Novel data sources for business cycle analysis and seasonal adjustment

Daniel Ollech*

Abstract
The Covid-19-pandemic has increased the need for timely and granular data that allow the assessment of the state of the economy in real time. As these time series often contain weekly or even daily observations, they present a challenge to traditional seasonal adjustment methods. We show how the calendar and seasonal effects of such economic indicators can be estimated reliably using the DSA approach for daily time series and an ARIMA-based adjustment strategy for weekly data. Drawing on a set of time series namely electricity consumption, traffic, and Google Trends data used in many countries to assess the economic development during the pandemic, we discuss obstacles, difficulties, and best-practices.

Key Words: Covid-19; DSA; Calendar adjustment; higher frequency time series

1. Motivation

During the last decade, the number of seasonal adjustment methods for higher frequency time series has seen a considerable increase (see Ladiray et al, 2018; Webel, 2020). This goes hand in hand with new data sources becoming available both from official data providers and private companies. This development has gained a lot of traction during the Covid-19-pandemic, as policy makers and economists craved more timely and granular information on the state of the economy.

These time series often present methodological challenges that differ substantially from those encountered in lower frequency data. In this paper, we will therefore analyse an exemplary set of series that is typical for the kind of data relied upon by economists and business cycle analysts during the Covid-19-pandemic. Here, we will focus on the features and idiosyncracies of these time series relevant in the context of seasonal adjustment.

To this end, Section 2 presents the seasonal adjustment methods employed. Section 3 analyses and adjusts the daily truck toll mileage and electricity consumption and a weekly Google Trends series while section 4 derives the characteristic features of higher frequency time series from a seasonal adjustment perspective. Section 5 summarises.

2. Seasonal adjustment methods

Let $\{Y_t\}$ denote a series of length T with τ observations per year on average. The basic time series decomposition is then given by

$$Y_t = T_t + S_t + C_t + I_t \tag{1}$$

which includes the trend-cycle (T_t), seasonal (S_t), calendar (C_t) and irregular component (I_t). It can easily be adapted to capture a multiplicative relationship between the components by log-transformation of the original time series.[1] For multiple periodic effects with cycle lengths $\tau_1, \tau_2, \ldots$, the basic time series model can be generalised to

$$Y_t = T_t + \sum_i S_t^{(\tau_i)} + C_t + I_t \tag{2}$$

*Deutsche Bundesbank, Central Office, Directorate General Statistics, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany

[1] For a thorough discussion of further decomposition models, see U.S. Census Bureau (2016).

Remark. *In Equation 2 the periodic effects are subsumed as seasonal effects to allow a parsimonious notation. Here, effects related to calendar constellation are not considered to be periodic effects, even though the calendar constellation recurs with a cycle length of* 400 *years in the Gregorian calendar. Yet, this recurrence is not exploited in the modelling of the time series and usually effects arising from these calendar constellations will not recur in similar intensity every* 400 *years.*

As we will see, for some series, the seasonal and calendar effects are interdependent. If all interactions are relevant, Equation 2 needs to be augmented so that

$$Y_t = T_t + \sum_i S_t^{\tau_i} + C_t + \sum_i \sum_{j>i} S_t^{(\tau_i)} * S_t^{(\tau_j)} + \sum_i S_t^{\tau_i} * C_t + I_t. \tag{3}$$

In line with the methods used for official seasonal adjustment - namely X-13 and TramoSeats - the adjustment of the higher frequency time series presented here combines a seasonal adjustment routine with a RegARIMA based pre-adjustment. The latter includes the estimation and elimination of calendar effects and is given by

$$\phi_p(B)\phi_P(B^\tau)(1-B)^d(1-B^\tau)^D\left(Y_t - \sum_{i=1}^{r}\beta_i X_{it}\right) = \theta_q(B)\theta_Q(B^\tau)\varepsilon_t \tag{4}$$

where $\phi(B)$ and $\theta(B)$ are AR and MA polynomials of order p and q, number of differences d, and capitals indicating seasonal terms while B is the backshift operator, i.e. $B(y_t) = y_{t-1}$. β_i captures the impact of the i-th regressor X_{it} on the time series Y_t and ε_t is the error term. The ARIMA part of Equation 4 can be abbreviated by $(p\,d\,q)(P\,D\,Q)_\tau$. Extensions of this model to multiple seasonalities are known and available (e.g. Svetunkov, 2017), but at the time of writing, these are rarely used in seasonal and calendar adjustment.

2.1 Seasonal adjustment of daily time series

The iterative DSA procedure described by Ollech (2018) combines the aforementioned RegARIMA model with STL (Cleveland et al., 1990).[2]

STL decomposes a time series into T_t, S_t and I_t using a series of Loess regressions and moving averages. Loess regressions are local weighted regressions (Cleveland and Devlin, 1988), where the locality is defined by the number of neighbouring observations included, given by γ_τ. As STL only extracts one periodic pattern at a time, DSA combines multiple runs of STL with RegARIMA:

- Step I: Adjust intra-weekly seasonality with STL.
- Step II: Calendar- and outlier adjustment with RegARIMA.
- Step III: Adjust intra-monthly seasonality with STL.
- Step IV: Adjust intra-annual seasonality with STL.

2.2 Seasonal adjustment of weekly time series

On average a year contains a non-integer number of weeks, namely 52.18. Ladiray et al. (2018) describe how in these cases ARFIMA processes can be exploited that adapt the seasonal differencing part of Equation 4 to incorporate fractionally integrated processes.

[2] DSA can integrate other seasonal adjustment methods as well. Ollech et al. (2021) discuss how to flexibilise X-11 to handle higher frequency time series.

