusting CPI for seasonality in an international context. Therefore, in this exercise, the All items CPI from the 12 lower-level COICOP (1999 Classification) price indices is recomputed using a Lowe index and expenditure weights (see Annex C). This is the most common aggregation method according to the Manual on CPI Concepts and Methods (2020) (IMF et al., 2020). The month-on-month changes produced by the manually aggregated All items CPIs to those figures provided by NSOs, display very similar patterns.

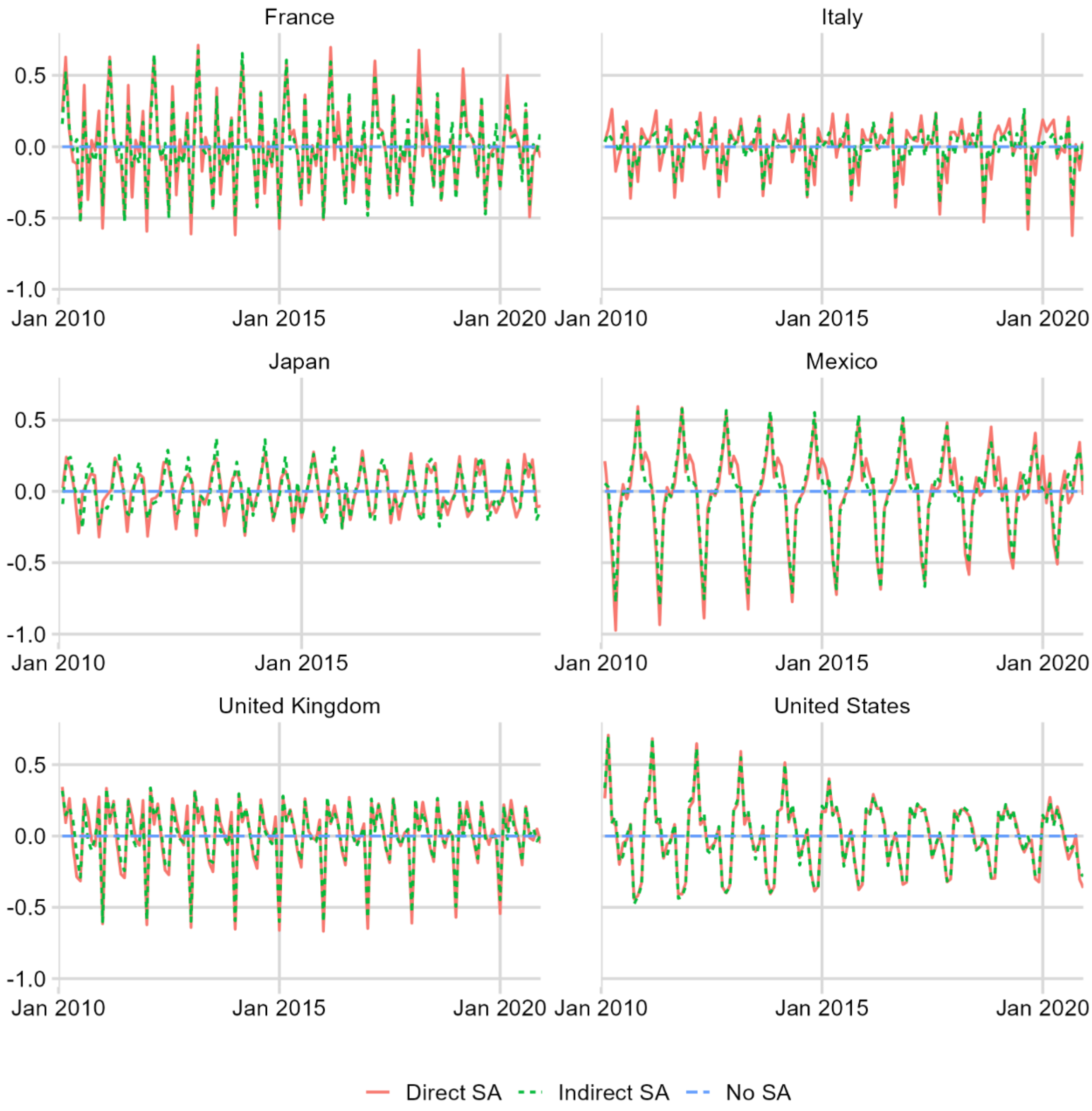
33. The resulting direct and indirect seasonally adjusted time series are nearly identical, regardless of whether one plots the absolute index values (Figure 7), the month-by-month changes (Figure AC.3 & 6) or the year-on-year changes (Figure AC.4 & 7). The month-on-month changes produced by the manually aggregated All items CPIs to those figures provided by NSOs, display very similar patterns (Figure AC.8).

34. The direct method of seasonal adjustment seems to produce a smoother series than the indirect method. The peaks and troughs of the deviations for direct seasonal adjustment are more pronounced than for indirect seasonal adjustment (Figure 8). The order of magnitude of the differences between direct and indirect methods is, however, extremely small. For the month-on-month rate, the direct method is on average -0.001 percentage point lower than the indirect method, when using X-13 for the adjustment. With TRAMO-SEATS, the direct approach produces a series that is on average almost identical the indirect method. The maximum difference between the direct and indirect approach is close to zero when using X-13 and 0.01 percentage point when using TRAMO-SEATS.

35. More importantly, the combined test for seasonality, when applied to the seasonally adjusted series, no longer detects the presence of seasonality for both measures. Across the 36 OECD countries, the average test statistic for residual seasonality was 0.32 with an average p-value of 0.87.

Figure 7. Direct versus indirect seasonal adjustment with X-13, deviation from aggregated CPI

Percentage point



Note: The charts plot to the deviation from the unadjusted all-items CPI, meaning that values further away from 0 imply more adjustment and hence a smoother series.
Source: Authors' calculations using the OECD CPI database.

41. Given that overall differences between direct and indirect seasonal adjustment methods for CPI are small and that both seem to fully capture seasonality, either approach appears to be justified from a statistical point of view. This also suggests that there is no need for a reconciliation method in case the indirect approach is used. But this finding may need to be revisited should the data sample change significantly.

4. Seasonal adjustment in times of large shocks

What are the issues?

42. Large shocks and especially recessions are expected to distort seasonally adjusted estimates (Luca and Wright, 2021; Dagum and Morry, 1985) The basic intuition is that seasonal adjustment procedures cannot disentangle precisely the impact of a large shock from a change in seasonality. This alters not only the identification of seasonal effects during the period of the shock but in the period that follows. This happens because seasonal adjustment filters use a weighted average of recent comparable periods to estimate the "normal" seasonal pattern for that period. A large but temporally limited disruption to economic activity and prices, such as with the pandemic or the Great Recession, might therefore introduce spurious seasonal patterns, an "echo" of the shock, in the subsequent data.

43. Several papers looking at different macroeconomic indicators document the presence of residual seasonality in time series in the aftermath of a shock. Wright (2013) highlights how the financial crisis 2008/2009 appears to have distorted seasonally adjusted US nonfarm payrolls data. He finds that initially this effect pushed reported seasonally adjusted figure up in the first half of the year and down in the second half of the year, by slightly more than 100,000 in both cases. As the relative weight of the recession period faded, the effect declined steadily. Lengermann et al. (2017) and Rudebusch et al. (2015) all point out that first estimates of seasonally adjusted GDP growth appear to have been biased, because of residual seasonality and the influence of the Great Recession.

44. In the specific case of prices, Peneva and Sadée (2019) and Peneva (2014) have shown that US consumer price inflation indices (in particular the CPI) have tended to be higher in the first half than in the second half of the year, despite the time series having been seasonally adjusted. Although they show that in seasonally adjusted vintages the difference between the first and second half of the year is smaller, it is still noticeable. They attribute this finding to residual seasonality without conclusively settling on an explanation.

How was it addressed by the NSOs during the pandemic?

45. The specific nature of the pandemic, which did not originate from the economic sphere added another layer of complexity in the treatment of seasonality. NSOs and academics disagreed over how best to treat the disruptions that arose from the lockdowns during the COVID-19 pandemic. In a brief overview Bógalo et al. (2021) classify NSO efforts to account for the COVID-19 shock into two main categories: the use of concurrent seasonal adjustment with additive outlier interventions and the projection of estimated seasonal factors for the year preceding COVID-19. They conclude that the projection of the preceding year's seasonal factors is a worse option than the intervention treatment.

46. In practice NSOs implement the adjustment in a more nuanced way, especially when it comes to the intervention models. Certain NSOs advocated for different outlier treatments than Additive Outliers (AO), for example Level Shifts (LS) or Transitory Changes (TC) (Annex B). There are also differing opinions on the preferred time-period(s) NSOs treat as outliers. For example, the US Bureau of Labor Statistics used an intervention analysis consisting of a level shift from June 2020 through September 2020. By contrast, Eurostat suggested modelling the COVID shock as an Additive Outlier (AO), with the crisis period defined as the first quarter of 2020, though this recommendation would later be extended for 2020 and 2021. The Australian Bureau of Statistics, although also having contemplated using an Additive Outlier (AO), preferred to create a "pseudo control" derived from RegARIMA factors on past dynamics of the series and using these factors to adjust the TC during the lockdown.

48. There are more sophisticated intervention approaches present in the literature on outlier treatments in seasonal adjustment, such as quadratic ramps (Lytras and Bell, 2013). For sake of simplicity, these are not considered in the subsequent analysis, which is illustrative only.

Outlier detection on CPIs in times of large shocks

36. Outlier detection is a key step in the seasonal adjustment process, which can be significantly influenced by the length of the time series. For instance, recent events—such as the COVID-19 pandemic in 2020 and the war in Ukraine have significantly impacted the CPI, and when included in the sample they may affect outlier detection and in turn both the preliminary adjustment phase and the extraction of seasonal components.

37. To test this, a total of 80 seasonal adjustments were conducted on CPI for 36 countries. Specifically, adjustments were made considering various starting points (1990, 1995, 2000, 2005) and end points (2020, 2021, 2022, 2023), applying five specifications (Table 4). Although the analysis covers all the periods, the focus of the section is on episodes of large shocks.

38. A notable finding was that the X13 method typically detected more outliers than the TRAMO-SEATS method across various time spans (Table 4). This is consistent with prior research indicating that although the preliminary adjustment processes of X-13 and TRAMO-SEATS are conceptually similar, differences in their methodological details could result in variations in the type and timing of detected outliers (Pavía Miralles et al., 2023). While TRAMO-SEATS consistently identified a similar number of outliers regardless of the series length, X-13 tended to detect more outliers as the series length increased. The choice of specification within each method, however, did not significantly affect the outcomes of outlier detection.

39. Furthermore, new outliers may be identified, or previously selected outliers excluded, when the number of data points increases, with changes occurring in both the type and timing of outliers (Table 5). This phenomenon was consistently observed in both the TRAMO-SEATS and X-13 methods, indicating that the outlier detection process is influenced by various factors and is sensitive to changes in these factors.

40. The inclusion of data from the latest year, and resulting changes in outlier detection, holds significant implications. National statistical agencies commonly perform seasonal adjustments annually or as new data becomes available, and differences in the results of seasonal adjustment compared to previously published series are usual. The use of X-13 will in this respect imply more revisions coming from the detection of additional outliers as the length of the time series increases. At the same time it is likely that this does not affect residual seasonality.

Table 4. Number of selected outliers by time interval and specifications

outlier year	time span	SPECIFICATIONS				
		T_R4	T_R5	T_Rf	X_R4	X_R5
2020	1990-2020	11	11	11	16	16
	1990-2021	12	12	12	16	16
	1990-2022	10	10	10	20	21
	1990-2023	10	10	10	23	23
	1995-2020	8	9	9	15	14
	1995-2021	9	10	9	19	19
	1995-2022	11	11	11	16	17
	1995-2023	9	9	9	16	16
	2000-2020	7	7	7	12	12
	2000-2021	11	10	11	15	14
	2000-2022	9	9	9	13	13
	2000-2023	7	7	7	11	11
	2005-2020	7	7	7	11	9
	2005-2021	9	11	11	11	12
	2005-2022	10	10	11	11	11
	2005-2023	9	9	9	10	10
2021	1990-2021	7	6	7	15	15
	1990-2022	10	9	10	18	18
	1990-2023	9	9	9	26	26
	1995-2021	10	9	10	13	13
	1995-2022	12	12	10	20	20
	1995-2023	11	10	10	21	20
	2000-2021	9	8	9	15	15
	2000-2022	12	12	12	18	18
	2000-2023	9	9	9	15	15
	2005-2021	10	10	10	12	12
	2005-2022	11	13	12	18	18
	2005-2023	11	11	11	15	16
2022	1990-2022	35	35	36	66	65
	1990-2023	41	41	41	82	83
	1995-2022	38	39	39	75	75
	1995-2023	41	41	41	75	77
	2000-2022	40	39	40	56	60
	2000-2023	40	42	42	70	71
	2005-2022	41	40	40	55	55
	2005-2023	43	43	43	57	60

Note: "T_R4" refers to the RSA4 specification of Tramo-Seats, while "T_R5" indicates the RSA5 specification of Tramo-Seats. "T_Rf" denotes the RSAfull specification of Tramo-Seats. "X_R4" represents the RSA4 specification of X-13, and "X_R5" corresponds to the RSA5 specification of X-13. See Annex B for more details on those specifications.
Source: Authors' calculations

Table 5. Changes in outliers with the addition of time series data

Country	Start year	Outlier type & time	TRAMOSEATS (Spec : Rfull) 2020	2021	2022	2023	X13-ARIMA(Spec : R5) 2020	2021	2022	2023
Austria	1990	LS (5-2020)			O	O				O
		TC (5-2020)					O	O	O	
	1995	LS (5-2020)			O	O				O
		TC (5-2020)					O	O	O	
	2000	LS (5-2020)			O	O				O
		TC (5-2020)	O	O			O	O	O	
	2005	LS (5-2020)			O	O				O
		TC (5-2020)	O	O			O	O	O	
United Kingdom	1990	AO (7-2020)	O		O		O	O	O	O
		AO (8-2020)		O		O				
	1995	AO (7-2020)	O	O	O	O	O	O		O
		LS (8-2020)							O	
	2000	AO (7-2020)	O	O	O	O	O	O	O	O
	2005	AO (7-2020)	O	O	O	O	O	O	O	O

Note: "O" indicates that an outlier has been detected, while a blank signifies that no outlier has been selected.
Source: Authors' calculations

Performance of TRAMO-SEATS and X-13 in the presence of outliers

41. Once outliers are detected in a series, national statistical agencies have several options within JDemetra+ to control for outliers. To test which approach is best, simulations similar to those reported in Annex B were conducted, albeit with an important addition. For the last two years of each time series, outlier effects are added to mimic one of three situations: Additive Outliers, Level Shifts and Transitory Changes. In times of shocks, NSOs need to make decisions about which outlier treatments apply, without fully knowing whether and for how long such treatments will be necessary. Usually this revolves around deciding which of the three outlier types and for what period they apply.

42. Simulations point to two interesting results. First, the presence of outliers in a time series decreases the accuracy of all seasonal adjustment procedures (Table 6). When comparing the performance of both X-13 and TRAMO-SEATS on simulated time series without outliers to time series that include outliers, the MAE is lower for series without outliers, by 3.3 percentage points and 3.8 percentage points respectively.

43. Second, the differences between the different outlier types are so small that they are unlikely to be statistically significant. Although manually specifying the correct outlier type seems to improve the performance of TRAMO-SEATS, the same does not hold for X-13. This finding echoes Bell et al. (2022) who conclude that, within a run of outliers, i.e. while the extreme events have not ceased to influence the time series yet, the choice of outlier type is largely irrelevant. Since outliers are simulated for the last 2 years of the time series, the specification of the outlier type does not appear to significantly improve seasonal adjustment.

44. Where the type of outlier does matter, is once a period of extreme values comes to an end, for example, because a recession has ended, or because lockdowns have been lifted (Bell et al., 2022). More research is needed to understand exactly how to determine when outliers start to affect time series values and how misspecification of outlier types impacts the performance of seasonal adjustment. This is

particularly true if a seasonally adjusted time series is meant to inform future decisions and will be used to forecast future values.

