Study on the leap year effect

Contents

1 Introduction

In order to keep the calendar year synchronized with the astronomical year, the calendar year can contains one additional day in February. It happens almost every four years[1] and a year that contains 29 days in February is called a leap year (the next leap year is in 2020). The leap year has to be taken into account in the interpretation of some economical indicators. Indeed, one additional working day in February implies one additional day of production and therefore an increase of the production, even if the productivity has remained constant. To compare changes within a year or between months of different years, lots of economical series are seasonal adjusted and correcting for working days, and so for leap year. This is for example the case of the industrial production indexes.

Not all the time series are affected by leap years. For example, when the Labor Force Survey asks respondent for their employment status over a particular week, unemployment estimates are not likely to be affected by leap years. Nevertheless, when time series are likely to be affected by leap years, it may not be possible to correctly estimate it for multiple reasons:

- If the leap year effect is too weak to be estimate. For instance, it can be the case when production is measured in "quantities produced" as for the French production index relate to airplane airframes. Indeed, an additional day in February may not implies an additional number of finished planes during the month and therefore the leap year may not be statistically significant. For the similar reason it may be difficult to estimate the leap year effect in quarterly series.

- If the series span is too short to estimate a leap year effect because it can conduce to misleading results. Indeed, if the series span doesn't contain any leap year it makes no sense to estimate its effect. However, if the series span only contains one leap year, is it enough to correctly estimate its effect? If not, how

[1] More precisely it happens in years which are multiples of four, with the exception of years divisible by 100 but not by 400.

many observations do we need to have a robust estimation? Another question comes arise: does the leap year effect is similar to all data, as it would expect to be?

Is to those three questions that we aim to respond. The data use are monthly and correspond to the turnover indexes in industry, services and in wholesale and retail trail and the industrial production indexes of the European Union members[2]. Series are study at three levels of the NACE Rev 2 (two, three and four digits) and are retains only if more than 12 years of data is available: 2 198 series are examine. For all of them, the leap year effect is economically explainable.

2 How and when carry out the leap year adjustment?

2.1 The leap year adjustment

The leap year adjustment is usually made by two different methods:

- Adding an explanatory variable in a regression model with ARIMA times series errors (regARIMA model). The leap year regressor (L_t) used in X13ARIMA-SEATS is defined as:

$$L_t = \begin{cases} 0.75 & \text{in February during leap years} \\ -0.25 & \text{in February during non leap years} \\ 0 & \text{for all other months} \end{cases}$$

  Therefore, the leap year effect can be defined as the estimate value of the coefficient of the leap year regressor. This value depends in particular on if series are assumed that unobserved components of the time series combine multiplicatively or additively; it can also depends, to a lesser extent, on the other components of the regARIMA model (for example in the case of model failures).

- Correcting values prior to modelling multiplying by a fixed proportion:

$$\begin{cases} \frac{28.25}{29} & \text{in February during leap years} \\ \frac{28.25}{28} & \text{in February during non leap years} \\ 1 & \text{for all other months} \end{cases}$$

  In the case of multiplicative models, this is equivalent to adding a leap year regressor if and only if the coefficient is approximately equal to 0.035 (which is the approximate value expected in the multiplicative model; see appendix A) [1]. This isn't true for additive models.

2.2 When does the leap year adjustment should be done?

According to the ESS Guidelines on seasonal adjustment [2], the leap year adjustment should only be done for time series for which there is an economic explanation and a statistical evidence. This is for example what is done by the Office for National Statistics of the United Kingdom [3].

In this study, there is an economic explanation for all time series. Therefore adjusting from the leap year effect make sense and the statistical evidence will then be tested.

[2] The *sts_intv_m*, *sts_setu_m*, *sts_sepr_m* and *sts_inpr_m*, datasets of Eurostat.