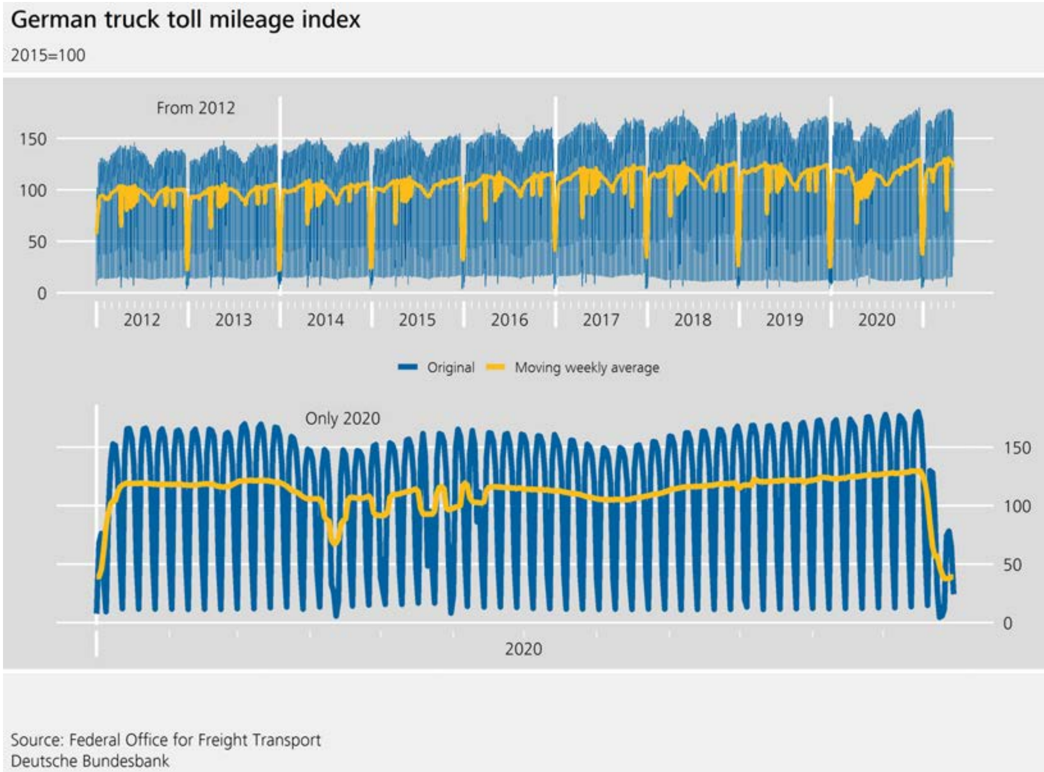

Figure 1: German index of truck toll mileage.

The authors develop a fractional variant of the well-known airline model, i.e. the seasonal ARIMA model of order $(0,1,1)(0,1,1)_\tau$. For a fractional seasonal period of $\tau = \lfloor \tau \rfloor + \alpha$, with $\alpha \in [0,1]$, the fractional differencing operator $\tilde{\nabla}_\tau$ can be approximated by a first order Taylor series expansion so that

$$\tilde{\nabla}_\tau Y_t \approx Y_t - (1-\alpha)B^{\lfloor \tau \rfloor} Y_t - \alpha B^{\lfloor \tau \rfloor + 1} Y_t \tag{5}$$

The fractional airline model can be used for linearisation of a time series analogously to Tramo.[3] We combine this pre-processing with a SEATS-type time series decomposition of the fractional airline model to seasonally adjust weekly time series with $\tau = 52.18$.

3. Exemplary analysis of higher frequency data

We will present a small set of higher frequency time series that have been of high importance for business-cycle analysis during the Covid-19-pandemic. We will discuss key features of this series and show how these might be accounted for when conducting calendar and seasonal adjustment.

3.1 Truck toll mileage index

The Federal Office for Freight Transport in Germany is responsible for a distance-based toll on trucks, which has been implemented in January, 2005. The truck toll mileage index has been developed together with the Federal Statistical Office and is based on the raw mileage

[3]The fractional airline model based estimation and decomposition are included in the R Package {rjdhf} available on github: https://github.com/palatej/rjdhighfreq. As of version 0.0.4, the Tramo-type estimation of the model does not allow forecasting of the time series. For weekly time series, we therefore use seasonal ARIMA models for pre-processing of the time series with $\tau = 52$.

and free of structural breaks resulting from changes in the vehicles that have to pay the toll. The data are available as a monthly and a daily index (Deutsche Bundesbank, 2020). The daily time series analysed here is available from January 1, 2005 up to April 3, 2021, and thus contains **many observations**.

As can be seen from Figure 1 the series is characterised by **multiple periodic effects**, namely a strong week-day pattern, with a trough at Sundays and an annual pattern. These effects are interdependent: the **cross-seasonality** presents itself as a week-day pattern that changes throughout the year. In part this is due to governing laws and regulations: with only a few exceptions trucks are not free to drive on motorways on Sundays and public holidays. During July and August, this is extended to Saturdays as well. We further observe different patterns around Christmas related to different changes in the consumption behaviour.

Remark. *Other series contain* ***unusual periodic effects****. Ollech (2018) finds a monthly recurring pattern in the German currency in circulation.*

The series is further marked by a weakly positively sloping trend that is halted temporarily in early 2020, as a consequence of the Covid-19-Pandemic. After the Covid-slowdown we can observe **breaks in periodic effects**, in particular the weekday pattern. At least temporarily, the difference in the truck mileage on workday and weekday is less pronounced. The annual seasonal pattern is **non-isochronous**, i.e. the number of observations per cycle is not the same for all cycles as years contain either 365 or 366 observations.

Remark. *Non-isochronicity can impact a time series in a number of ways. Especially with a continuously evolving seasonal pattern, we may observe that the seasonal pattern in longer cycles is just a stretched-out version of the pattern in shorter cycles. In series with a more variable pattern, the seasonal impact of additional observations in a given cycle, such as February 29 in a leap year, may be barely related to adjacent observations.*

For higher frequency time series, it is often possible to observe the impact of holidays more directly, and thus identify **series-specific calendar effects**. Figure 2 shows the truck toll mileage index on All Saints' Day as well as the three days leading up to and the three days after the holiday. All Saints' is a public holiday in 5 out of 16 German federal states. In these states, all being located either in the south or west of the country, trucks are prohibited from driving on motorways. Therefore, the impact of All Saints' Day is cross-seasonal: the magnitude of the decrease depends on the weekday.[4] As driving trucks is already restricted on Sundays, there is no additional impact of All Saints' Day on the truck toll mileage index if it falls on a Sunday. In contrast, there is a considerable reduction if it falls on any weekday.

If All Saints' Day falls on a Tuesday (Thursday) the neighbouring Monday (Friday), i.e. any **bridge days**, also appear to have a reduced number of trucks on motorways. In contrast to lower frequency time series, these bridge day effects can be seen and modelled more directly.

The Covid-19-pandemic hitting Germany in March, 2020 has a two-fold impact on the time series. First, it abruptly leads to a decreasing number of kilometers driven by trucks. Second, it impacts the observed weekday pattern, due to changes in the restrictions and the consumption pattern.

To obtain a calendar and seasonally adjusted series, DSA has been used with $\gamma_7 := 7$ and $\gamma_{365} := 11$ using an additive decomposition. For the seasonal adjustment of the monthly index, we found a multiplicative decomposition to be more appropriate. Yet, due

[4] Reformation day, which is the day before All Saints' Day is a public holiday in all of the federal states in East Germany. In 2017, Reformation day was a national public holiday. From 2018, Reformation day is a public holiday in the federal states in northern Germany.