Table 6. Results of seasonal adjustment with manual adjustment for outliers

Average of 100 simulated time series over 10 years, error terms

Method	Simulated Outliers	Manual SA for Outliers	RMSE	MAE	MAPE
X -13	None		1.25	0.9	0.9
	Additive Outliers	AO	8.6	2.7	3.7
		TC	8.6	2.7	3.8
		LS	8.6	2.7	3.8
	Level Shift	AO	18.0	8.7	6.2
		TC	18.0	8.7	6.2
		LS	18.0	8.7	6.2
	Transitory Change	AO	3.2	1.2	1.1
		TC	3.2	1.3	1.2
		LS	3.2	1.2	1.2
TRAMO-SEATS	None		1.0	0.5	0.6
	Additive Outliers	AO	8.8	3.5	4.6
		TC	8.5	2.5	3.6
		LS	8.6	2.5	3.5
	Level Shift	AO	18.0	8.6	6.1
		TC	18.0	8.7	6.2
		LS	18.0	8.5	6.1
	Transitory Change	AO	3.6	2.0	2.0
		TC	3.1	1.0	1.0
		LS	3.4	1.8	2.0

Note: The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE measures how concentrated the estimated data is around the simulated values. The Mean Absolute Error (MAE) is the arithmetic average of the absolute errors between the estimated and the simulated values. It tells us how far the simulated values are off the estimated ones, on average. The Mean Absolute Percentage Error (MAPE) converts the MAE into relative values. It calculates how far off the simulated values the estimated values are on average, in terms of percentages.
Source: Authors' calculations using the simulated CPI database.

Structural changes to seasonality

45. The next question is how much large shocks have the potential to change post-shock seasonality. In the case of the pandemic, for prices, this might occur due to changes in household consumption behaviour, or due to the effects of supply-constraints on the availability of certain goods. In other words, large shocks are likely to cause or accelerate structural changes in the underlying time series, that potentially require fine-tuning of the usual seasonal adjustment procedures.

46. Determining the presence and potential impact of structural breaks in a time series is an open question. In the case of CPI the question is further complicated by the period of high inflation that followed shortly after the conclusion of the pandemic. Since inflation is a shock in its own right, a period of high inflation has the potential to cause a structural break in seasonality (Ari et al., 2023).

47. While it seems as though the pandemic has had little impact on seasonality so far (Section 2), it is possible that sufficient data is lacking to fully appreciate the presence of any structural breaks. This is

certainly true in the case of the most recent inflation shock, which although showing signs of having peaked, has by no means abated yet. In any case the question of structural breaks presents an interesting topic for further research. For starters, the Australian Bureau of Statistics (2021) offers a conceptual breakdown of the potential issues (Table 7).

Table 7. Conceptually determining structural breaks in seasonal adjustment

		Real world effect on series	
		Pandemic causes a structural change	***Pandemic does not cause a structural change***
Decision on treatment based on limited data observations	***Determine a structural change***	Minimise the delay in settling ultimate decomposition:	The adjustment for a structural break is later removed:
		- Users sooner interprete behaviours	- Reduced clarity in short-term direction
		- Reduce size of revisions	- Earlier revision in data points potential reversed
	Determine no structural change	Delay in identifying s structural change:	Correct decision not to adjust for a break:
		- Longer period of higher volatility in SA series, reducing clarity in short-term developments	- Avoiding increased volatility which reflected a temporary disruption
		- More points get revised when structural break adjustment is implemented	- Avoiding some points being revised in opposite directions

Source: Based on Australian Bureau of Statistics (2021).

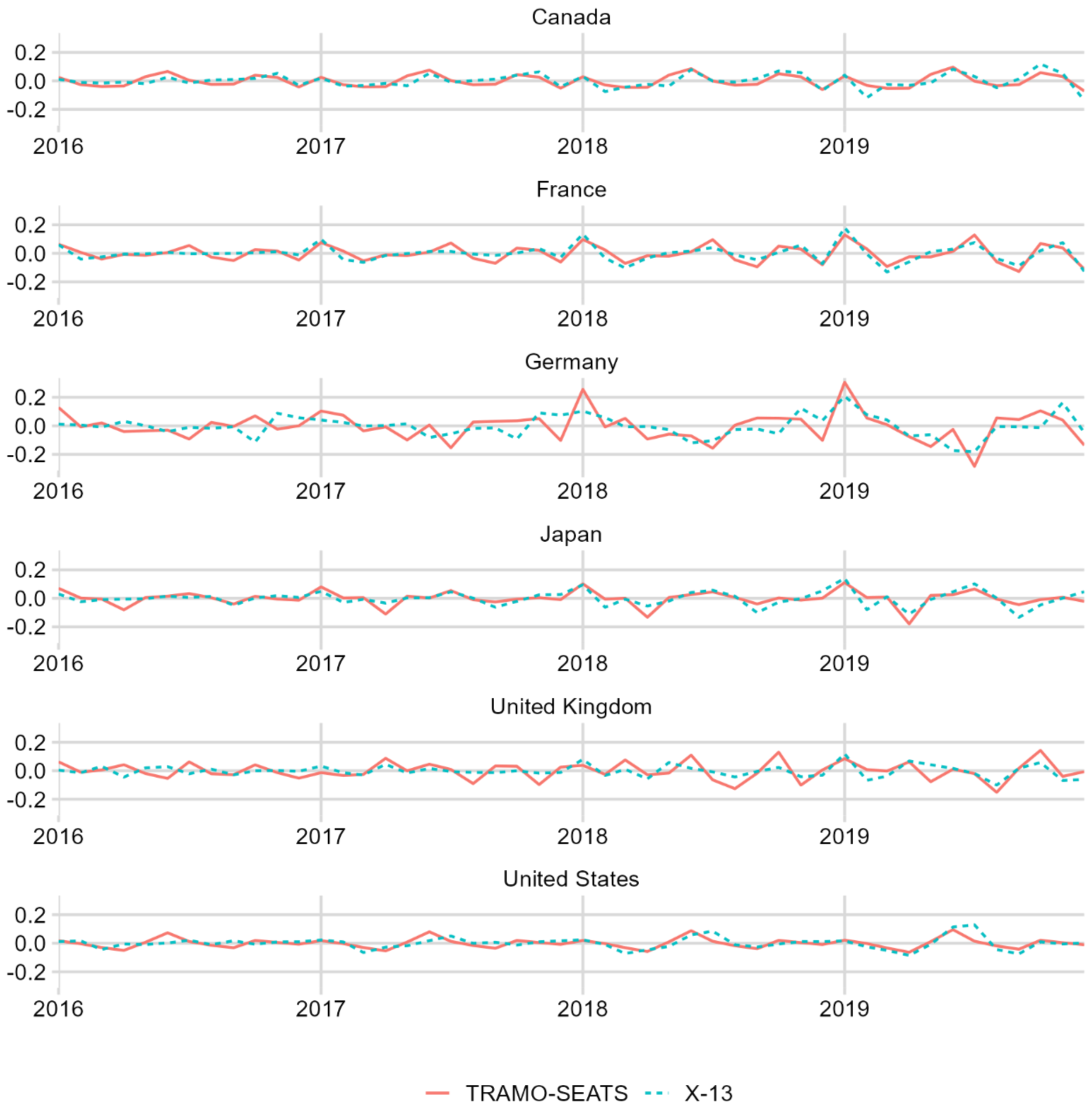
Revisions to seasonally-adjusted CPIs

48. Revisions have been put forward as the major reason to exercise caution when publishing seasonally adjusted CPIs (see Section 6). To give an upper bound of the magnitude of the changes in seasonality stemming from data revisions, the exercise focuses on an extreme situation combining the COVID-19 crisis and the recent period of high inflation using both X-13 and TRAMO-SEATS. CPIs are adjusted over two periods: from 1980 to 2019 (i.e. before the outbreak of the pandemic) and from 1980 to 2023 (including the large shocks). Note that the absolute seasonally adjusted values are in this exercise converted into month-on-month percentages, which analysts working on short-term developments are usually most interested in.

49. Regardless of the method, revisions deviate only slightly from the original time series (Figure 9). Across the G7 countries, the average absolute difference stemming from revisions amounts to 0.02 percentage point for TRAMO-SEATS and 0.01 percentage point for X-13. The magnitude of the differences varies markedly across countries. The maximum revision amounts to 0.395 percentage point in the United Kingdom for TRAMO-SEATS and 0.207 percentage point in Germany for X-13. They are estimated to be large in the United Kingdom and Germany but negligible for Canada and France. The largest difference between the two methods is observed for the United Kingdom, where X-13 outperforms TRAMO-SEATS on average by about 0.04 percentage point. Except for Japan, the X-13 outperforms TRAMO-SEATS in all countries.

Figure 8. Differences between X13 and TRAMO-SEATS revisions in G7 countries

Deviation from first estimation, percentage points



Note: The seasonal adjustment was performed with the default specification for both methods, but it does not include any outlier treatment, as this would have made an equal comparison more difficult. The seasonal adjustment was performed for the entire time series (1980 – 2020). To increase visibility, only a subset is presented here; the message would be unchanged if the whole time series was presented here.
Source: Authors' calculations using the CPI dataset.

50. It is important to mention that the differences in revisions are likely to be influenced by the length of the time series. As outlined in Section 4, shorter time series tend to be more volatile, and TRAMO-SEATS outperforms X-13 in those circumstances. Therefore, a decision based on the differences in revisions should equally be informed by the length of the available time series.

5. Experimental methods and seasonal adjustment

51. 59. NSOs that seasonally adjust CPI usually use an X-12/X-13 method, and standard econometric tests. Standard tests usually provide a consistent assessment of the presence of seasonality in the majority of cases but can sometimes lead to contradictory results for a non-negligible number of series. There may be scope for improvement, especially as seasonal patterns can vary over time and seasonal components and non-seasonal components are not independent and thus not separable, contrary to what is assumed in standard methods (Hylleberg, 1994).

52. Being a flexible modelling tool, neural networks which have been increasingly used in recent years, can, in principle, model any type of relationship in the data with high accuracy when sufficient data points are available. This makes them a potential good candidate to simultaneously detect both the nonlinear trend and the seasonality in the data (Gorr, 1994). Empirically, a number of papers have found that neural networks can detect seasonal patterns (Franses and Draisma ,1997; Nam and Schaefer, 1995; Williams, 1997). Other machine learning methods can also prove useful. Ollech and Webel (2020) show that tree-based methods are useful to classify whether a series presents a seasonal pattern or not. Misclassification rates are particularly low for random forests, independently of the time series length, as opposed to some single seasonality tests.

53. Beyond the detection of seasonality, evidence is mixed as to whether machine or deep learning would outperform traditional methods in adjusting data for seasonality. For instance, neural networks have been sometimes found to outperform traditional statistical models in identifying both seasonal and trend variations in time series data (Wang et al., 2011; Rahman et al. 2019; Hamzaçebi, 2008; Mitrea et al., 2009). By contrast Zhang and Qi (2005) find that that neural networks are not able to capture seasonal or trend variations effectively. There is also evidence that complexity in neural networks is sometimes required to model seasonality adequately (Curry, 2006).

54. Overall, the usefulness of machine learning – in particular neural networks – in improving the seasonal adjustment is still an open question that needs to be further investigated. One promising area could be to use machine learning to improve seasonality detection tests.

6. Communication and guidance to users

55. Only a few OECD countries currently produce seasonally adjusted (SA) CPIs. Some explanation for this practice can be inferred from the main CPI methodological guidance document, the CPI Manual (IMF et al., 2020). It stipulates that “CPI are not normally seasonally adjusted, although some countries do produce a SA CPI”. Using year-on-year index changes is recommended to avoid seasonality affecting the resulting inflation rates or using “core” CPI that excludes highly volatile or seasonal items like food or energy. However, seasonality patterns can vary across periods (e.g. due to moving holiday dates, like Easter) and consequently using year-on-year indices may be insufficient to account for seasonality in the data. At the same time, month-on-month and quarter-on-quarter CPI index changes are key metrics used to examine short-term developments and for forecasting.

56. In terms of dissemination, the Manual recommends the SA series should be marked as series for analytical purposes. Their explanation to the user, detailed methodology and reasons why a particular seasonal adjustment procedure has been followed, should be made available. The main reason for this cautious approach is the fact that SA series can be (and often are) revised. The revisions to CPI series are more generally perceived as not permissible, except in exceptional cases (e.g. in case of an error exceeding a threshold). Hence, data revisions play an important role for the communication of SA CPI to users, as publishing SA data is usually the only instance when a CPI series is revised.

57. The prudence with respect to revisions is, in turn, related to perceptions among the public who sometimes perceives the headline inflation rate as underestimating the overall increase in prices they observe in their everyday life. This can impact the credibility of the index. Revisions may further aggravate this issue. Moreover, CPI is often used for indexation both in the private and public sectors (e.g. for wages, pensions, public transfers, private contracts, etc.), in contexts where revisions may be problematic. CPI flash estimates in principle face similar communication issues as SA CPI, as the flash estimate also can be “revised” in the subsequent official release. In general, the prerequisite to publishing any alternative CPI estimates or series, including those SA, is that they are well communicated, documented and explained to the users. Lastly, the extent to which revisions may pose a communication problem is possibly closely related to their size and frequency – if the revisions are small and/or scarce, the CPI series can be more easily perceived as stable and hence credible. Therefore, it is advisable that the extent of revisions is among the decision criteria when choosing from several seasonal adjustment methods.

58. The practice of communicating CPI revisions and SA series varies in the sample of OECD countries or zones that regularly publish seasonally adjusted CPI (Table 8 and Box 2 for a more detailed overview).

Table 8. Practices of communicating seasonally adjusted data in selected OECD countries

Communication-related practice	Countries applying it
Commonly revising consumer inflation data (other than due to seasonal adjustment)	United States (the Personal Consumption Expenditure Index)
Marking SA data as experimental or analytical	Australia
Presenting SA data as the default series in communicating month-on-month or quarter-on-quarter inflation	United States
Presenting and/or mentioning SA data in the regular CPI press release	United States, Australia, Canada, France
Publishing methodology used for SA	All 9 countries
Publishing simplified explanatory/methodological information on SA (in non-technical language)	United States, Canada, Japan, Mexico, Norway
Publishing information on which data to use for indexation	United States, Australia, Canada
Pointing out and/or explaining revisions related to SA data	United States
Explaining generally that SA data imply revisions and providing information on their frequency and period impacted	United States, Canada, Japan, Norway, France (no information on period impacted), Australia (no information on frequency or impacted period)
The extent of data revisions mentioned as one of the criteria when choosing among SA methodologies/parameters	Euro area (ECB), France, Germany

Note: SA refers to seasonal adjustment. Table sums up information from 9 OECD countries or zones that regularly publish seasonally adjusted CPI data (Australia, Canada, euro area, France, Japan, Germany, Mexico, Norway and the United States).
Source: OECD compilation.

59. Some countries such as the United States or Canada publish a lot of information materials on SA data, ranging from those in easily accessible language to more technical documents and from general

materials to specific ones (for instance guidance on what indices to use for indexation). In addition, some countries feature the SA data in their regular CPI press release (e.g. Australia, France, Canada) or even present them as the default series when communicating short-term inflation rates (the United States). Alternatively, some countries only produce SA CPI series seemingly "on the side", i.e. they do not promote them in any way and it is less straightforward to find information on them (e.g. Germany, Mexico).

60. In general, it seems good practice also for easier communication of SA series to consider the extent of implied revisions as one of the decision criteria when choosing among SA methods, as it is done by the ECB. Moreover, a transparent communication of seasonal adjustment in non-technical language is advisable, including explanation on what seasonal adjustment is, why it is performed, that it implies data revisions, in what frequency and impacting what periods of data series. This information should be easily accessible with the data and can be part of the explanatory note accompanying the data and the regular CPI press release, as is the case in the United States, or prominent on the NSO's dedicated CPI website section, as in Japan, Australia, Norway or Canada. It would be very useful to advise users not to use SA CPI, which are prone to revisions, for indexation and in contracts, administrative and judicial provisions. This could be done by publishing short notes as done in the United States, Australia or Canada. This would complement more technical methodological documentation available to users which should be the norm. Lastly, it would be useful to point out specific data revisions, either in the press release and/or in the online communication of CPI. Revised data should also be flagged in the database.