3 From the regARIMA model to the study of the convergence of the leap year estimation

3.1 The regARIMA model

In this study, all time series are working days and leap years adjust (even if there isn't statistical evidence) and the estimation is done in the presence of possibly several types of outlier. In the case of additive model, the following regARIMA model is used:

$$Y_t = \beta_0 L_t + \beta_1 WD_t + \underbrace{\sum_i \gamma_i O_{i,t}}_{\text{possibly absent}} + \varepsilon_t \tag{1}$$

where Y_t is the observed time series, L_t the leap year regressor defined in the section 2.1, WD_t the working day regressor, $O_{i,t}$ possibles outliers and ε_t has mean zero and follows an airline ARIMA model (an ARIMA(0,1,1)(0,1,1)).

In the case of multiplicative combination, a logarithm transformation is applied to the observed series and then modelled additively: the same model than (1) is used but replacing Y_t by $\log(Y_t)$:

$$\log(Y_t) = \beta_0 L_t + \beta_1 WD_t + \sum_i \gamma_i O_{i,t} + \varepsilon_t \tag{2}$$

The leap year effect is thus β_0. In the multiplicative model the value we would expect for β_0 is closed to 0.035. However estimates values of β_0 can deviate from 0.035 due to various model failures such as presence of autocorrelation of heteroskedasticity on regARIMA residuals (since this bias the estimates of the regression parameters). Because all the time series used in this study are indexes of base 100, we could also assume that value we would expect for β_0 in the multiplicative model is closed to 3.5 (this is the case for almost all series study here).

For all time series the methodology described as followed is used to study the estimation of the leap year effect:

- First, we study each series throughout the entire sample in order to determine the decomposition of the times series (multiplicative or additive).

- Then, we identify a regARIMA model (detecting outliers) and estimate it from the beginning of the time series with a four years span fixing the decomposition (determine in the first step)[3]. In the same way we re-identify and re-estimate the model for each additional date to the end of the sample (always from the beginning of the time series).

3.2 Study of the convergence of the leap year estimation

In the study we defined the *time convergence* of the leap year estimation as the elapsed time required for the leap year estimation to remain statistically significant and not statistically different from the last estimation (or from the expected value defined in 3.1). That is to say that the *time convergence* is the the elapsed time required for the decision to perform a leap year adjustment to remain permanent and to the estimation to stabilize.

For instance, taking a times series starting in January 2000 and ending in December 2015, a time convergence equal to 8 years means that if we re-identify and re-estimate the model from January 2000 to any date after January 2008:

[3] Four years are needed to have at least one leap year.

- the leap year is statistically significant;
- the leap year is not statistically different from value estimate from January 2000 to December 2015 (or from the expected value defined in 3.1).

When the leap year estimation isn't statistically significant, we then consider that we don't have convergence and so we don't have any time convergence.

4 Results

4.1 Preliminary study

Almost all series have converged to the expected value (92.7%; table 1), whether in a multiplicative or an additive model. Only 3 series can't be force to multiplicative model (because some data are too closed to zero) and it affects negligently the significativity of the leap year effects and it's convergence to the expected value. So, in order to make further comparison easier, all series will be forced to multiplicative decomposition[4] and we will study the leap year convergence to the value of 0.035 (for which adding a regressor is equivalent to pre-adjusting the initial times series to correct from the leap year effect).

Table 1: Statistics on convergence of the leap year (LY) effect by decomposition

	As converged to the expected value[a]		The LY coefficient is significant	
	Yes	No	Yes	No
Initial decomposition[b]				
Additive	676	70	235	511
Multiplicative	1361	91	429	1023
Forced to multiplicative[c]				
Initially additive	673	70	239	504

Notes: Tests are done with 5% significance level;
[a] 0.035 for multiplicative decomposition and 3.5 for additive decomposition;
[b] Indentified estimating the model throughout the entire sample;
[c] 3 series can't be forced to multiplicative decomposition.