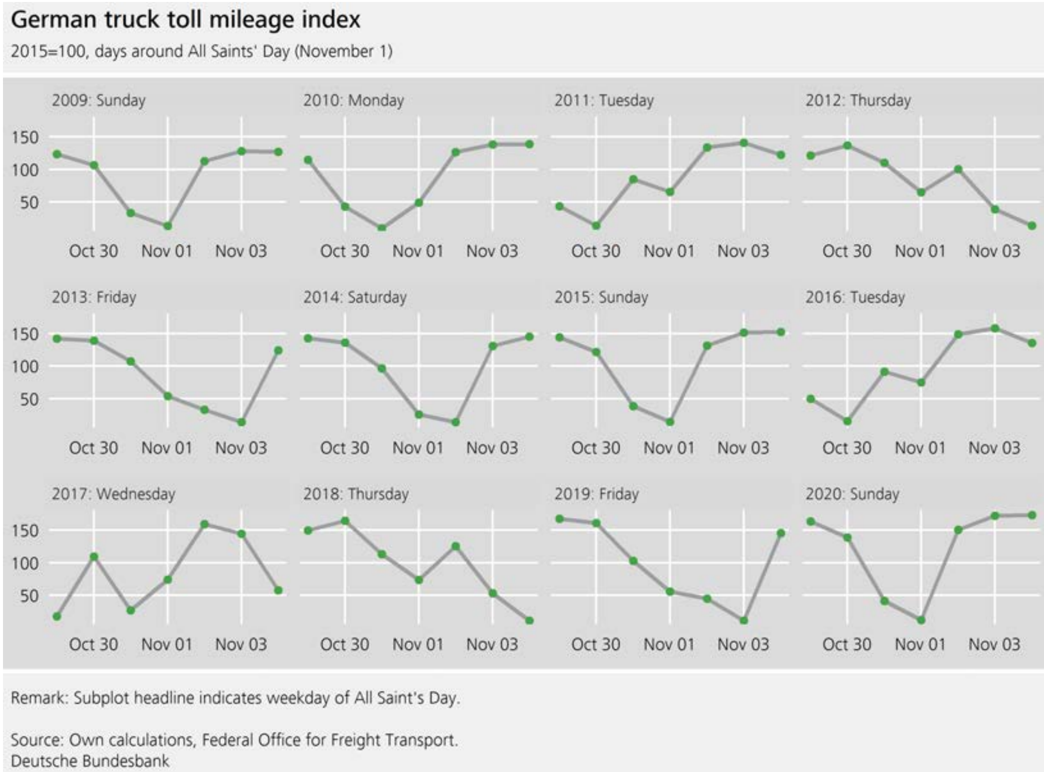

Figure 2: German index of truck toll mileage.

to the **higher volatility** of the daily index - which is typical for higher frequency time series - the seasonal factors from a multiplicative model would inflate a substantial number of unusual observations. The series does not contain day-of-the-month effects, accordingly, step III of the DSA procedure is omitted. The choice of a very short filter to estimate the day-of-the-week effect, i.e. $\gamma_7 := 7$, is a means to allow the day-of-the-week-effect to change during the year and thereby to capture this particular cross-seasonal effect. As discussed above, this is important, because we already know that the weekday pattern is different in different parts of the year, due to restrictions. An important tool to gauge an appropriate parameter for γ_τ are SI-ratios (Cleveland and Terpenning, 1982).

To estimate the impact of moving holiday effects and the interaction between weekdays and fixed holidays, a RegARIMA model is used. As described by Ollech (2018) this model combines a non-seasonal ARIMA model with trigonometric terms that capture deterministic seasonality. For the truck toll mileage index, 30 cosine and sine terms are used. Ollech (2018) correctly states that multiples of 12 capture intra-monthly pattern, yet here, the high number of trigonometric terms used reflects the complexity of the seasonal pattern and does not indicate a day-of-the-month effect.

Remark. *To obtain the unadjusted monthly truck toll mileage index, the daily raw truck toll mileage is summed up for each month and transformed to an index. If likewise a* ***temporal aggregation*** *is performed on the seasonally adjusted truck toll mileage index via summation, the resulting time series will contain a length-of-the-month effect - i.e. months with more days will have a higher value - and thus be seasonal. Using the average monthly index instead is a simple remedy, but it reduces the comparability of the directly adjusted monthly and daily time series.*

The calendar and seasonally adjusted truck toll mileage index is especially volatile around holidays (see Figure 3). This **Heteroskedasticity** is due to estimation uncertainty

Table 1: Estimated moving holidays and cross-seasonal effects for the German truck toll mileage index. Based on time series only adjusted for intra-weekly seasonal effects. A RegARIMA(2,1,1) model with 2x30 trigonometric terms has been estimated.

	Estimate	S.E.		Estimate	S.E.
Carnival Monday	−8.6	1.6	NH[1] (Mon)	−92.3	2.3
Mardi Gras	−5.0	1.6	NH[1] (Tue)	−105.9	2.4
Holy Thursday	−20.0	1.6	NH[1] (Wed, to 2015)	−90.2	3.1
Good Friday	−99.6	1.9	NH[1] (Wed, from 2015)	−118.7	2.7
Holy Saturday	−17.3	1.9	NH[1] (Thu)	−106.4	2.1
Easter Sunday	−11.1	2.0	NH[1] (Fri)	−74.5	1.9
Easter Monday	−116.8	1.9	NH[1] (Sat)	−21.2	2.1
Easter Mon. (t+1)	−20.4	1.6	Xmas Period (Mon)	−31.7	1.8
Ascension (t–1)	−13.6	1.6	Xmas Period (Tue)	−42.6	1.8
Ascension (to 2015)	−108.9	2.3	Xmas Period (Wed)	−42.0	1.8
Ascension (from 2016)	−122.1	2.5	Xmas Period (Thu)	−41.7	1.8
Ascension (t+1)	−17.4	1.7	Xmas Period (Fri)	−29.0	1.8
Corpus Chr. (t–1, to 2015)	−7.3	2.0	3d before Xmas (Sun)	120.1	18.9
Corpus Christi (to 2015)	−73.5	2.3	Xmas Eve (Sat)	11.4	4.4
Corpus Christi (from 2016)	−77.6	2.3	Xmas Eve (Sun)	42.1	5.2
Corpus Chr. (t+1)	−7.3	1.6	Xmas Day (Sat)	22.0	5.9
Pentecost (t-1)	4.2	1.3	Xmas Day (Sun)	59.2	4.3
Pentecost (to 2015)	−107.2	1.9	Boxing Day (Sat)	18.7	3.2
Pentecost (from 2016)	−119.3	2.3	Boxing Day (Sun)	33.2	5.9
Pentecost (t+1)	−13.4	1.6	10d post Dec 26 (Sat)	31.8	2.0
German Unity (bridge)	−2.4	2.9	10d post Dec 26 (Sun)	14.8	1.9
Labour Day (bridge)	−17.7	2.6			
All Hallows (bridge)	−20.6	2.4			

[1] *NH includes the following holidays: Epiphany, Labour Day, Assumption Day, Day of German Unity, Reformation Day and All Saints' Day. The weights of the regional holidays are given by: Epiphany 0.2, Assumption Day 0.1, Reformation Day from 2018 on 0.2 and All Saints' Day 0.6.*
Note: Ascension and Labour Day 2008 both fell on May, 1. Because the effect is not additive, the effect has been assigned to Labour day only, i.e. the regressor for Ascension is 0 *on that day.*

and the fact that truck drivers are restricted by laws with regards to the number of hours that they are allowed to drive per week and per day. The latter determines the optimal logistics and thus the interaction between holidays, consumer demands and kilometers driven by trucks on a given day resulting in the observed local volatility increases.