Box 2. Communication in selected OECD countries with respect to revisions and seasonally adjusted series

The United States is a special case, as the typical user of US inflation figures may be more used to revisions than users in other countries where CPI is the main inflation measure. Indeed, in addition to CPI, there is another prominent inflation measure – the target measure used for monetary policy purposes – the (core) personal consumption expenditure (PCE) index. Unlike CPI series, even the unadjusted PCE is subject to both relatively big (in extreme cases by as much as 6 percentage points, and more usually between 0 to +/-0.6 percentage point for core PCE, in annualised terms), as well as relatively frequent revisions (both in the months following the respective month's estimate, as well as in the following years) (Audoly et al., 2023). The Bureau of Economic Analysis (BEA) publishes PCE revisions in each PCE press release, usually without any special comment, unless the revision is due to for instance a major update of the National Accounts, in which case a more detailed information is provided. In data tables, revised figures are flagged. At the point of a new release, the previous releases are marked as superseded and the user is pointed in a short note to the latest data. PCE is published both in a SA (default) and non-adjusted version. Besides denoting the (lack of) seasonal adjustment, they are not specifically distinguished and the SA series is not marked as analytical. Both types of series are presented as estimates. General documentation available online explains why the PCE is revised.

The US month-on-month CPI published by the Bureau of Labor Statistics (BLS) in its press release is also by default presented as SA and is not marked as analytical. The technical note of the press release includes a comprehensive section on explaining the SA data, including how often the seasonal factors are updated and what part of the series is impacted, treatment of extreme events and links to a more detailed documentation on the adjustment. It also provides an explanation in what cases the SA series may be more useful than unadjusted ones (e.g. for analysing short-time trends), and use cases where unadjusted series should be preferred (e.g. indexation purposes). The BLS also publishes various CPI documentation materials e.g. a Factsheet on indexation (escalation) or a Handbook of methods, where it also points out to differences and appropriateness of using SA vs unadjusted series, if relevant.

Statistics Bureau of Japan (SBJ) publishes both SA and non-adjusted CPI data. The former are not denoted as analytical or experimental. On its English CPI webpage, the SBJ announces if CPI data was revised due to seasonal adjustment. It also lists detailed methodological information, including on seasonal adjustment, as well as several appendices on more detailed SA information with parameters of the used SA methodology. The SBJ also mentions seasonal adjustment in its Frequently Asked Questions. However, the information available in English does not seem to mention caveats or appropriate uses of SA versus non-adjusted data.

In **the European Union**, a regulation lays down the methodological requirements regarding the Harmonised Index of Consumer Prices (HICP), including revisions. In principle, revisions to HICP should be made only if there are errors in the data; other revisions need to be coordinated with the European Commission and they have to be published with their explanation. Each HICP press release also includes the flash estimate for the respective reference month in the revisions section. Eurostat maintains a database of first releases of HICP starting in 2016. To aid short-term inflation analysis, the ECB produces SA HICP for the euro area aggregates, the SA data being published on the ECB website (with no press release) at the same time as Eurostat's publication of unadjusted HICP series. Documentation of methodology (ECB, 2000), having minimisation of implied revisions as one of the criteria in choosing the method of adjustment, is available. The series are not explicitly marked as analytical series.

The Federal Statistical Office of Germany (DESTATIS) publishes SA CPI data. They are not, however, featured in the regular CPI press release that only contains non-adjusted data. The dedicated DESTATIS CPI webpage also does not mention the SA series, though they are published together with the non-adjusted series in the short-term statistics database. The SA methodology is separate from the CPI methodological information. Documentation of general SA methodology (Linz, Fries and Voelker, 2018; for all short-term indicators, not only for CPI), including reasons for seasonal adjustment, is published separately.

The French National Institute of Statistics and Economic Studies (INSEE) publishes SA figures for headline CPI alongside the default non-adjusted figures in its CPI press release, including mentioning them in the text. The release does not include any disclaimers or explanations of SA data. It is accompanied by a general abbreviated CPI methodology note that includes a short section informing the user that INSEE also publishes SA data for headline CPI and components and that these are annually revised. The methodological note (Smyk and Tchang, 2021) is related to seasonal adjustment in general and not specific to CPI.

The Australian Bureau of Statistics (ABS) started publishing monthly CPI data experimentally in 2022. While the quarterly data are revised only exceptionally, monthly data, while still experimental, may be revised more frequently. Along with non-adjusted quarterly CPI, the ABS also publishes SA series, marked in the press release as analytical (alongside trend CPI measures). The press release also contains link to SA methodology documentation and mentions that this data may be subject to revisions. Similarly to the US practice, the ABS CPI online documentation also includes information on how to use price indices in contracts, where also seasonal adjustment and revisions are discussed.

Statistics Canada (StatCan) publishes SA CPI series alongside the default non-adjusted in the monthly CPI press release, including mentioning them in the release text. They are not marked as analytical or experimental. The data table accompanying the release includes a short note on what seasonal adjustment means, that it implies revisions and what period may be impacted (last three years), a disclaimer that the series should not be used for indexation and a link to general CPI methodology that also includes a section on seasonal adjustment method. StatCan also has an online section dedicated to SA data more generally (not only CPI data). This includes frequently asked questions related to both conceptual issues, as well as issues related to analysis and interpretation, including in what contexts SA data and non-adjusted data should be used.

The **Mexican National Institute of Statistics and Geography** (INEGI), published non-adjusted CPI data. For monetary policy purposes, Banco de Mexico (Banxico) seasonally adjusts INEGI's CPI data. These are published in monetary policy publications but are not accessible via a database. A short explanatory note on seasonal adjustment of inflation, as well as a more technical methodological document (Capistran, Constandse and Ramos-Francia, 2009) on SA, are available at Banxico's website.

Statistics Norway (SN) publishes SA series for its headline CPI and CPI adjusted for tax changes and excluding energy products (CPI-ATE). They are not featured in the monthly CPI press release. They appear in the list of available CPI series at the SN CPI website. In addition, the methodological section of the CPI website also mentions information on SA: from explanation on what is SA and why it is done, that it implies revisions, how often they are made and how far back they apply; to basic methodological information on SA, as well as links to further, more comprehensive but still reader-friendly material on SA methodology by SN.

7. Conclusion

61. This paper examines the presence of seasonality in CPIs in 36 OECD economies and reviews the properties of standard methods, namely X-13 and TRAMO-SEATS, in performing the adjustment. It contributes to the empirical literature in that it investigates seasonal patterns and performance of main seasonal adjustment methods in a relatively large sample of OECD countries, using simulations and an observational approach. In addition, as official statistics in general, and CPI in particular, can have specificities in terms of their communication to the wider public, the paper also summarises communication practices in OECD countries already regularly performing seasonal adjustment of CPI and draws lessons from these practices.

62. Evidence from statistical tests point to the presence of seasonality in headline CPI and its components, with stronger seasonality in some categories. There are also indications of changes in seasonal pattern from 1980 to 2022, but it is not systematic across countries. Based on the information so far, the COVID-19 pandemic does not appear to have changed CPI seasonal patterns. This issue will, however, need to be revisited when the period of heightened inflation that immediately followed the pandemic and made identification of its effects in isolation more difficult, has subsided.

63. Turning to the assessment of methods, the paper focuses on the most standard approaches, X - 13 and TRAMO-SEATS which are also used by most NSOs (Table 9). Simulations suggest that differences between the two methods are small when applied to CPI in OECD countries in normal times. Differences between the direct and the indirect approaches are also minimal, limiting the need for a reconciliation method, and the combined test for seasonality no longer detects the presence of residual seasonality after adjustment. In times of a large shock, evidence suggests that the type of outlier used can make a difference only after the impact of the shock has dissipated. The number of outliers detected by X-13 increases with the sample, while it stays constant with TRAMO-SEATS.

64. Finally, in terms of publication and dissemination, country experience suggest that the release of SA CPI should be accompanied by good and transparent communication, including alerting users not to use seasonally adjusted CPI for indexation purposes or in legal contracts.

Table 9. Summary of the assessment

	X-13	TRAMO-SEATS
Performance in normal times without outliers	MAE: 0.470	MAE:0.555
Performance in normal times with outliers	MAE: 4.21	MAE: 4.34
Difference direct indirect	Small difference	Small difference
Outlier detection	Increase with sample	No change with sample
Size of revisions, Average G7, since 2010	Mean: 0.01 percentage point Max: 0.207 percentage point	Mean: 0.02 percentage point Max: 0.395 percentage point

Note: The MAE comparing the performance in normal times refers to the best performing specification for each method.
Source: Authors' compilation.

65. Against this background, the paper can serve as a practical initial guidance to NSOs or other institutions considering seasonal adjustment of their CPI data. It provides an overview of the basic methods used for SA more generally, a summary of the specific methods used in NSOs already publishing SA CPI, as well as an assortment of testing exercises that can be performed to aid the decision-making process on the final SA method and its parametrisation. It is advisable that these tests are revisited from time to time, so that the seasonal adjustment method is always up to date and reflects the needs and characteristics of the data. The extent of revisions implied by the SA should be among the criteria for choosing the final method. The communication of revisions, and more generally of the SA data, is an important aspect of publishing SA CPI. To this end, the paper provides a summary of how communication is handled by selected OECD countries and provides a list of best practices that can be drawn upon by an NSO aiming to publish SA CPI.

66. The paper provides only a first step in the examination of seasonality in CPI and could be usefully complemented by additional analysis. In particular, it could be worth investigating whether machine learning could help improve seasonality detection tests. In addition, investigating the current period of heightened inflation and its effect on seasonal patterns once it has subsided, could be another interesting shock to be analysed, as it may or may not have caused structural shifts which will require more incoming data to assess.

References

Abeln, B. and J. Jacobs (2015), "Seasonal Adjustment with and Without Revisions: A Comparison of X-13ARIMA-SEATS and CAMPLET", CAMA Working Paper No. 25/2015.

Ari, A., Mulas-Granados, C., Mylonas, V., Ratnovski, L., & Zhao, W. (2023). "One Hundred Inflation Shocks: Seven Stylized Facts." IMF Working Paper WP/23/190

Audoly, R., M. Almuzara, R. Crump, D. Melcangi, and R. Xing (2023), "How Large Are Inflation Revisions? The Difficulty of Monitoring Prices in Real Time", Federal Reserve Bank of New York Liberty Street

Australian Bureau of Statistics (2021), "Seasonal Adjustment throughout Periods of Significant Disruption and Uncertainty", September 7.

Baxter, M. (1999), "The Seasonal Adjustment of RPIY", *Economic Trends*, 546.

Bell, W. R. and T. McElroy (2022), "Seasonal Adjustment of Time Series During the Pandemic", Presentation United States Census Bureau.

Andrew R. Biehl & John P. Judd, 1993. "Inflation, interest rates, and seasonality," FRBSF Economic Letter, Federal Reserve Bank of San Francisco, issue October 15.

Bógalo, J., P. Poncela and E. Senra (2021), "Circulant Singular Spectrum Analysis to Monitor the State of the Economy in Real Time", Mathematics 9, no. 11: 1169.

Bógalo, J., M. Llada, P. Poncela and E. Senra (2022), "Seasonality in COVID-19 Times.", Economics Letters 211, February.

Bryan M.F. and S. Cecchetti (1993), "Measuring Core Inflation", NBER working paper No. 4303.

Capistran, C., C. Constandse and M. Ramos-Francia (2009), "Using Seasonal Models to Forecast Short-run Inflation in Mexico", Banco de Mexico Working Paper No. 2009-05.

Cleveland, R. B., W. S. Cleveland, J. E. McRae and I. J. Terpenning (1990), "STL: A Seasonal-Trend Decomposition Procedure based on Loess", *Journal of Official Statistics*, *6*(1), 3–33.

Curry B. (2007), "Neural Networks and Seasonality: Some Technical Considerations", European Journal of Operational Research 179, 267–27.

Dagum, E.B., (1980), "The X-11-ARIMA Seasonal Adjustment Method", Ottawa: Statistics Canada.

Dagum E.B. and M. Morry (1985), "Seasonal Adjustment of Labour Force series during recession and non recession periods",*Statistics Canada: Survey Methodology*, Vol. 11, No. 2

Daniel, S., V. Haller and B. Bellone (forthcoming), "Real-time and Multi-Frequency Seasonal Adjustment: Traditional vs Alternative Data", *Paper presented at the OECD-INSEE-Eurostat-TSACE workshop on times series*, 13-14 December 2023.

Dokumentov, A. and R. J. Hyndman (2015), "STR: A Seasonal-Trend Decomposition Procedure Based on Regression", Technical report, Monash University, Department of Econometrics and Business Statistics.

Durbin, J, and S. J. Koopman (2012"", 2nd edn, Oxford Statistical Science Series.

ECB (2016), "Harmonised Index of Consumer Prices – Easter effects and improved seasonal adjustment", Economic Bulletin, Issue 3.

ECB (2014) "Calendar Effects in Recent Inflation Developments", Monthly Bulletin May.

ECB (2000), "Seasonal Adjustment of Monetary Aggregates and HICP for the euro area",

Eurostat (2018), *Handbook on Seasonal Adjustment*

Eurostat (2015), *ESS Guidelines on Seasonal Adjustment*.

Franses, P.H. and G. Draisma (1997), "Recognizing Changing Seasonal Patterns using Artificial Neural Networks", Journal of Econometrics, 81, 273–280.

Gorr, W.L. (1994), “Research Prospective on Neural Network Forecasting”, International Journal of Forecasting, 10, 1–4.

Gómez, V. and A. Maravall (1997), "Guide for Using the Programs TRAMO and SEATS (Beta Version: December 1997)," Working Papers 9805, Banco de España.

Hamzaçebi, C. (2008), “Improving Artificial Neural Networks’ Performance in Seasonal Time Series Forecasting”, Journal of Information Sciences, 178, 4550–4559.

Harvey, A. (1990), “Forecasting, Structural Time Series Models and the Kalman Filter”, Cambridge University Press.

Hylleberg S. (1994), “Modelling Seasonal Variation” in Hargreaves, C.P. (Ed.), Nonstationary Time Series Analysis and Cointegration. Oxford University Press, Oxford, pp. 153–178.

Hansen J.V. and R.D. Nelson (2003), “Forecasting and Recombining Time-Series Components by Using Neural Networks”, Journal of the Operational Research Society, 54 (3), 307–317.

ILO, IMF, OECD, EU, UN, WB (2020), ”Consumer price index manual: concepts and methods”, Washington, DC: International Monetary Fund.

Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (2000). “STAMP 6.0 Structural Time Series Analyser, Modeller and Predictor”, Timberlake Consultants Press, London http://www.stampsoftware.com.

Lengermann, P., N. Morin, A. Paciorek, E. Pinto, and C. Sahm (2017), “Another Look at Residual Seasonality in GDP”, FEDS Notes, no. 2031, July.

Linz, S., C. Fries and J. Voelker (2018), “Seasonal Adjustment of Short-Term Statistics using X-12-ARIMA and X-13 in J-Demetra+”, WISTA Wirtschaft und Statistik, Statistisches Bundesamt (Federal Statistical Office).

Lucca D. and J. Wright (2021), “Reasonable Seasonals? Seasonal Echoes in Economic Data after COVID-19”, Liberty Street Economics, March 21.

Lytras, D. and W. R Bell (2013), “Modeling Recession Effects and the Consequences on Seasonal Adjustment,” US Census Bureau.

Lis L. and M. Porqueddu (2018), “The Role of Seasonality and Outliers in HICP Inflation Excluding Food and Energy”, ECB Economic Bulletin, Issue 2.

Manski, C. F. (2015), “Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern”, Journal of Economic Literature 53, no. 3: 631–53. https://doi.org/10.1257/jel.53.3.631.

McElroy, T. and A. Roy (2022), “A Review of Seasonal Adjustment Diagnostics”, International Statistical Review 90, no. 2, 259–84, August.