Table 2: Statistics on regARIMA residuals, forcing the multiplicative decomposition

	The LY* coefficient is significant	
	Yes	No
RegARIMA residuals are autocorrelated or heteroskedastics		
As converged to the expected value	285	564
Asn't converged to the expected value	28	43
RegARIMA residuals are correct		
As converged to the expected value	326	859
Asn't converged to the expected value	29	61

Notes: Tests on residuals are done with 1% significance level;
Significativity test done with 5% significance level.
* LY=Leap year

[4] Same results are found without forcing multiplicative decomposition.

Results must be carefully considered when regARIMA residuals are autocorrelated or heteroskedastics. Indeed, in this case the estimated variance of the leap year coefficient is biased (the estimated coefficient can also be biased) and so tests on time convergence are wrong. This is the case for 920 series (table 2). On the 2 195 series examined, only 14.3% have a leap year statistically significant with correct regARIMA residuals. That is to say that, studying the time series throughout the entire sample, we would have doubtlessly proceed to the leap year adjustment for only 14.3% of them!

4.2 Study of the time convergence

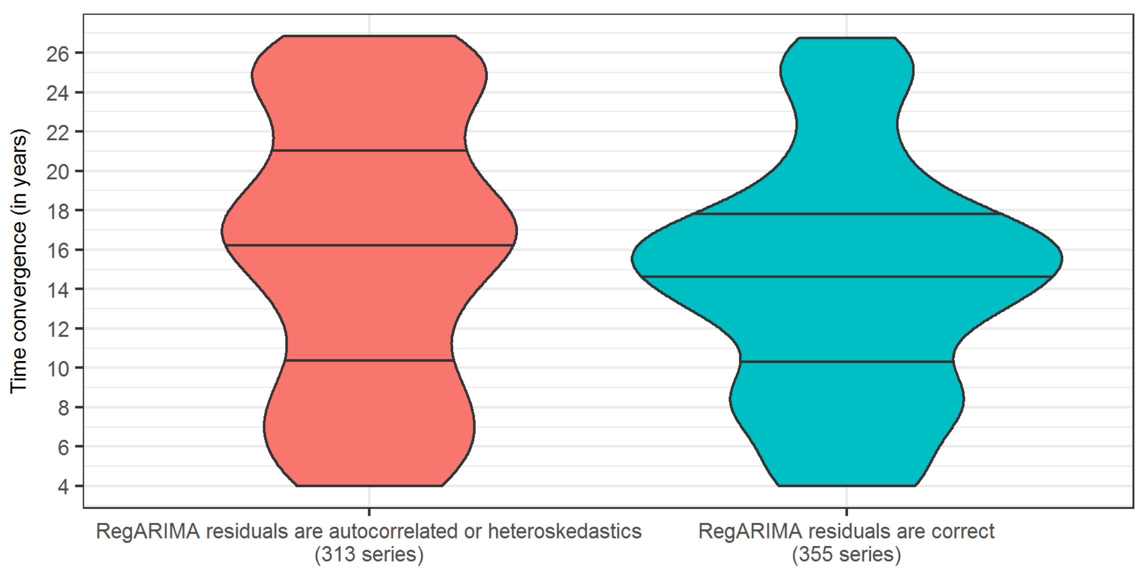
Table 3: Quantiles on time convergence (in years) of the leap year coefficient to 0.035 with a free ARIMA model

	Minimum	65%	70%	80%	90%	Maximum
RegARIMA residuals are correct						
The LY coefficient is significant	4.0	4.0	4.0	5.2	10.2	26.3
The LY coefficient isn't significant	4.0	4.0	4.0	5.2	10.5	26.6
RegARIMA residuals are autocorrelated or heteroskedastics						
The LY coefficient is significant	4.0	4.0	4.0	6.3	13.6	26.5
The LY coefficient isn't significant	4.0	4.0	4.5	8.1	14.5	26.9

Notes: Tests on residuals are done with 1% significance level;
Significativity test done with 5% significance level.
How to read it: For 80% of the times series for which regARIMA residuals are correct and the leap year statistically significant, it takes less than 5.2 years to converge to 0.035.