Table 1 shows the estimated impact of moving holidays and cross-seasonal effects. All moving holidays that are public holidays on a national or federal state level are included in the regression. This extends to surrounding days, as usually a transition phase can be observed, so that both the day before and after a given public holiday show a decrease. The only exception is the Sunday just before Pentecost Monday, which is slightly positive compared to a typical Sunday in spring. Additionally, we include the main event days of the carnival season for which a significant impact can be detected.

The impact of fixed holidays is usually captured by STL in the last step of DSA. Yet, as discussed above, the interdependency between that impact and the weekday cannot be captured by STL. Here, we model it using interaction dummies in the RegARIMA model. The impact of each weekday is combined across all national and federal state level holidays to increase the parsimony of the estimated model. Days that are public holidays only

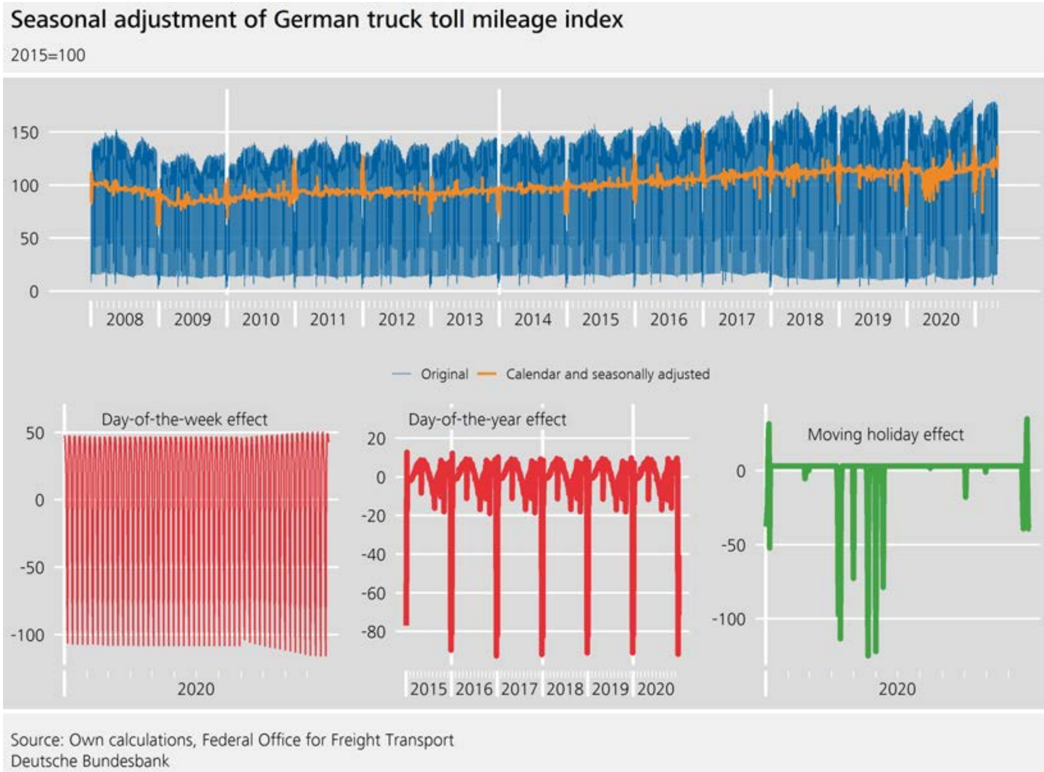

Figure 3: German index of truck toll mileage.

regionally are weighted accordingly (for details on the weights see Table 1 and Deutsche Bundesbank, 2020).

The changes in the weekday pattern after the start of the Covid-19-pandemic in Germany discussed above recur weekly and may thus be considered to be part of the $S^{(7)}$-component. A different view is that as it is only temporary and a result of the irregular nature of the crisis it should be captured in the irregular component and thus be visible in the calendar and seasonally adjusted series.[5] For the period from March 23 to August 30, i.e. from the beginning of the lockdown until after the summer holidays, we chose the latter approach. To seasonally adjust this period, forecasted calendar and seasonal factors were used, which are obtained by restricting the estimation span from time series beginning to March 22, 2020. For the period after August 30, we again use all available data in a controlled current adjustment scheme. This means that the seasonal and calendar components are re-estimated monthly, but we control weekly, whether a re-estimation is necessary.

3.2 Electricity consumption

The German electricity consumption is compiled by the German Federal Network Agency using data from the network providers and published from January, 2015. For the most recent observations, the series is subject to **unreliable data delivery** as some network providers do not provide data immediately.

Remark. *More broadly, unreliable data delivery addresses the issue that often the data providers do have no legal or contractual obligation to provide the data or restrictions*

[5] At the time of writing the ESS Guidelines on Seasonal Adjustment does not discuss this matter and therefore does not stipulate any action (Eurostat, 2015).

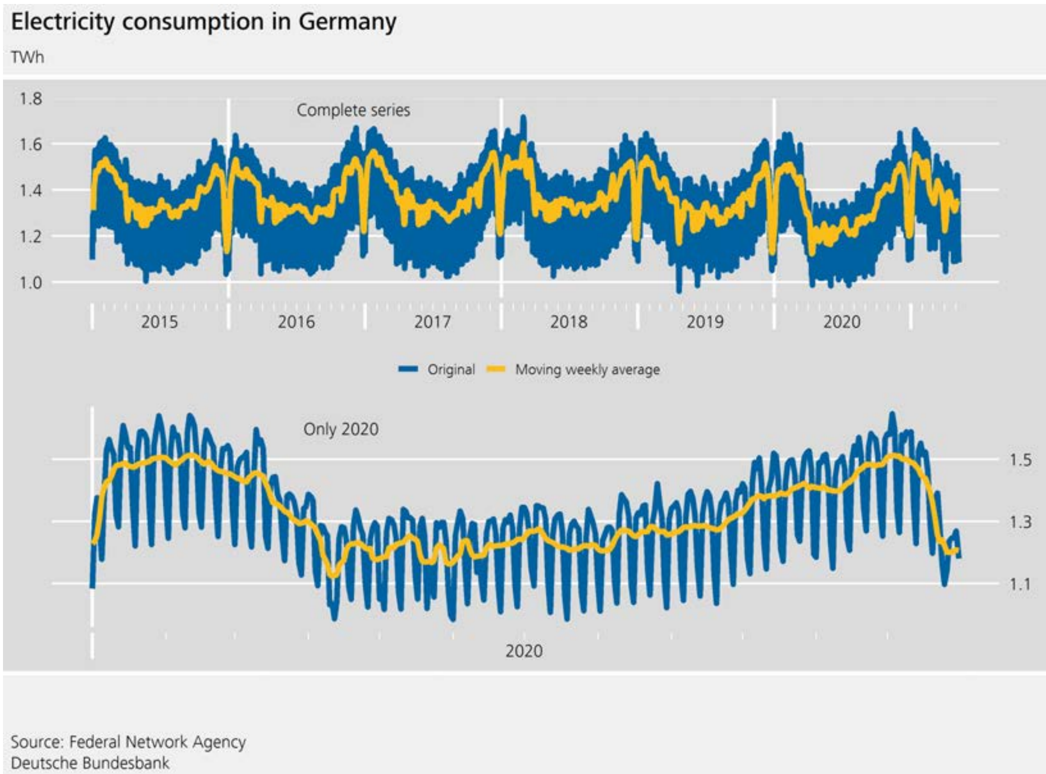

Figure 4: German electricity consumption in TWh.

regarding the data quality or dates at which the data need to be provided. Further, if the data are just a by-product, it may be difficult to analyse the quality of the data.