Mitrea C. A., C. K. M. Lee and Z. Wu (2009), “A Comparison between Neural Networks and Traditional Forecasting Methods: A Case Study”, International Journal of Engineering Business Management, 1(2);19- 24.

Nam K. and T. Schaefer (1995), “Forecasting International Airline Passenger Traffic using Neural Networks”, Logistics and Transportation Review 31 (3), 239–251.

Nelson M., T. Hill, O. Remus. and M. Connor (1999), “Time Series Forecasting using NNs: Should the Data be Deseasonalized First?”, Journal of Forecasting 18, 359–367.

Ollech, D. (2018), “Seasonal Adjustment of Daily Time Series”, Deutsche Bundesbank Discussion Paper No. 41.

Ollech, D. and K. Webel (2020), “A Random Forest-based Approach to Identifying the Most Informative Seasonality Tests”, Deutsche Bundersbank Discussion Paper, No. 55.

Pavía Miralles, J. M. B. Cabrer Borras and D. Iranzo Pérez (2012), “Outlier detection with automatic modelling: TRAMO/SEATS versus X-12-ARIMA”, Model Assisted Statistics and Applications, vol. 7, no. 3, pp. 229-244.

Peneva, E. (2014), "Residual Seasonality in Core Consumer Price Inflation", FEDS Notes. Washington: Board of Governors of the Federal Reserve System, October.

Peneva, E. and N. Sadee (2019), "Residual Seasonality in Core Consumer Price Inflation: An Update" FEDS Notes 2019, no. 2318, February.

Rahman A., Z. Hossain, M L. Rahman and A. K Majumder (2019), "Seasonal Adjustment Based on Neural-Network in Time Series Analysis: A Comparative Study", JUJSS Vol. 35, 2019, pp 65-75

Riley H. E. (1961), "Some Aspects of Seasonality in the Consumer Price Index", Journal of the American Statistical Association, vol. 56, no. 293, 1961, pp. 27–35.

Rudebusch, G.D., D.J. Wilson and B.D. Pyle, (2015), "Residual Seasonality and Monetary Policy" , in 2015 FRBSF Economic Letters.

Smyk, A. and A. Tchang (2021), "R Tools for J-Demetra+ Seasonal Adjustment Made Easier", Document de travail no. M2021/01, INSEE.

Tiller, R.B. and T. Evans (2014), "Seasonal Adjustment and the Great Recession: Implications for Statistical Agencies," BLS.

US Census Bureau (2023), X-13ARIMA-SEATS Reference Manual.

Wang S., L. Yu, L. Tang and S. Wang (2011), "A Novel Seasonal Decomposition Based Least Squares Support Vector Regression Ensemble Learning Approach for Hydropower Consumption Forecasting in China", Energy, 36; 6542-6554.

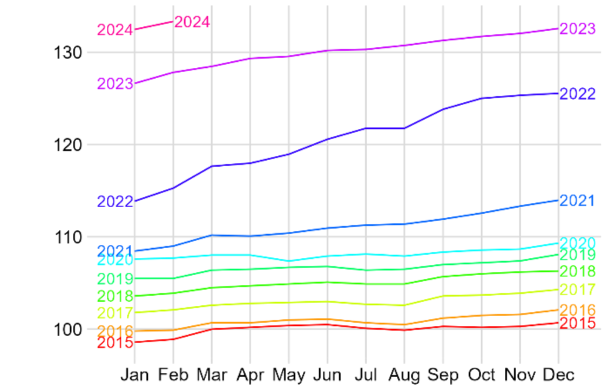
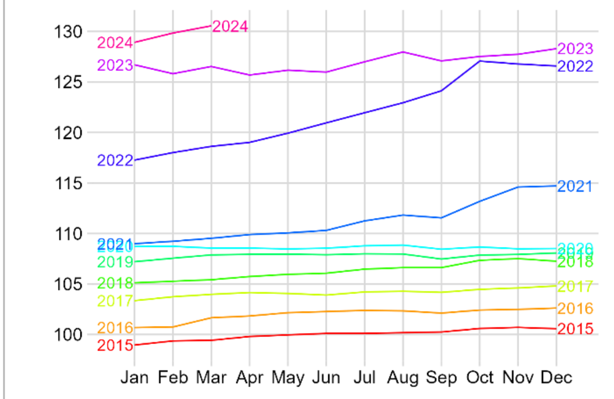
Williams P.M. (1997), "Modelling Seasonality and Trends in Daily Rainfall Data" in Touretzky, D.S., Mozer, M.C., Hasselmo, M.E. (Eds.), Advances in Neural Information Processing Systems, vol. 10, pp. 985–991.

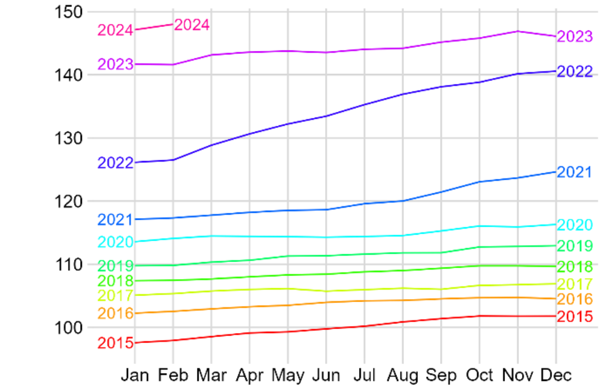
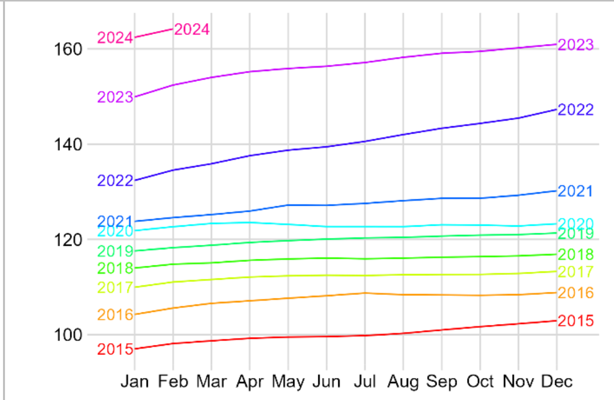
Wright, J. H (2013), "Unseasonal Seasonals?", Brookings Papers on Economic Activity, no. 2, 65–126.

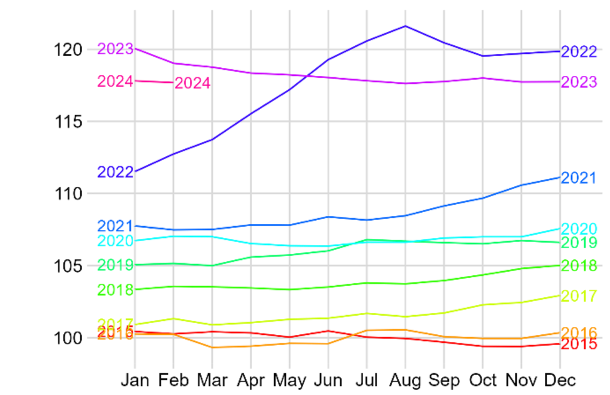
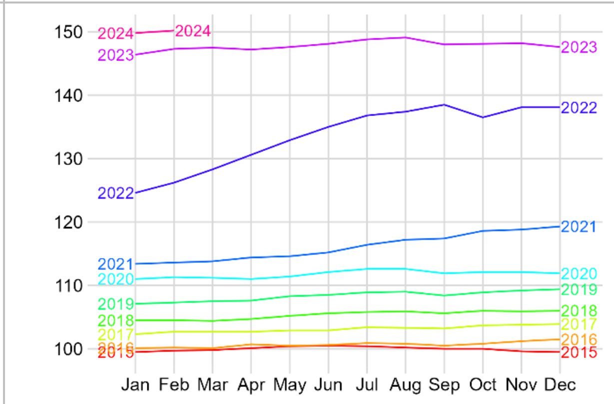
Zhang G. P. and M. Qi (2005), "Neural Network Forecasting for Seasonal and Trend Time Series", European Journal of Operation Research, 160; 501-514.

Annex A. Visual inspection of seasonality in headline CPI in non-G7 economies

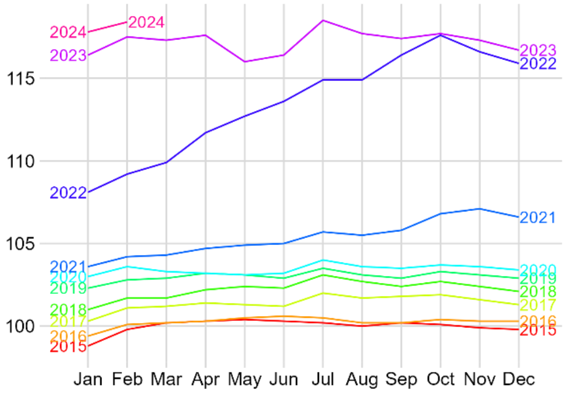
Index =100 in 2015

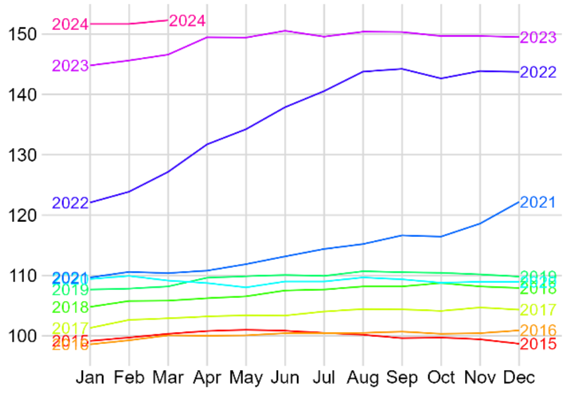
Austria	Belgium \i
	

Chile	Colombia
	

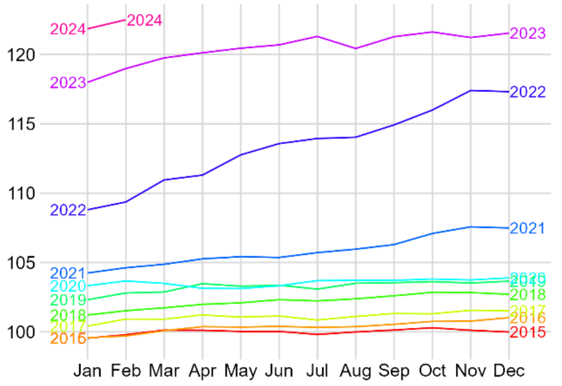
Costa Rica	Czechia
	

Denmark	Estonia

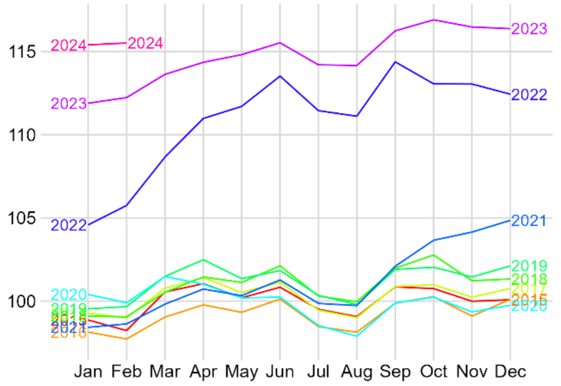




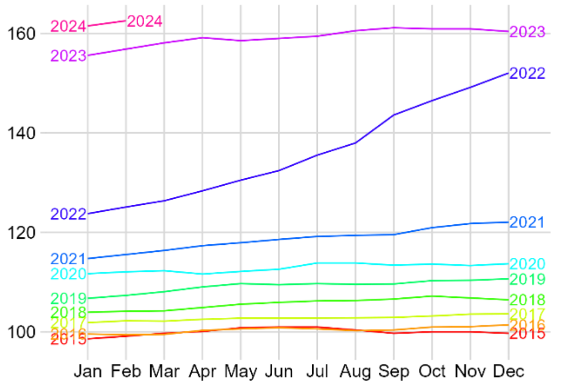
Finland



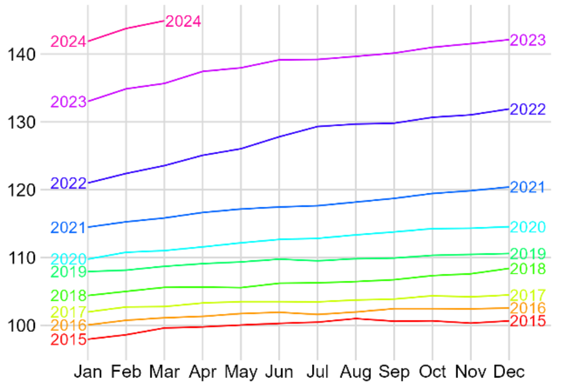
Greece

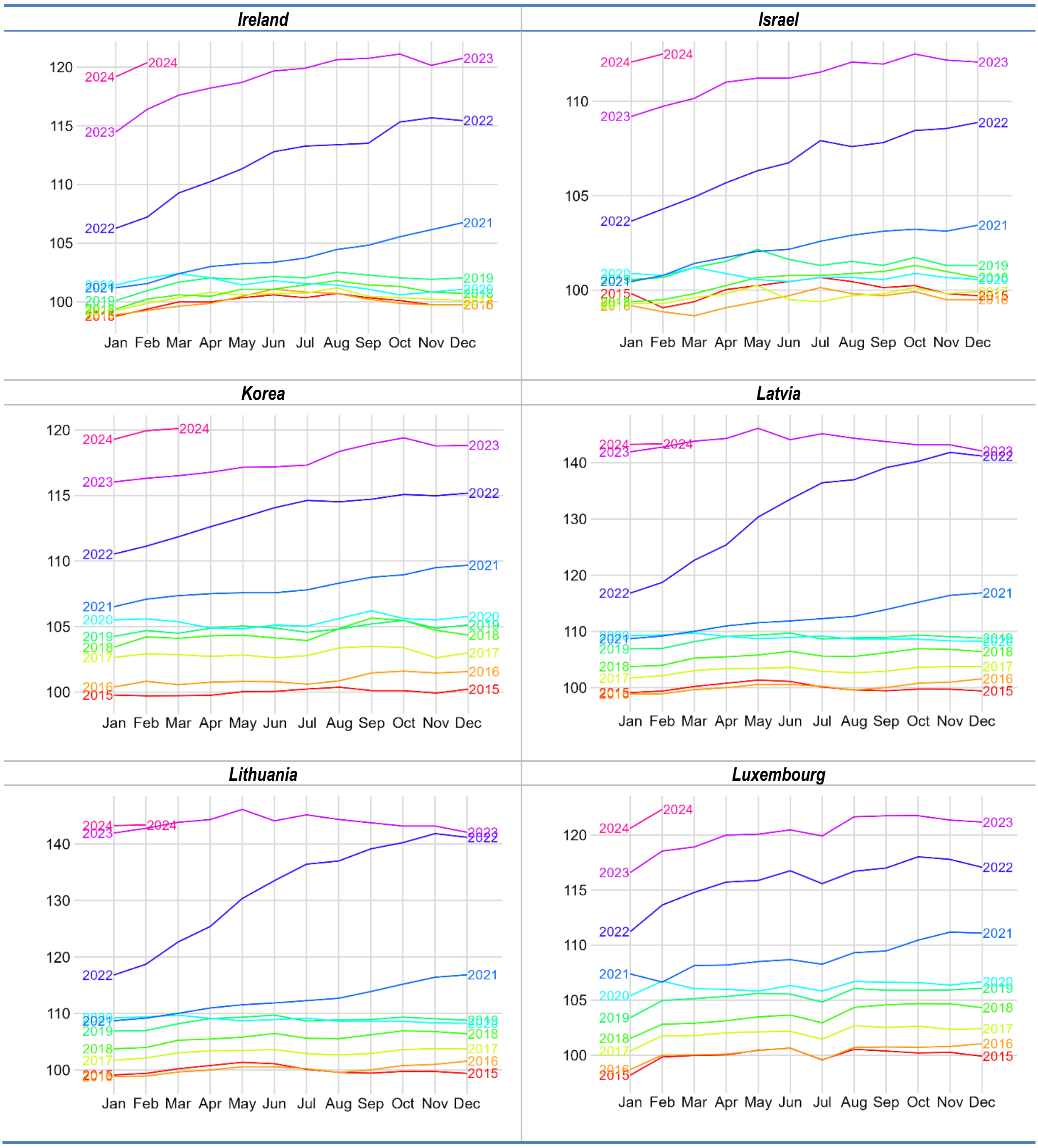


Hungary

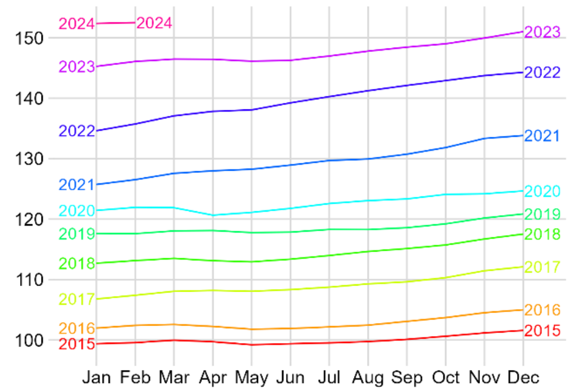


Iceland

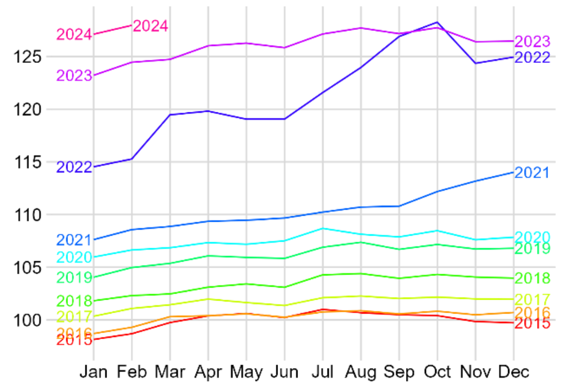



Ireland
Israel
Korea
Latvia
Lithuania
Luxembourg
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024