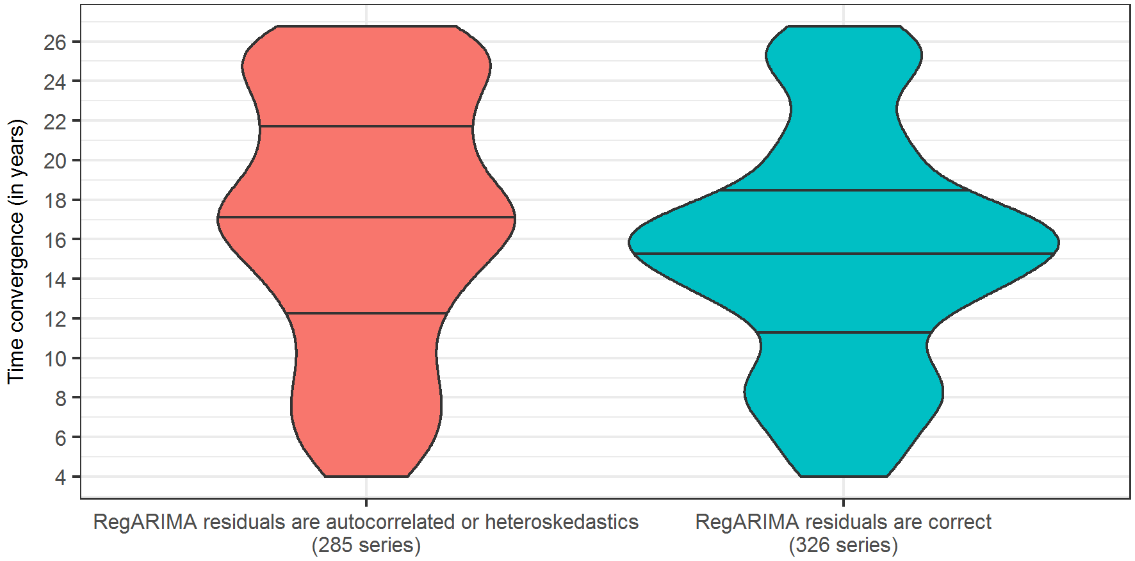
The elapsed time for the leap year estimation to remain not significantly different from 0.035 is very short (table 3): it takes less than 6 years for 80% of the time series. However, the elapsed time for the leap year estimation to remain significant is much longer (figure 1): more than 10 years for 75% of the 624 times series for which the leap year is statistically significant. That implies a long time convergence of the leap year estimation (figure 2).

Figure 1: Time for the leap year coefficient to remain significantly different from zero



How to read it: Violin plots represent a rotated kernel density plot, horizontal lines representing quartiles.