In the case of the electricity consumption, this leads to temporarily **missing values** that might require an interpolation - either before or as part of the seasonal and calendar adjustment.

Remark. *Some higher frequency time series do have structural missing values, e.g. time series that only contain observations on working days. If this is the case, the data may be* ***non-equidistant****, i.e. the distance in time between observations is not the same for all observations. In other words, the distance between a Monday and a Tuesday is less than between a Friday and a Monday.*

As can be seen from Figure 4, the electricity consumption is characterised by a weekday pattern and - especially if we disregard Christmas and fixed holidays - an almost sinusoidal annual seasonality. The latter can be termed **autocorrelational seasonality**, as the seasonal impact of each day-of-the-year strongly correlates with the seasonal influences observed on the neighbouring days. Such dependence structure is usually not exploited fully in the seasonal adjustment of lower frequency time series. Yet, for higher frequency time series it may improve the estimation of the seasonal effects given the high volatility of the time series and the corresponding estimation uncertainty.

The uncertainty problem is aggravated as the daily electricity consumption is a relatively **short series**.[6] Table 2 shows the estimated moving holiday and cross-seasonal

[6]For very short series, it may not be advisable to estimate all seasonal and periodic components. An estimation of the day-of-the-year effect requires at least two years, but for a reliable and sensible estimation usually three or more years of observations are needed. The day-of-the-week effect may often be estimated using less than two years of data.

effects for this time series. The cross-seasonal effects included are again interactions between weekday and fixed holidays. Due to the length of the series, not all possible calendar interactions can be observed and overall, only very few observations per constellation are available.

At times higher frequency time series show rather **unusual calendar effects**. In the case of the electricity consumption, we have to include Daylight Saving Time, which has an obvious - albeit small - effect on the time series.

Table 2: Estimated moving holidays and cross-seasonal effects for the German electricity consumption, in percent. Based on time series only adjusted for intra-weekly seasonal effects. A RegARIMA(5,1,1) model with 2x24 trigonometric terms has been estimated.

	Estimate	S.E.		Estimate	S.E.
Carnival Monday	−1.7	0.6	Labour Day (bridge)	−5.2	1.5
Holy Thursday	−2.1	0.7	German Unity (bridge)	−7.4	1.1
Good Friday	−22.5	0.9	All Hallows (bridge)	−4.9	1.1
Holy Saturday	−7.1	0.9	Reformation Day (bridge)	−7.4	1.5
Easter Sunday	−6.9	0.9	National Holidays (Mo-Fr)	−19.7	0.5
Easter Monday	−25.6	0.9	National Holidays (Sat)	−5.7	0.9
Easter Mon. (t+1)	−2.8	0.7	3d before Xmas (Sat)	2.1	1.2
Ascension (t–1)	−2.1	0.7	Xmas Eve (Sat)	4.9	1.9
Ascension	−24.2	0.9	Dec 26 (Sat)	4.0	1.2
Ascension (t+1)	−10.3	0.7	10d post Dec 26 (Sat)	5.0	0.9
Corpus Christi (t–1)	−1.3	0.7	3d before Xmas (Sun)	7.3	1.4
Corpus Christi	−15.4	0.9	Xmas Eve (Sun)	9.7	1.7
Corpus Chr. (t+1)	−6.2	0.7	Xmas Day (Sun)	11.8	1.9
Pentecost (t–1)	2.0	0.6	10d post Dec 26 (Sun)	6.6	0.9
Pentecost	−24.6	0.7	Xmas Period (Mo-Fr)	−4.6	0.7
Pentecost (t+1)	−2.9	0.7	Daylight Saving Time Spring	−2.9	0.6
			Daylight Saving Time Autumn	5.1	0.6

To obtain the seasonally adjusted series shown in Figure 5, DSA has been used with $\gamma_7 := 11$ and $\gamma_{365} := 13$ employing a multiplicative decomposition. The short filter for the day-of-the-week is again tailored towards allowing the weekday effect to vary across the year, as we observe that the difference between working days and weekends tends to increase during the warmer season.

The start of the Covid-19-pandemic in Germany again evokes a temporarily declining trend and some slight and gradual changes in the weekday pattern, reflecting the reduction in output in the production sector and possibly changes in the share of people working remotely. For the seasonal adjustment of lower frequency time series, the pandemic is treated using series of level shifts [LS], additive outliers [AO] or less frequently temporary change outliers [TC] (Eurostat, 2020). The downturn observed for the electricity consumption is gradual enough so that one or multiple level shifts are not necessary to model the crisis based on daily data. In general, we more frequently observe **non-traditional outlier pattern** in higher frequency time series than for monthly or quarterly series. This is illustrated well by the finding that the decay rate of 0.7 for TCs, the default for monthly series, will often not be suitable for daily or weekly series.

In the beginning of the pandemic, it may be a good policy not to re-estimate the seasonal and calendar factors immediately, to avoid over-estimation. In the summer, we established a controlled current adjustment scheme discussed above, i.e. a monthly re-estimation while

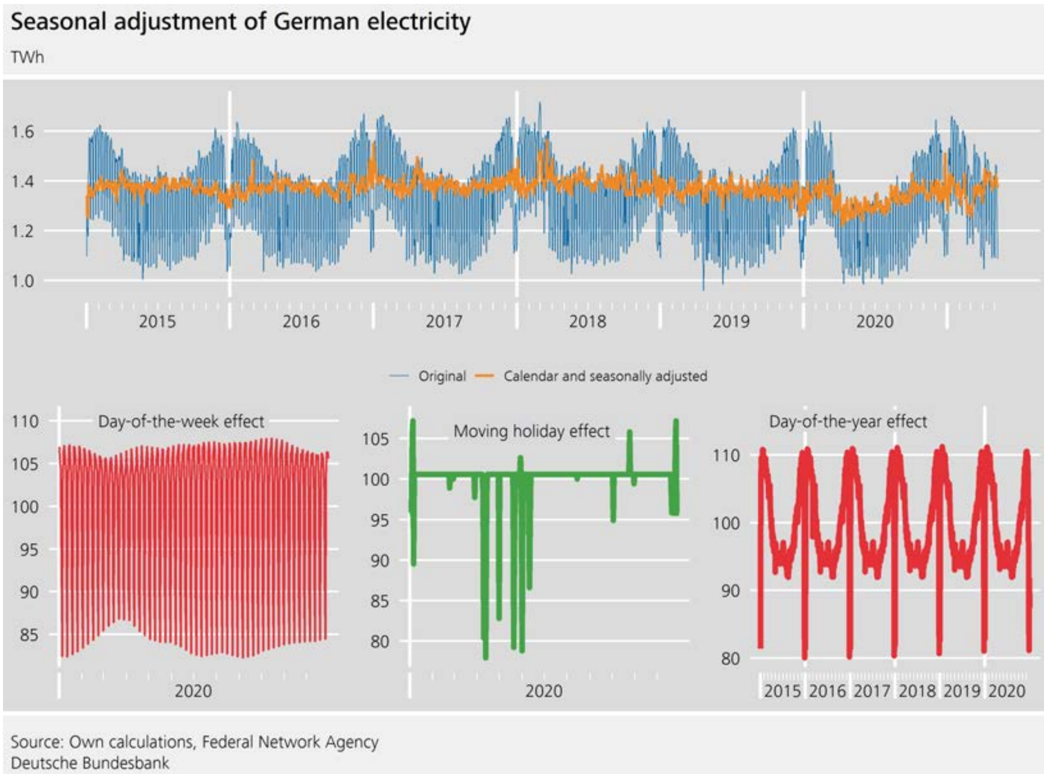

Figure 5: German electricity consumption in TWh.