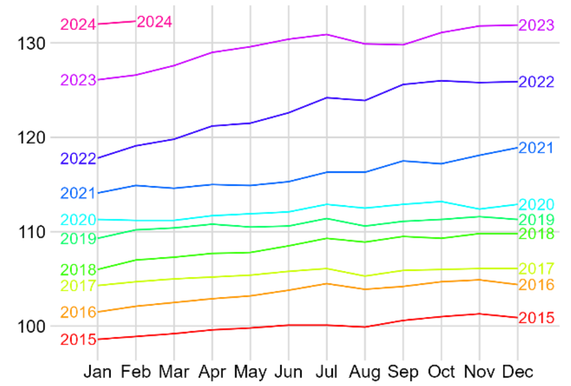
Mexico



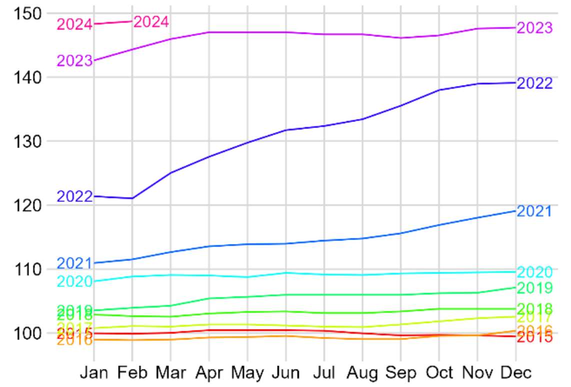
Netherlands



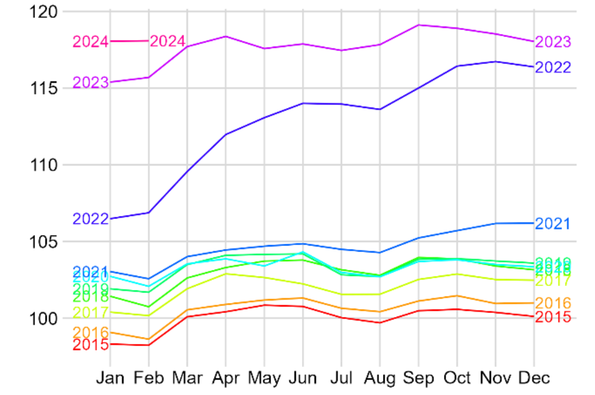
Norway



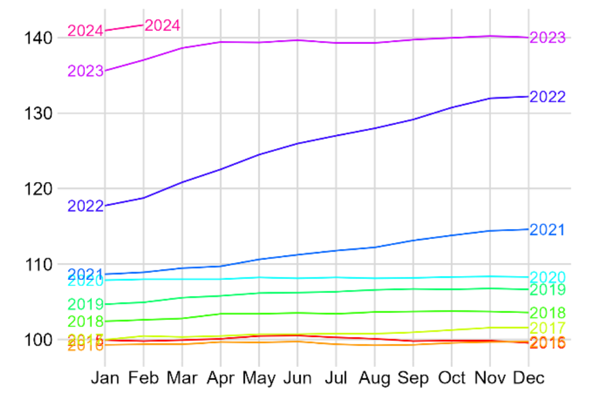
Poland



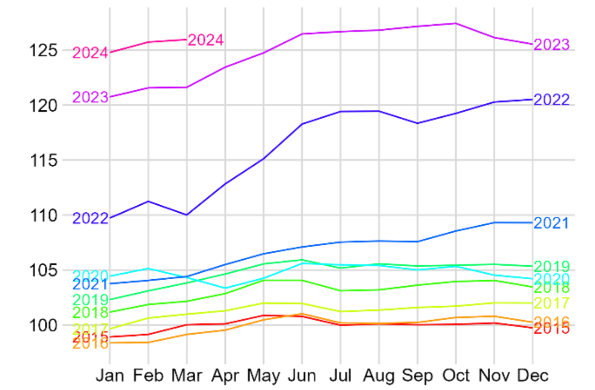
Portugal



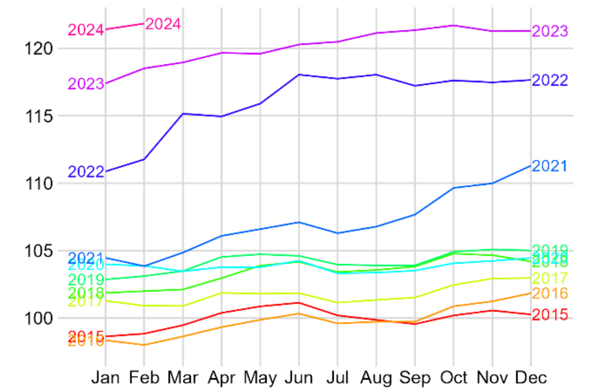
Slovak Republic



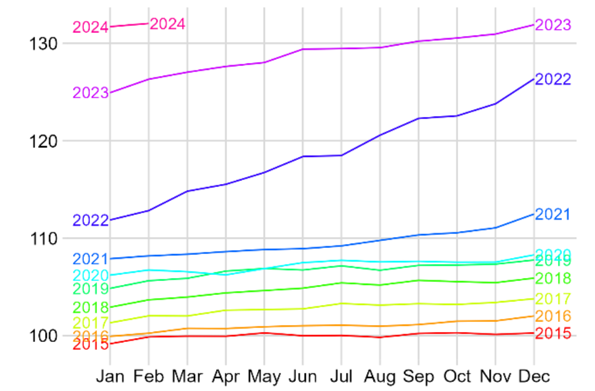
Slovenia



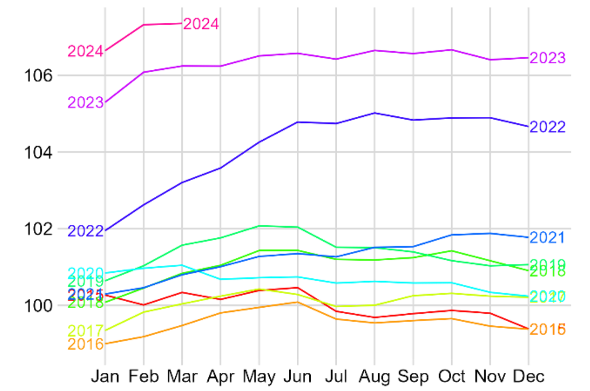
Spain



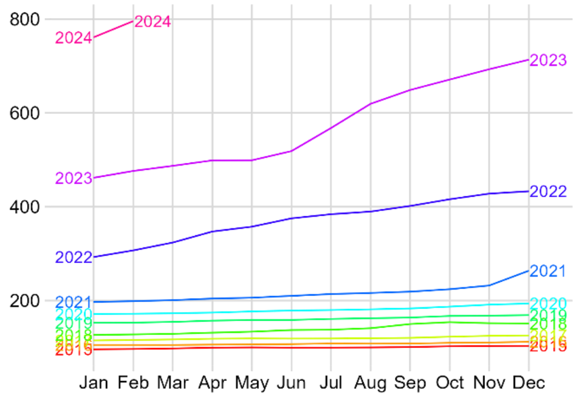
Sweden



Switzerland



Türkiye



Note: Costa Rica is based on COICOP 2018, other countries are based on COICOP 1999 and Australia and New Zealand based on quarterly data

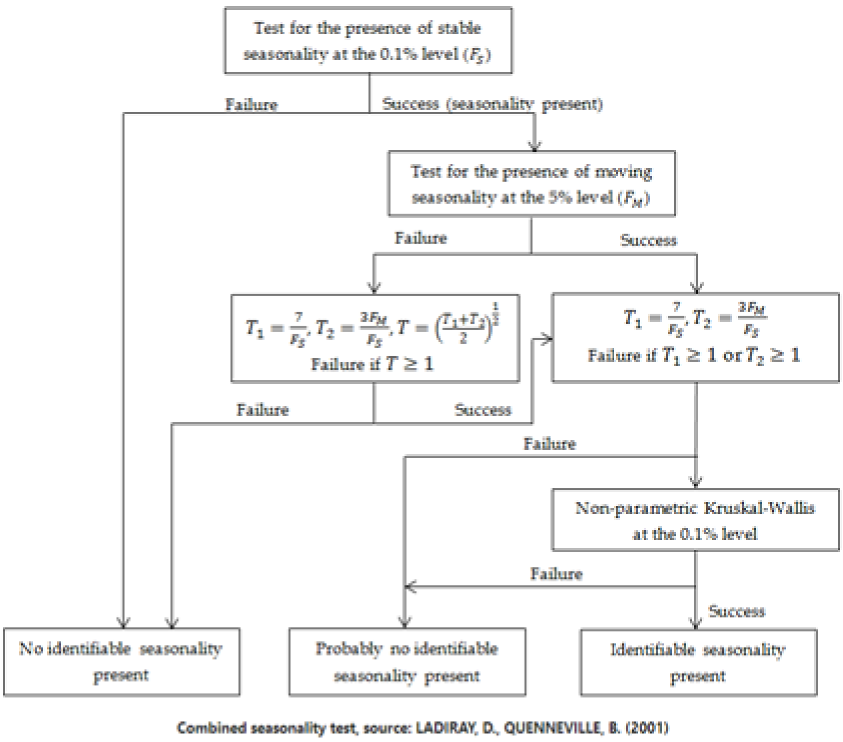
Source: OECD CPI database.

Annex B. Standard tests and methods

1. Test to detect seasonality

JDemetra+ "combined test" combines the Fisher statistic values of the parametric tests for stable (F_S) and evolving (F_M) seasonality, with the Kruskal-Wallis non-parametric test for the presence of stable seasonality (see: Seasonal Adjustment with the X-11 Method, p.65 &143). The diagramme below shows how the test works.

Figure A B.1. Seasonality test

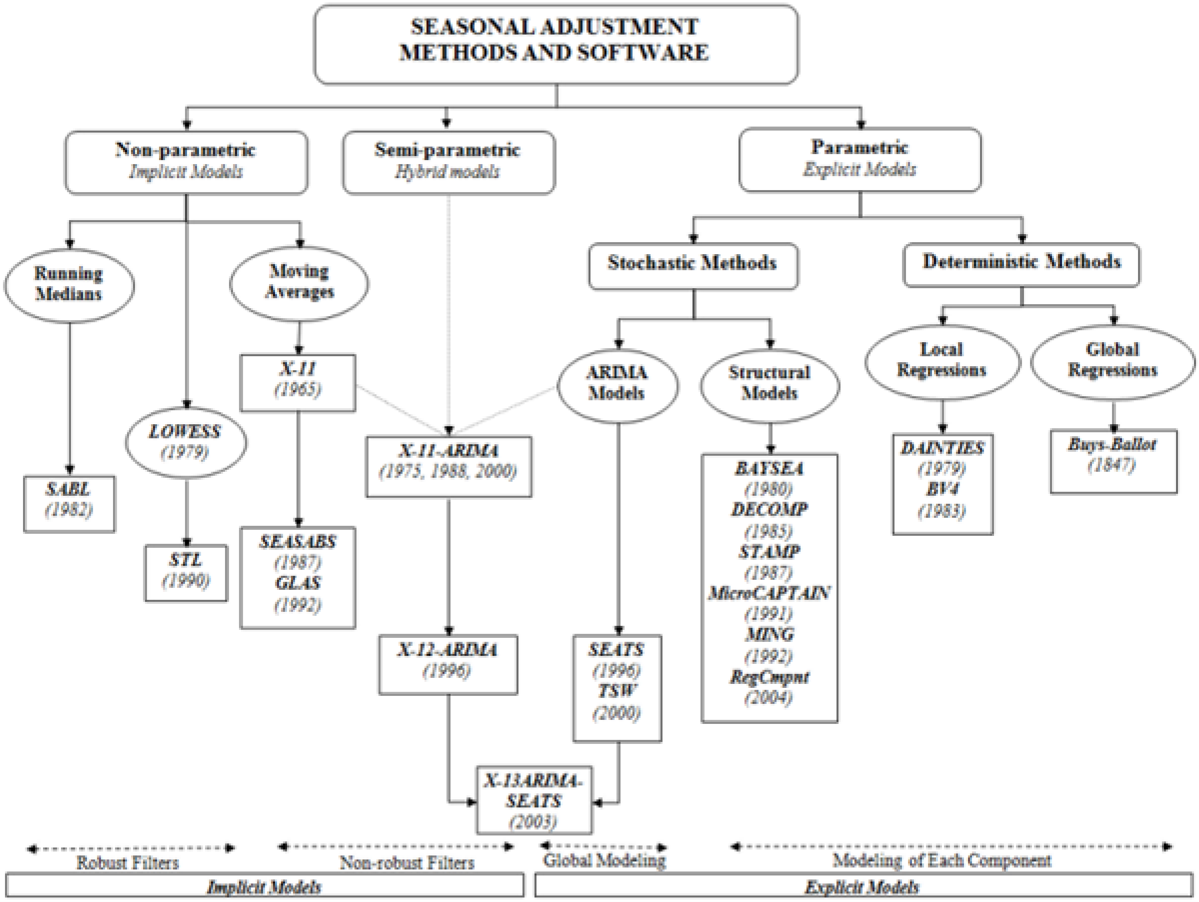


Source: Combined seasonality test (jdemetradocumentation.github.io).

2. Standard methods to adjust for seasonality

Eurostat provides a brief history and description of the different methods that are available for seasonal adjustment in its 2018 Handbook, including non-parametric, semi-parametric and parametric methods (Figure 5). Only the most common methods are discussed in this section.

Figure A B.2. Standard seasonal adjustment methods



Source: [Handbook on Seasonal Adjustment, edition 2018".](#)

Among the non-parametric methods, the best known is X-11, developed by the US Census Bureau. X-11 uses an iterative smoothing of symmetric and asymmetric moving averages and decomposes the time series into orthogonal components: the trend-cycle, seasonality and the irregular component. The method suffers from end-point bias due to the asymmetry of the filter and leads to important revisions when a new data point becomes available.

A variant of X-11, called X11-ARIMA, classified as semi-parametric method and developed by Dagum (1980), uses an ARIMA model and the Box and Jenkins methodology to backcast and forecast (from one to three years) to prolong the series and reduce the end-point bias. X12-ARIMA is a refined version of X11-ARIMA developed by the US Census Bureau in 1998 and integrates a pre-treatment "regARIMA" to estimate outliers, trading days, calendar effects with a seasonal ARIMA (SARIMA) model.