Figure 2: Time convergence of the leap year estimation

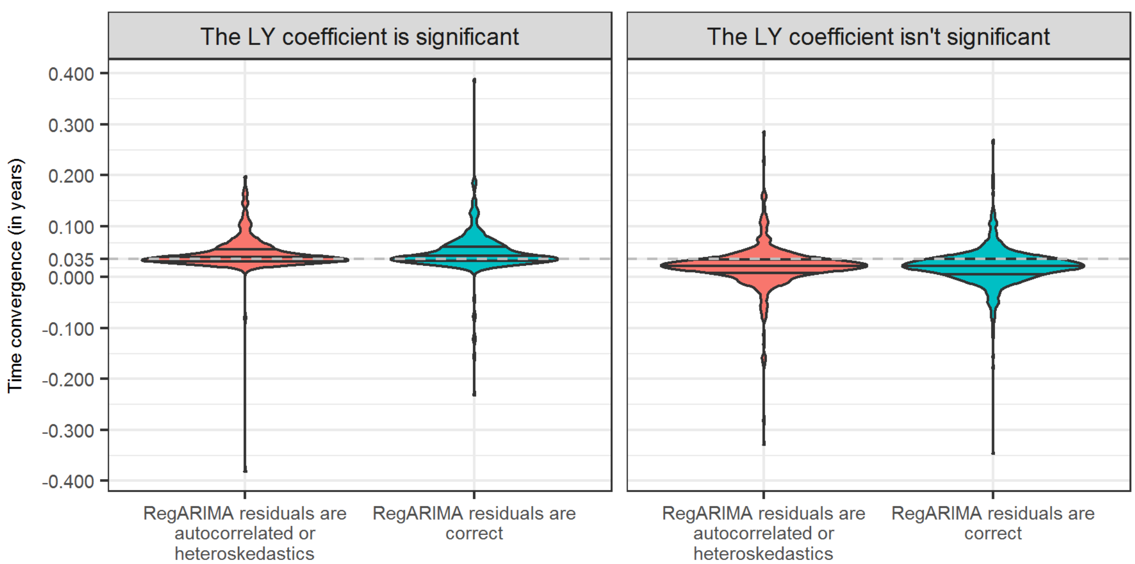


How to read it: Violin plots represent a rotated kernel density plot, horizontal lines representing quartiles.

4.3 Estimate value of the leap year coefficient

Until there, we have study the convergence of the leap year estimation without looking at the estimate value of the coefficient. Even if it isn't significantly different from the value of 0.035 for almost all series, the regARIMA estimation can lead to senseless negative values (graph 3) or questionable high coefficients (above 0.1: it means that the production in February has increased by more than 7.8% just because of the additional day).

Figure 3: Distribution of the estimate values of the leap year coefficient



How to read it: Violin plots represent a rotated kernel density plot, horizontal lines representing quartiles.

5 Changing the ARIMA model

In the regARIMA model (equations (1) and (2)) we assume that errors have mean zero and follow an airline ARIMA model. This hypothesis can be discussed, in particular regarding to the presence of autocorrelation or heteroskedasticity. In this section we study the time convergence of the leap year estimation but with a free ARIMA model: the model is identified by the estimation throughout the entire sample and is use for each estimation (see section 3.2). Although it reduce the number of time series for which regARIMA residuals aren't autocorrelated or heteroskedastics (table 4), it doesn't change the previous results. While the elapsed time for the leap year to remain not significantly different from 0.035 is usually short (table 5), it needs a long period to remain statistically significant (figure 4) and so a long time convergence.

Table 4: Statistics on regARIMA residuals, forcing the multiplicative decomposition with a free ARIMA model

	The LY* coefficient is significant	
	Yes	No
RegARIMA residuals are autocorrelated or heteroskedastics		
As converged to the expected value	224	486
Asn't converged to the expected value	31	35
RegARIMA residuals are correct		
As converged to the expected value	377	936
Asn't converged to the expected value	34	72

Notes: Tests on residuals are done with 1% significance level;
Significativity test done with 5% significance level.
* LY=Leap year

Table 5: Quantiles on time convergence (in years) of the leap year coefficient to 0.035 with a free ARIMA model