only changing parameters if necessary and controlling weekly whether the forecasted factors are still suitable or in need of re-estimation.

3.3 Google Trends: Unemployment

Weekly Google Trends data have to be downloaded in chunks of five years. Each of this chunks is a random sample of all search queries and therefore subject to noticeable revisions.[7] The chunks are chain-linked together and a value of 100 is added to avoid values too close to 0.

The definition of the week does not adhere to the ISO 8601 standard, instead it is defined to start on Sunday and end on Saturday. This is of relevance for regressor construction. For example, Easter Sunday and Easter Monday fall in the same week, their effect cannot be disentangled. In turn, we only need one joint regressor for Easter Sunday and Easter Monday.

For lower frequency time series, Christmas and New Year are seasonal effects as they fall in the same period each year. Depending on the modelling strategy, especially if the data are modified so that every year has 365 observations, this holds for daily data too. For weekly time series, as a consequence of their non-isochronicity and date conventions, Christmas' Day can fall in the 51st or 52nd week-of-the-year, while New Year's Day falls in week 52, 53 or 1 (ISO 8601 standard) or week 0 or 1 (US-convention).

The results from the calendar estimation are included in Table 3. As discussed, the construction of the data implies that some holidays always fall in the same week. Their impact is therefore indistinguishable. The inclusion of appropriate regressors for the Christmas and New Year's period can be challenging. Christmas Day always falls in the same week

[7]The average mean revisions are **7.14** points from week to week in the original series and **6.90** points in the calendar and seasonally adjusted data.

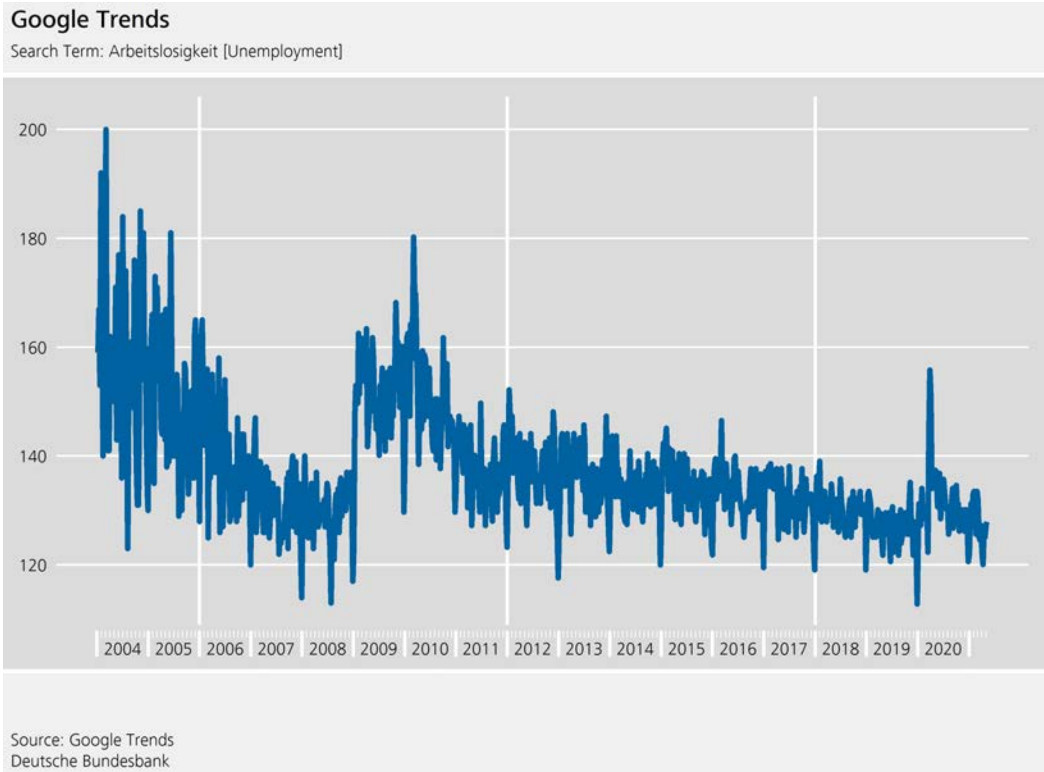

Figure 6: Google Trends, search term: Arbeitslosigkeit [unemployment].

as either Christmas Eve or New Year's Eve. If both Christmas Eve and New Year's Eve are included as regressors, then, particularly for short series, separating the impact of Christmas Day from the other two days may be intricate resulting in an instable estimation.

Table 3: Estimated ARMA coefficients and moving holidays effects (in percent) for Google Trends, search term: Arbeitslosigkeit [unemployment]. An RegARIMA(0,1,1)(0,1,1) model is estimated.

	Estimate	S.E.		Estimate	S.E.
MA	−0.87	0.02	SMA	−0.80	0.03
Good Friday	−6.5	1.1	Christmas Eve	−10.0	1.1
Easter	−4.8	1.1	New Year's Eve	−8.9	1.4

The week is defined to start on Sunday and end on Saturday. Thus, Easter covers both Sunday and Monday.

As can be seen from Figure 6 the series is very volatile, especially at the beginning of the series. It may be difficult to assess the seasonality of the time series visually, but seasonality tests[8] and the coefficients of the ARIMA model indicate seasonality. To obtain seasonal factors we rely on the default settings of the fractional airline decomposition.

As the original data are revised considerably every week, the revision policy of this weekly series differs from the daily series discussed above. For the Google Trends series we follow a partial concurrent adjustment scheme, i.e. the calendar and seasonal compo-

[8]We use the seasonality tests recommended by Ollech and Webel (2020) for monthly time series which are implemented in the {seastests} package in R. Setting the frequency of the time series to 52 ignores the fractional nature of the seasonality of the series. Yet, in practice this works well as an approximation. Both the QS- and the Friedman-test reject the null hypothesis of no seasonality at the 0.1-% level of significance. This holds true as well if we only include the last 5 years of observations.