Finally, within parametric methods, based on spectral analysis, two main approaches can be distinguished:

- TRAMO-SEATS was developed by Gomez and Maravall (1997) and composed of 2 steps. First, TRAMO (Time series Regression with Arima noise, Missing observations and Outliers) detects, estimates and corrects time series for deterministic effects such as outliers, missing values and structural breaks. Second, SEATS (Signal Extraction in Arima Time Series) finds the best ARIMA model to the above TRAMO stationary linear series and decomposes it into orthogonal components (trend-cycle, seasonal and irregular) by signal extraction from the spectral density of the raw series. Parameter of the trend-cycle and seasonal components are estimated using a Wiener-Kolmogorov filter.

- STAMP (Structural Time series Analyser, Modeller and Predictor) models trend, cycles and irregular components with an ARIMA and its estimation is done with the Kalman filter method (Harvey, 1990). The software was developed by Koopman et al. (2000).

Methods have also been developed combining non-parametric and semi-parametric approaches. X13-ARIMA-SEATS is an expansion of X12-ARIMA, developed by the US Census Bureau and Bank of Spain that allows seasonal adjustment with either X-11 or SEATS method within the same interface, and comparison of results with a common set of diagnostics.

Parametric methods, require assumptions on key parameters and standard softwares such as JDemetra+ usually provide default options (Table AB.1). Five default specifications are included in the JDemetra+ package. The first two follow a simple ARIMA (0,1,1) process, which is likely not flexible enough for CPI, given the observed changes in seasonal patterns. Credible options are therefore only the other specifications, which automatically detect the ARIMA model that best fits the data. The distinction between these default specifications rests in their assumptions regarding the presence of calendar effects (absence of calendar effects, presence of a variety of calendar effects, ranging from Working Days, Trading Days, Leap year and Easter effects).

Table A B.1. Default Specifications in JDemetra+

A - TRAMO/SEATS

Identifier	Outliers' detection	Calendar effects
RSA3	Additive Outlier/Level Shift/Temporary Change	no
RSA4	Additive Outlier/Level Shift/Temporary Change	Working Day + Easter
RSA5	Additive Outlier/Level Shift/Temporary Change	Trading Day + Easter
RSAfull	Additive Outlier/Level Shift/Temporary Change	automatic

B - X-13

Identifier	Outliers' detection	Calendar effects
RSA3	Additive Outlier/Level Shift/Temporary Change	no
RSA4c	Additive Outlier/Level Shift/Temporary Change	Working Day + Easter
RSA5c	Additive Outlier/Level Shift/Temporary Change	Trading Day + Easter

Note: All the default specifications automatically test for the necessity of log-transforming in the initial time series, as well as automatically detecting and including outliers and estimating the ARIMA process of the time series. They differ when in their inclusion of calendar effects. RSAfull allows for a fully automated detection of a series' calendar effects, even deciding whether to include any Easter effects at all. RSA5 and RSA5c can detect calendar effects for trading days but assume there is an Easter effect and estimate its scale. RSA4 and RSA4c can detect calendar effects for working days, but not for days of the week. These specifications also assume there is an Easter effect and estimate it.
Source: Authors' compilation.

Most common outliers in time series

Outliers are abnormal values of a time series. In general, they cannot be properly explained by the ARIMA model and its underlying normality assumption. They tend to be associated with irregular special events that produce a distortion in the series. The presence of outliers has an adverse effect on the quality of seasonal adjustment because outliers can lead to model misspecification, biased parameter estimates, poor forecasts and inappropriate decomposition of a series in regular and irregular components.

In the automatic outlier detection and correction procedures, three outlier types are considered by default (Figure A.B. 3):

- additive outlier (AO) – an abnormal value at a limited number of isolated points of the series;
- transitory change (TC) – a series of outliers with a temporarily increasing/decreasing effect on the level of the series;
- level shift (LS) – series of outliers that have a constant long-term effect on the level of the series, i.e., a sudden permanent shift in the level of the series.

Figure A B.3. Three types of outliers (AO, TC, LS)

Source: Authors' illustration.

Annex C. COICOP 99 categories used in the analysis

- CPI All Items (COICOP 01 to 12)
- CPI Food and non-Alcoholic beverages (COICOP 01)
- CPI Alcoholic beverages, tobacco and narcotics (COICOP 02)
- CPI Clothing and footwear (COICOP 03)
- CPI Housing, water, electricity, gas and other fuels (COICOP 04)
- CPI Furnishings, household equipment and routine household maintenance (COICOP 05)
- CPI Health (COICOP 06)
- CPI Transport (COICOP 07)
- CPI Communication (COICOP 08)
- CPI Recreation and culture (COICOP 09)
- CPI Education (COICOP 10)
- CPI Restaurants and hotels (COICOP 11)
- CPI Miscellaneous goods and services (COICOP 12)

Annex D. Additional information on the simulation

Simulation: comparison X11 and JDemetra+ in normal times

Deriving the benchmark

Following Ollech (2020), the simulated CPI series with seasonality is described as follows:

$$Y_t = SA_t + S_t^{(12)} \qquad (1)$$

where SA_t is the time series free of seasonality. It follows an ARIMA (p, d, q) process of *p,* a random number following the discrete uniform distribution over the set {0, 1, 2, 3}, and *d* and *q* are random numbers drawn from the uniform set {1, 2}. The orders are drawn from an inspection of ARIMA processes underlying the observed All item CPI time series. Moving averages and autocorrelation are drawn from $MA \sim U(-1, 1)$ and $AR \sim U(-0.1, 0.1)$ so that $\sum|AR| < 1$, respectively. Error terms are drawn from a standard normal distribution. The CPI time series have been estimated for a period of 20 years using real calendars starting in January 2000.

The monthly seasonal pattern $S_t^{(12)}$, is derived from $p^{(12)} \sim N$(1.5, 0.3). Formally it resembles:

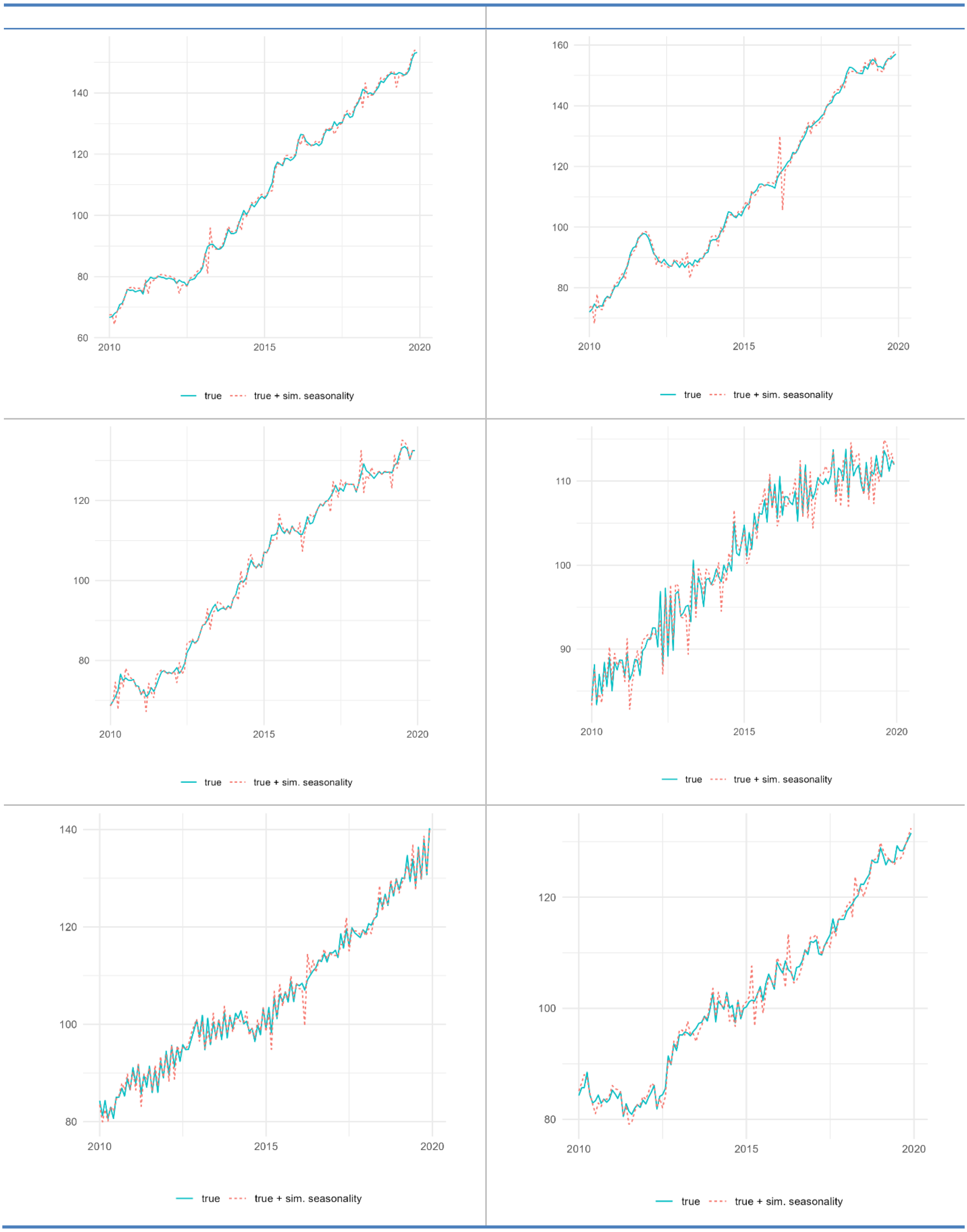
$$S_t^{(12)} = \sum_{j=1}^{J^{(12)}} \left(\beta_{j,t}^{(12)} sin\left(\frac{2\pi jG(t)}{f}\right) + \beta_{j,t}^{(12)} cos\left(\frac{2\pi jG(t)}{f}\right)\right) \qquad (2)$$

$$\beta_{j,t}^{(12)} = p^{(12)} \beta_{j,t-1}^{(12)} \qquad (3)$$

Visual inspection of Figure AC.1 confirms the assertion that the simulation produces series that closely mimic actual CPI time series.

Following this bottom-up construction, both X-13 and TRAMO-SEATS procedures are applied to Y_t and compared with the values generated by $S_t^{(12)}$.

Figure A C.1. Visualisation of simulated time series



Notes: The simulated time series are drawn randomly from all 1000 simulated time series to illustrate their validity in mimicking all-item CPI. The red line displays the simulated Y_t and the blue lines show the simulated "true" SA_t.
Source: Authors' calculations using the simulated CPI dataset.

Calendar effects

While the composition of the calendar, e.g. the number of working or trading days in a given period, can have a strong effect on activity indicators, such as industrial production, there is no a priori theoretical reason for it to have an impact on consumer prices at relatively low-frequency data (e.g. monthly time series). The only possible exceptions are moving holidays like Easter, which can take place in different months depending on the year and which have been shown to impact certain subcomponents of headline HICP in the euro area (ECB 2014, 2016). However, the examination of calendar effects in CPI series remains scant outside of the euro area.

The Easter effect was statistically tested for significance across member countries. Countries like Germany, Portugal, Italy, Belgium, Spain and France generally exhibit a significant Easter effect (Table AC.1). Japan represents a particular case; despite not having a Christian majority that widely observes Easter, statistical significance was found. This could be due to the overlap with ‘Golden Week,’ a series of holidays from the end of April to May 5th, which may have influenced the results. By contrast, in Canada, Finland, Ireland, Sweden, Switzerland and the United Kingdom, the Easter effect is very weak. It is non-existent in the United States, Mexico or Korea.

Table A.C.1. Presence of an Easter effect by country

	Strong	Weak	Absent
Europe	BEL, DEU, ESP, FRA, ITA, PRT	AUT, CHE, CZE, FIN, GBR, GRC, IRL, ISL, NLD, SWE, TUR,	DNK, EST, HUN, LUX, NOR, SVK, LAT, LTH, SVL, POL
Asia-Oceania	JPN	CHN, IDN,	AUS, KOR, NZL
America		CAN	CHL, COL, CRI, MEX, USA
Africa-Middle East		ISR, ZAF	

Source: Authors’ calculations.

There is no evidence that other calendar effects like Trading Days or Working Days effects are present in the data.

Performance of different approaches

In the first step, tests are performed using the default options RSA3, RSA4 and RSA5/RSAfull (for more details, see Annex B). These are used on the simulated raw series generated by Equation (1) to measure how closely the estimated seasonally adjusted values mirror the “true” simulated seasonally adjusted values.

The MAE of the estimated seasonal factors ranges from 0.555 to 0.558, for TRAMO-SEATS and from 0.470 to 0.738 for X-13 (Table AC2). The average for the 3 specifications also suggests that there are no differences between TRAMO-SEATS and X-13. The same conclusions can be drawn from the RMSE and MAPE. One finding that stands out, however, is that for the X-13 method, the RSA3 significantly underperforms the other alternatives. For TRAMO-SEATS no such difference is observed.

Table A.C. 2. X-13 and TRAMO-SEATS performance using default specifications

Average of 1000 simulated time series over 20 years

Method	Specification	RMSE	MAE	MAPE	Average
TRAMO-SEATS	RSAfull	1.24	0.555	0.930	0.908
	RSA5	1.24	0.558	0.930	0.909
	RSA4	1.24	0.555	0.930	0.908
	RSA3	1.26	0.554	0.922	0.912
X-13	RSA5c	1.07	0.470	0.810	0.784
	RSA4c	1.08	0.473	0.815	0.789
	RSA3	1.98	0.738	1.32	1.346

Note: The table presents a comparison of seasonally adjusted simulated CPI series (Y_t) according to the method and specification options in columns 1 and 2, with SA_t which is Y_t free of seasonality. The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE measures how concentrated the estimated data is around the simulated values. The Mean Absolute Error (MAE) is the arithmetic average of the absolute errors between the estimated and the simulated values. It tells us how far the simulated values are off the estimated ones, on average. The Mean Absolute Percentage Error (MAPE) converts the MAE into relative values. It calculates how far off the simulated values the estimated values are on average, in terms of percentages. All the default specifications in column 2 automatically test for the necessity of log-transforming in the initial time series, as well as automatically detecting and including outliers and estimating the ARIMA process of the time series. They differ when in their inclusion of calendar effects. RSAfull allows for a fully automated detection of a series' calendar effects, even deciding whether to include any Easter effects at all. RSA5 and RSA5c can detect calendar effects for trading days but assume there is an Easter effect and estimate its scale. RSA4 and RSA4c can detect calendar effects for working days, but not for days of the week.
Source: Authors' calculations using simulated CPI data.

Table A.C.3. Comparison of initial time series transformation choices using X-13 and TRAMO-SEATS

Average of 1000 simulated time series over 20 years

Method	Specification Options	RMSE	MAE	MAPE
	Transformation			
TRAMO-SEATS	Auto	1.20	0.627	1.34
	Log	1.88	0.970	1.63
	None	1.19	0.613	1.33
X-13	Auto	1.47	0.674	1.56
	Log	1.98	0.857	1.59
	None	1.47	0.674	1.56

Note: Note: The table presents a comparison of seasonally adjusted simulated CPI series (Y_t) according to the method and specification options in columns 1 and 2, with SA_t which is Y_t free of seasonality. The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE measures how concentrated the estimated data is around the simulated values. The Mean Absolute Error (MAE) is the arithmetic average of the absolute errors between the estimated and the simulated values. It tells us how far the simulated values are off the estimated ones, on average. The Mean Absolute Percentage Error (MAPE) converts the MAE into relative values. It calculates how far off the simulated values the estimated values are on average, in terms of percentages. When Auto Transformation is selected, JDemetra+ will conduct a log/level test to determine whether to use levels or logs. The test compares the sum of squares of the model without logs with the sum of squares multiplied by the square of the geometric mean from the model in logs. Logs are taken in case the last function is the maximum. Source: Authors' calculations using simulated CPI data
Source: Authors' calculations using simulated CPI data.