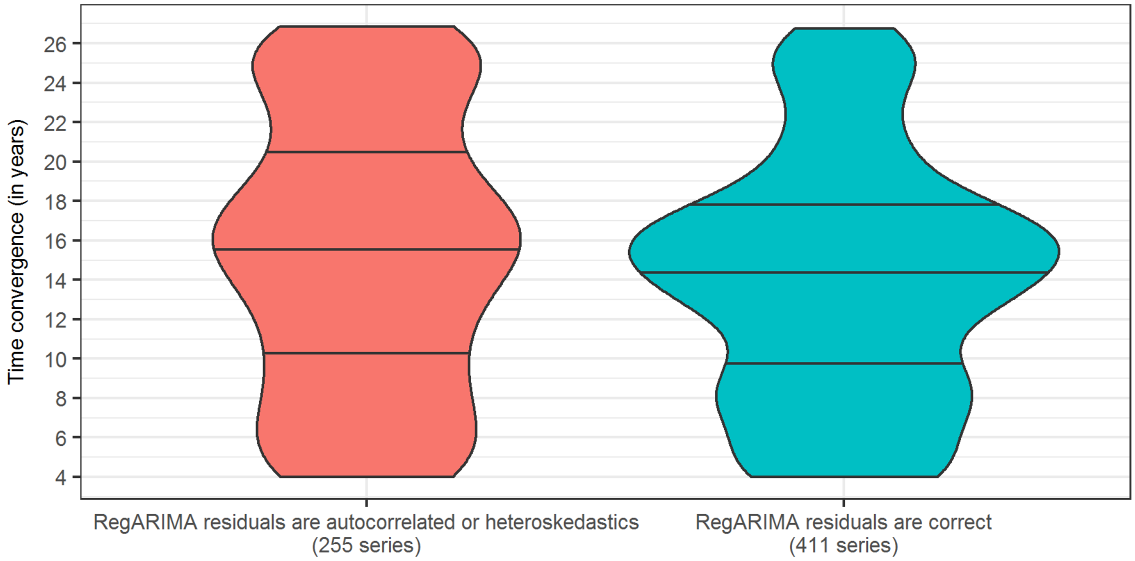
	Minimum	65%	70%	80%	90%	Maximum
RegARIMA residuals are correct						
The LY coefficient is significant	4.0	4.0	4.2	6.1	11.2	26.3
The LY coefficient isn't significant	4.0	4.0	4.1	5.9	11.0	26.4
RegARIMA residuals are autocorrelated or heteroskedastics						
The LY coefficient is significant	4.0	4.8	5.9	8.2	16.1	26.5
The LY coefficient isn't significant	4.0	4.7	5.7	8.8	15.2	26.1

Notes: Tests on residuals are done with 1% significance level;
Significativity test done with 5% significance level.
How to read it: For 80% of the times series for which regARIMA residuals are correct and the leap year statistically significant, it takes less than 6.1 years to converge to 0.035.

6 Conclusions

The leap year adjustment performed including in the regARIMA model must be done carefully, especially with short time series. Indeed, the time convergence of the leap year estimation is quite long (more than 10 years for most series): for short time series the estimation may not be stable. So we may erroneously conclude that the leap year is statistically significant and vice versa.

Figure 4: Time for the leap year coefficient to remain significantly different from zero with a free ARIMA model



How to read it: Violin plots represent a rotated kernel density plot, horizontal lines representing quartiles.

Conversely, it is not recommended to perform an adjustment for time series exceeding twenty years of length since it may produce sub-optimal results [2]. For long time series it's thus not recommended to use a regARIMA model to correct from the leap year effect.

Moreover the regARIMA model can lead to strange estimate leap year coefficient (and so to suspicious leap year adjustment) surely due to model failures.

For all time series where the leap year is economically explainable, correcting values from the leap year effect prior to modelling could be preferred. For multiplicative decomposition it's equivalent to adding a regressor in the regARIMA model and it avoids facing to model failures, suspicious adjustment or convergence problems. For additive decomposition it seems preferable due to the long time convergence which could lead to wrong decisions.

A Alternative approaches to leap year adjustment

Description of Bell's paper.

References

[1] William Robert Bell. Alternative approaches to length of month adjustment. *Statistical Research Division. US Bureau of the Census Statistical Research Division Report Number: Census/SRD/RR*, 92(01), 1992.

[2] Eurostat. Ess guidelines on seasonal adjustment. Technical report, Eurostat Methodologies and Working Papers, European Commission, 2015.

[3] Office for National Statistics. *A methodological note on leap year adjustments*, February 2016.