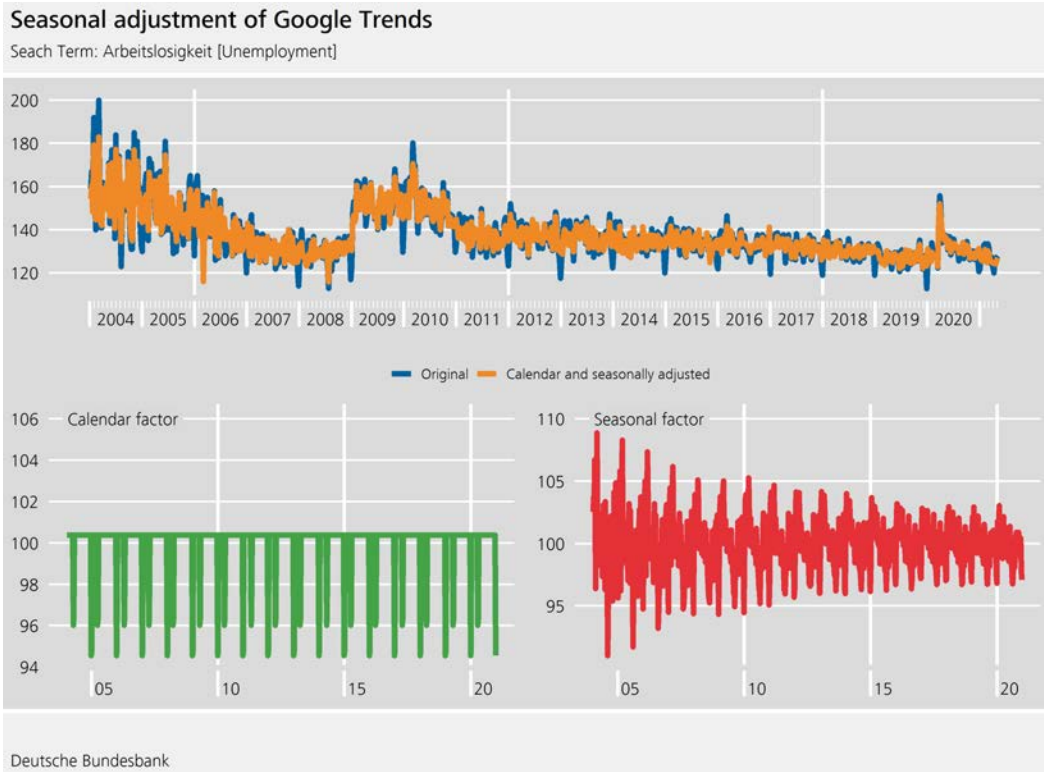

Figure 7: Seasonal and calendar adjustment of Google Trends series, search term: Arbeitslosigkeit [unemployment].

nents are re-estimated every week, but the model settings are fixed and only re-identified annually. The appendix includes a discussion of graphical tools that can be used for weekly time series if a controlled current adjustment scheme is used.

4. Key notions for the adjustment of higher frequency data

The examples discussed in the previous sections highlight features and idiosyncracies of higher frequency time series relevant to calendar and seasonal adjustment. Table 4 organises these characteristics into a taxonomy of relevant features.

The basic characteristics cover the most obvious features that concern the whole set of observations and are relevant for very many time series available. That higher frequency time series contain many observations especially in comparison to lower frequency time series is in itself trivial. Yet, it increases the computational burden and may render some methods inapplicable. Non-isochronicity and non-equidistance will usually either have to be addressed by adapting the methods used or by transforming the data, e.g. by interpolation or time-warping.

The periodic and calendar effects comprises of concepts that extend the definition of seasonality and calendar effects in lower frequency time series. As becomes clear from the examples of daily data, higher frequency time series may contain multiple noticeable periodic effects that may be interdependent and autocorrelational. For calendar effects, we may consider including constellations that are not recommended to be adjusted for in lower frequency time series, such as bridge days. This may be sensible as the higher number of observations allow us to observe and estimate the effect more directly.

Remark. *For weekly series, the impact of bridge days may even be inseparable from the* *moving holidays, if they always fall in the same week.*

Higher frequency time series are generally more volatile than traditional business-cycle indicators. This is mostly a result of the fact that irregular influences tend to partially offset each other over a longer time span. Additionally, many of the higher frequency time series used are merely a by-product and not compiled for economic analysis or official statistics and accordingly do not adhere to the same data standards. Accordingly, the methods applied to higher frequency time series usually need be robust against outliers and irregular observations.

5. Summary

The contribution of this paper is two-fold. First, we exemplarily analyse a small set of higher frequency time series. We discussed how these data differ from lower frequency time series and how this is relevant for seasonal adjustment in general and in light of the Covid-19-pandemic. Second, we developed a taxonomy of the central features of seasonal higher frequency time series.

Further research may evaluate different procedures that allow the seasonal adjustment of higher frequency time series with respect to their ability to handle all of these features.

REFERENCES

Almomani, A., Gupta, B., Atawneh, S., Meulenberg, A. & Almomani, E. A. (2013). Survey of Phishing Email Filtering Techniques. IEEE Communications Surveys & Tutorials 15 (4), 2070-2090.

Cleveland, R. B., Cleveland, W. S., McRae, J. E. & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6 (1), 3-73.

Cleveland, W. S. & S. J. Devlin (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American statistical association, 83 (403), 596-610.

Cleveland, W. S., & I. J. Terpenning (1982). Graphical Methods for Seasonal Adjustment. Journal of the American Statistical Association 77 (377), 52-62.

Deutsche Bundesbank (2020). Seasonal adjustment of the daily truck toll mileage index. Methodology report by the seasonal adjustment section. Available from https://www.bundesbank.de/en/statistics/economic-activity-and-prices/output/methodology-and-quality-795360

Eraslan, S. & Goetz, T. (2020). Weekly activity index for the German economy. Available from https://www.bundesbank.de/wai.

Eurostat (2015). ESS Guidelines on Seasonal Adjustment. 2015 edition. Publications Office of the European Union, Luxembourg.

Eurostat (2020). Guidance on Treatment of Covid-19-crisis Effects on Data. Methodological Note. Available from https://ec.europa.eu/eurostat/cros/system/files/treatment_of_covid19_in_seasonal_adjustment_methodological_note.pdf.

Ladiray, D., Palate, J., Mazzi, G. L., & Proietti, T. (2018). Seasonal Adjustment of Daily and Weekly Data. In: Handbook on Seasonal Adjustment, G. L. Mazzi, D. Ladiray & D. A. Rieser (eds.). Luxembourg: Publications Office of the European Union, 757–783.

Ollech, D. (2018). Seasonal Adjustment of Daily Data. Deutsche Bundesbank, Discussion Paper Series 41/2018.

Ollech, D., Gonschorreck, N. & Hengen, L. (2021). Flexibilisation of X-11 for Higher-Frequency Data. NTTS conference 2021.

Ollech, D. & Webel, K. (2020). A Random Forest-Based Approach to Identifying the Most Informative Seasonality Tests. Discussion Paper Series 55/2020. Deutsche Bundesbank, 1-36.