In a second step, tests are undertaken to move away from the default options, by exploring options relating to the underlying ARIMA model. This includes looking at the selection between additive and multiplicative model types (log-test) and identifying the ARIMA model fitting the time series, with a particular focus on the selection of the order of differentiation. No significant differences across the various methods and options can be observed (Table AC3 and AC4).

Table A.C.34. Comparison between first and second differenced ARIMA models using X-13 and TRAMO-SEATS

Average of 1000 simulated time series over 20 years

Method	Specification Options	RMSE	MAE	MAPE
	ARIMA Differencing			
TRAMO-SEATS	First	1.20	0.627	1.34
	Second	1.88	0.970	1.63
X-13	First	1.47	0.674	1.56
	Second	1.98	0.857	1.59

Note: The table presents a comparison of seasonally adjusted simulated CPI series (Y_t) according to the method and specification options in columns 1 and 2, with SA_t which is Y_t free of seasonality. The results are averaged over 1000 randomly generated CPI series. The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE measures how concentrated the estimated data is around the simulated values. The Mean Absolute Error (MAE) is the arithmetic average of the absolute errors between the estimated and the simulated values. It tells us how far the simulated values are off the estimated ones, on average. The Mean Absolute Percentage Error (MAPE) converts the MAE into relative values. It calculates how far off the simulated values the estimated values are on average, in terms of percentages.
Source: Authors' calculations using simulated CPI data.

In the third step, different specifications that explicitly try to capture the Easter effect are explored. In accordance with findings in the previous section, other calendar effects were not studied. It should be noted that the simulated time series include a randomly sized calendar effect in the week leading up to Easter, identified using the real calendar from 2000 onwards. The options to account for Easter Effects differ between TRAMO-SEATS and X-13. By default, in X-13 the Easter effect includes the 8 days leading up to and including Easter Sunday, while TRAMO-SEATS considers a period of 6 days. The duration of the Easter effect can be modulated for both TRAMO-SEATS and X-13. For TRAMO-SEATS the Easter effect can be defined as any duration between 1 and 15 days, whereas for X-13 it can extend from 1 to 20 days. TRAMO-SEATS offers further customisation options, such as the inclusion of Easter Monday in the Easter period. This can be useful when Easter Monday falls in a different month than the rest of the Easter period.

The inclusion of an Easter effect improves the performance of seasonal adjustment procedures by 0.29 percentage point (MAE) on average for TRAMO-SEATS and by 0.44 percentage point (MAE) on average for X-13, when an Easter effect is present in the time series (Table 6). However, due to the nature of the simulations, the size of this improvement might be a statistical artefact that is due to the randomly generated size of the calendar effects.

The standard duration (6 and 8 days respectively) provides the best performance for both TRAMO-SEATS and X-13. Since the simulated Easter effect is specified for a 7-day duration, the default specifications capture it best. Despite greater options to account for Easter effects using the TRAMO-SEATS procedure, X-13 slightly outperforms it.

Table A.C.5. Comparison of different Easter effect specifications using X-13 and TRAMO-SEATS

Average of 1000 simulated time series over 20 years

Method	Specification Options		RMSE	MAE	MAPE
	Easter Effect	Duration			
TRAMO-SEATS	None	NA	1.83	0.816	1.66
	Easter	3	1.11	0.534	0.970
	Easter	6 (default)	1.03	0.494	0.889
	Easter	14	1.14	0.555	0.991
	Easter + Easter Monday	3	1.09	0.523	0.903
	Easter + Easter Monday	6 (default)	1.04	0.503	0.875
	Easter + Easter Monday	14	1.10	0.537	0.960
	Auto	3	1.13	0.522	0.919
	Auto	6 (default)	1.03	0.484	0.851
	Auto	14	1.22	0.597	1.08
X-13	None	NA	1.42	0.640	1.22
	Auto	3	0.981	0.473	0.831
	Auto	8 (default)	0.981	0.473	0.831
	Auto	14	0.981	0.473	0.831

Note: When Auto Easter inclusion is selected, JDemetra+ will conduct a pre-test for the significance of the Easter effect based on the t-statistic, where the Easter effect is included via a regressor if the t-statistic is greater than 1.96. The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE measures how concentrated the estimated data is around the simulated values. The Mean Absolute Error (MAE) is the arithmetic average of the absolute errors between the estimated and the simulated values. It tells us how far the simulated values are off the estimated ones, on average. The Mean Absolute Percentage Error (MAPE) converts the MAE into relative values. It calculates how far off the simulated values the estimated values are on average, in terms of percentages.
Source: Authors' calculations using simulated CPI data.

Table A.C.6. Correlation between OECD and NSO seasonal adjustment (month-on-month series)

Country	X-13	TRAMO-SEATS
Canada	0.939	0.929
Germany	0.826	0.846
France	0.935	0.922
Japan	0.913	0.894
United States	0.952	0.955

Note: The X-13 column represents the Pearson's r correlation coefficient between the seasonally adjusted values obtained by the OECD using the X-13 default specification (RSA5c) and the seasonally adjusted values provided by the NSOs. The TRAMO-SEATS column represents the Pearson's r correlation coefficient between the seasonally adjusted values obtained by the OECD using the TRAMO-SEATS default specification (RSA5) and the seasonally adjusted values provided by the NSOs. The initial values are transformed into month-on-month changes.
Source: Authors' calculation using the OECD's CPI dataset.

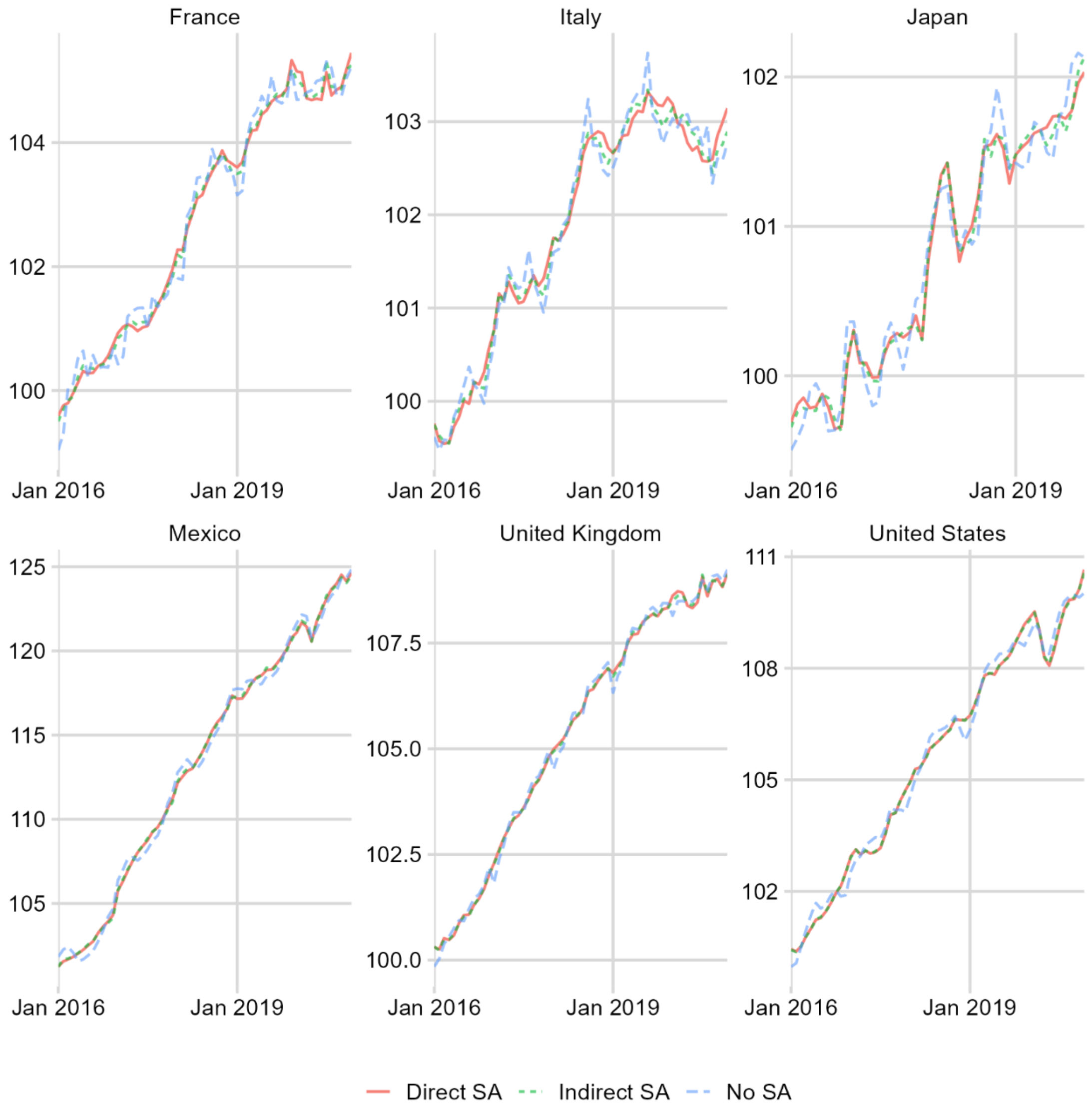
Simulation direct vs indirect

Lowe index

The formula employed is equivalent to $\mathrm{I}_{\mathrm{L}}^{0:t} = \sum \mathrm{w}_{\mathrm{j}}^{\mathrm{b}} \; \mathrm{I}_{\mathrm{j}}^{0:t}, \sum \mathrm{w}_{\mathrm{j}}^{\mathrm{b}} = 1$ where $\mathrm{I}_{\mathrm{L}}^{0:t}$ denotes the All items CPI, from period 0 to t, and $\mathrm{w}_{\mathrm{j}}^{\mathrm{b}}$ is the weight attached to each of the elementary price indices, recorded in period b (usually set to precede t by up to 1 year). $\mathrm{I}_{\mathrm{j}}^{0:t}$ is the corresponding elementary price index in the respective time period 0 to t. The 12 elementary indices are identified by the subscript j.

Figure A C.1. Direct vs Indirect SA indices (X-13)

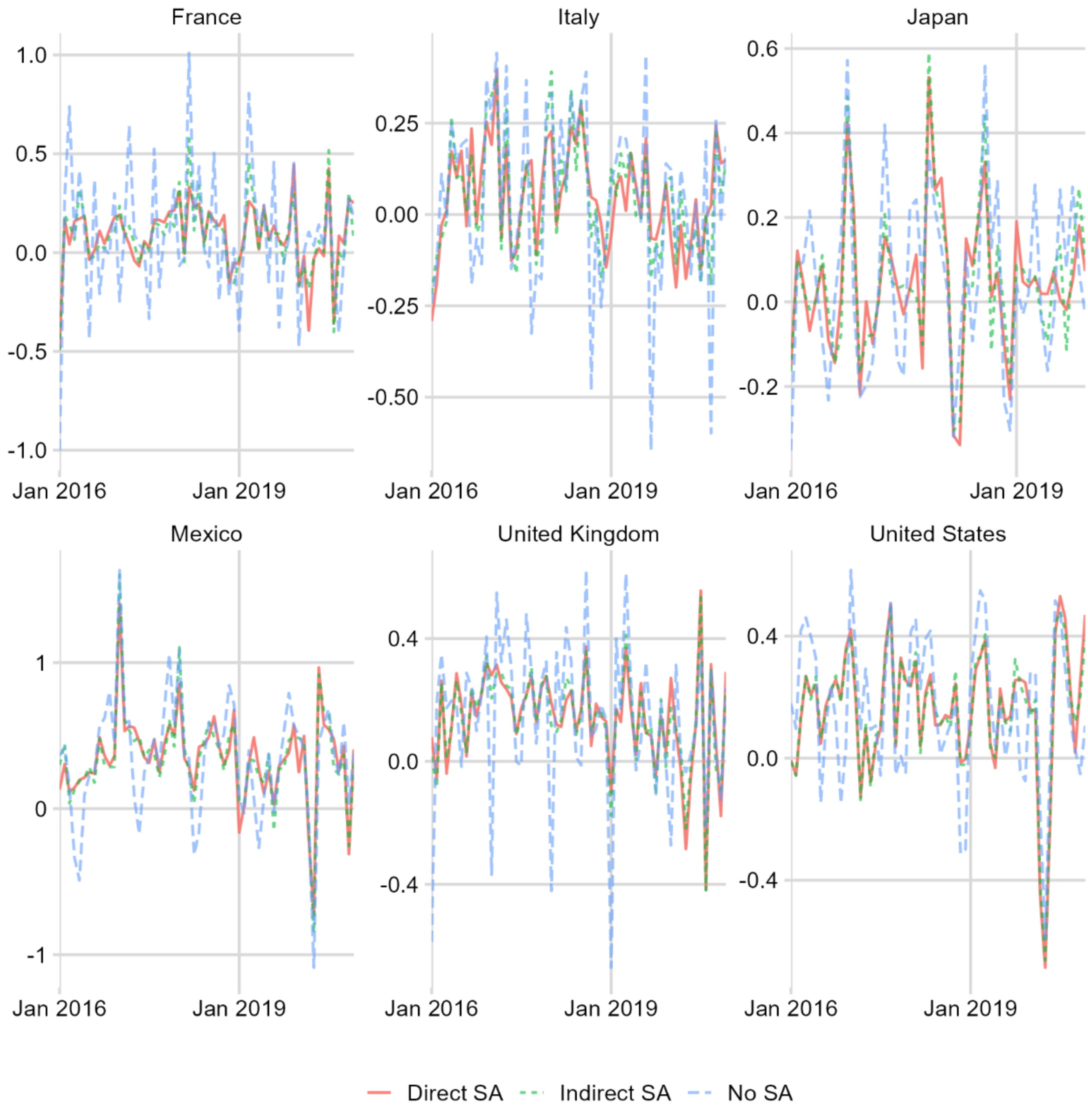
CPI All Item, absolute values



Source: Authors' calculations using the CPI dataset.

Figure A C.2. Direct vs Indirect SA month-on-month change (X-13)

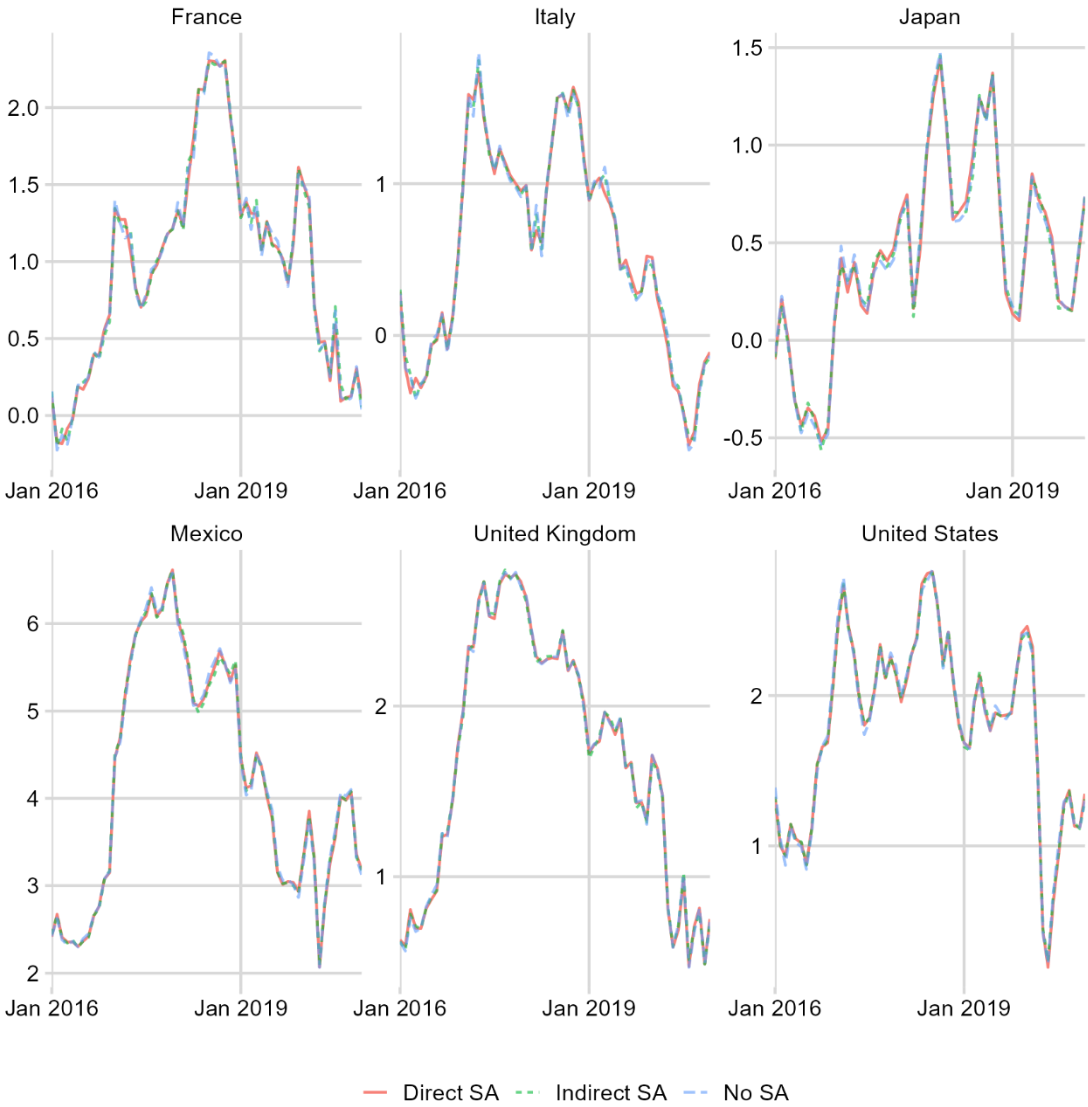
CPI All item, month=on-month changes (%)



Source: Authors' calculations using the CPI dataset.

Figure A C.3. Direct vs Indirect SA year-on-year change (X-13)

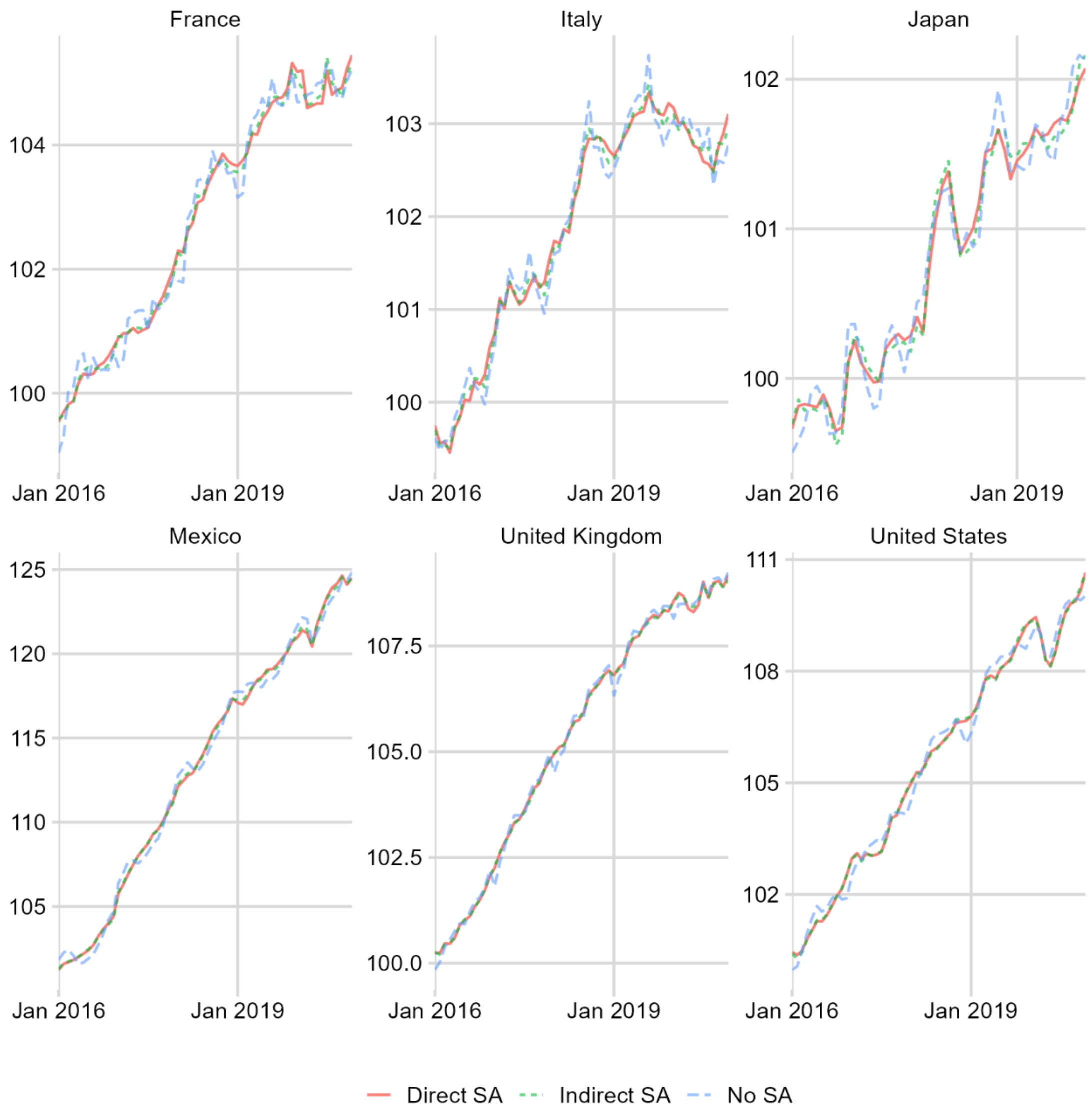
CPI All item, year-on-year changes (%)



Source: Authors' calculations using the CPI dataset

Figure A C.5. Direct vs Indirect SA indices (TRAMO-SEATS)

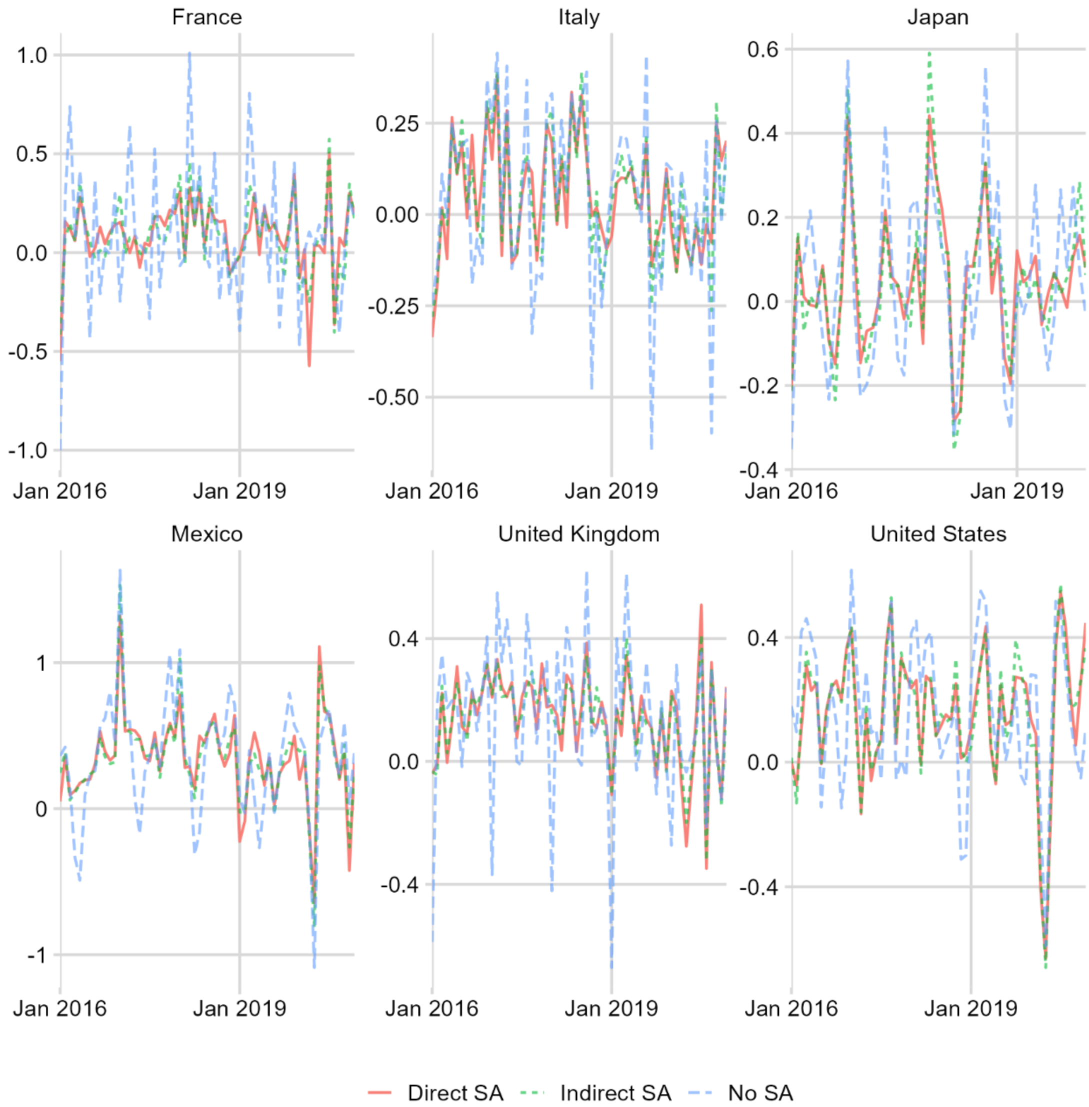
CPI All item, absolute values



Source: Authors' calculations using the CPI dataset.

Figure A C.6. Direct vs Indirect SA month-on-month change (TRAMO-SEATS)

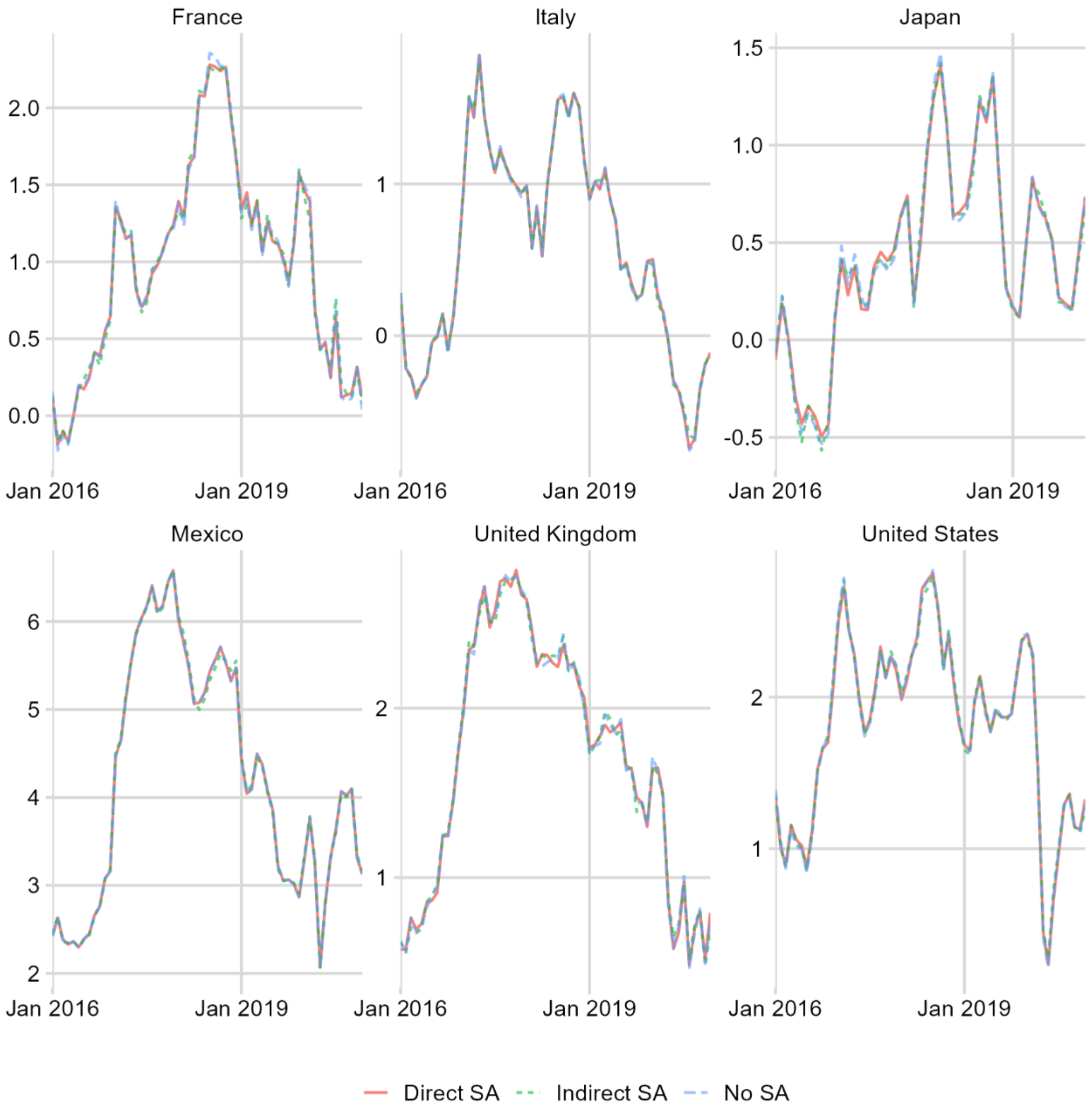
CPI All item, month-on-month changes (%)



Source: Authors' calculations using the CPI dataset.

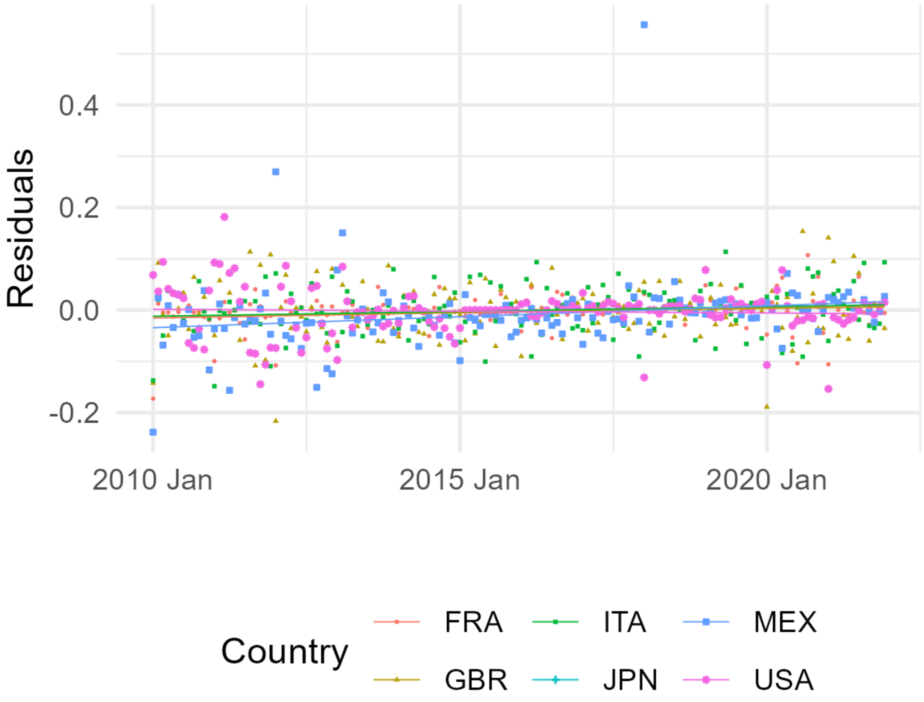
Figure A C.7. Direct vs Indirect SA year-on-year change (TRAMO-SEATS)

CPI All item, year-on-year changes (%)



Source: Authors' calculations using the CPI dataset.

Figure A C.8. Calculated All-items CPI and NSO All-items CPI comparison



Source: Authors’ calculations using the CPI dataset.