Svetunkov, I. (2017). Statistical models underlying functions of 'smooth' package for R. Lancaster University Management School.

U.S. Census Bureau (2016). X-13ARIMA-SEATS Reference Manual Version 1.1. Time Series Research Staff, Statistical Research Division, U.S. Census Bureau.

Webel, K. (2020). Challenges and Recent Developments in the Seasonal Adjustment of Daily Time Series. JSM Proceedings, Business and Economic Statistics Section, Alexandria, VA: American Statistical Association, 1971-1997.

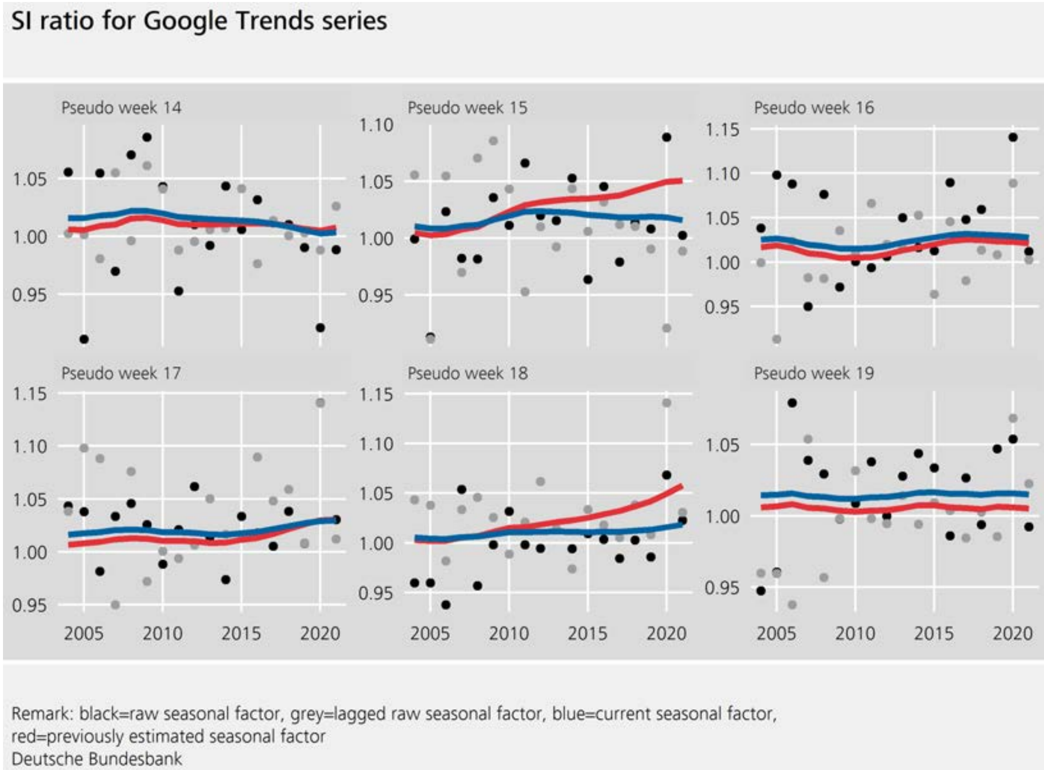

Figure 8: Example of SI ratio for Google Trends series.

A. Appendix

For a controlled current adjustment the results from the pre-adjustment are investigated and most importantly, the previously forecasted seasonal components ($\hat{S}^F$) are compared to the new seasonal factors ($\hat{S}^N$) obtained after re-estimation. As a frame for the comparison the raw seasonal factors ($\hat{S}^{raw}$), i.e. the combined seasonal-irregular component, are included in the analysis. The comparison is either based on tables or graphics, namely the SI ratio. The idea is to gauge whether $\hat{S}^N$ better captures the systematic development of $\hat{S}^{raw}$ compared to $\hat{S}^F$. If this is the case to a relevant extent, the previous components are discarded and the re-estimated results are used.

For integer-type periodic behaviour the construction of the SI ratios are straight forward. Each subplot shows the $\hat{S}^N$, $\hat{S}^{raw}$ and $\hat{S}^F$ for a given period, e.g. a subplot for each month of the year. For series with a non-integer number of observations per year the configuration is more intricate. For weekly series, the estimated seasonal component of a given observations depends both on observations (multiples of) 52 weeks ago as well as 53 weeks ago and likewise on future observations. Therefore $\hat{S}^{raw}$ should include all weeks that contribute to the estimation of the seasonal component. For the most recent observations, this is approximately the sequence of weeks with a step size of 52 and the respective previous week. Exemplarily, this is shown in Figure 8. The exact representation for any period of the seasonal component, the relevant raw seasonal factors are given by the sequence $(..., \hat{S}^{raw}_{t+\lceil 2\tau \rceil}, \hat{S}^{raw}_{t+\lfloor 2\tau \rfloor}, \hat{S}^{raw}_{t+\lceil \tau \rceil}, \hat{S}^{raw}_{t+\lfloor \tau \rfloor}, \hat{S}^{raw}_{t}, \hat{S}^{raw}_{t-\lfloor \tau \rfloor}, \hat{S}^{raw}_{t-\lceil \tau \rceil}, \hat{S}^{raw}_{t+\lfloor 2\tau \rfloor}, \hat{S}^{raw}_{t+\lceil 2\tau \rceil}, ...)$.

Table 4: Key notions for the adjustment of higher frequency data

Notion	Explanation
Basic characteristics	
Many observations	Daily data contain relatively many observations which can be a challenge for algorithms and users.
Short series	For many series only a few years or less are available.
Temporal aggregation	Restrictions regarding temporal aggregation of adjusted series depending on data type.
Non-isochronicity	Number of observations per periodic cycle is not necessarily the same for all cycles, e.g. varying number of weeks per year.
Non-equidistance	Some daily time series do not have observations for all days, e.g. banking-daily, then the distance between observations varies.
Periodic and calendar effects	
Multiple periodic effects	Commonly, daily time series contain day-of-the-week and day-of-the-year effects.
Unusual periodic effects	Other periodic effects, such as a week-of-the-month.
Breaks in periodic effects	Periodic effects may change rapidly, e.g. as a consequence of fundamental changes in the data generating process.
Cross-seasonality	Different periodic and calendar effects can be interdependent.
Autocorrelational seasonality	The seasonal impact of adjacent observations may be highly dependent.
Series-specific calendar effects	Calendar effects can be observed more directly, thus, regressor construction can more easily be tailored to the series.
Unusual calendar effects	Other calendar effects, such as daylight saving time.
Bridge days	The effect of bridge days can have a traceable effect on the series.
Outliers and missing values	
Missing values	Due to data availability, some series contain - at least temporarily - missing values, especially at the end of the series.
Unreliable data delivery	Data producers do not necessarily have an obligation to deliver data or provide additional information.
Higher volatility	The volatility of the time series usually decreases with a higher temporal aggregation.
Heteroskedasticity	The volatility may change over time, e.g. during crisis.
Non-traditional outlier pattern	Observed outlier pattern can be different from lower frequency series, e.g. slower convergence in a temporary change outlier.