ESTP: Introduction to seasonal adjustment

Calendar adjustment

Anna Smyk (Insee) - Daniel Ollech (Deutsche Bundesbank)

Objectives of this segment

In this segment, we will see the different types of calendar effects and how to remove them from a series.

At the end, you will know how to :

- identify the different types of calendar effects
- explain the reasons why we remove them
- model a series correction
- read the available diagnostics in JDemetra+

Table of Contents

Calendar heterogeneity

The calendar is heterogeneous :

- trading days :
  - days usually worked, taking into account the company's sector. Most frequently, they are Mondays through Fridays when not bank holidays.
- week-ends
- bank holidays

For a given year as well as throughout the years, every month doesn't have the same number of days per day-type :
$\Longrightarrow$ all months/quarters aren't "equal", even for a given type of month or quarter
This causes **calendar effects**.

The month/quarter length effect

Examples :

- the more trading days in a month, the higher the production is expected to be
- number of trading days in France (non-bank holidays Mondays, Tuesdays,. . ., Fridays) per quarter :

Year	Q1	Q2	Q3	Q4
2010	63	62	65	64
2011	64	62	64	63
2012	65	60	64	64
2013	63	60	65	63
2014	63	60	64	64
2015	63	60	65	64
Average	**63,50**	**60,67**	**64,50**	**63,67**

The “type of day” effect

Examples :

- Sales in retail trade are greater on Saturdays than on any other day
- average number of days of a given type per quarter in France, between 1940 and 2015

Trading day	Q1	Q2	Q3	Q4
Monday	12,57	11,02	12,86	12,72
Tuesday	12,74	12,80	12,88	12,71
Wednesday	12,74	12,82	12,86	12,74
Thursday	12,74	11,84	12,87	12,70
Friday	12,76	12,80	12,84	12,72
Saturday	12,76	12,80	12,84	12,70
Sunday	12,76	12,80	12,86	12,71

Bank holiday	Q1	Q2	Q3	Q4
Monday through Friday	0,91	3,75	1,42	2,14
Saturday	0,14	0,20	0,29	0,43
Sunday	0,13	0,20	0,29	0,42

Bank holidays effects (french calendar)

Type of bank holiday		Examples	Effects
Moving holiday	Depending on Easter	– WhitMonday – Ascension Day (Thursday) – Easter Monday	Type of day
	Easter		Gradual effect
Set date : these bank holidays always occur in the same month but not necessarily on the same type of day		01/01, 01/05, 08/05, 14/07, 15/08, 01/11, 11/11, 25/12	Type of day

Gradual Easter effect

For some series, variations linked to Easter can be seen over a few days prior or following Easter : it's the *gradual Easter effect*.

Example : flowers and chocolate sales might rise significantly as Easter approaches

(in practice very rare, better to deactivate by default detection)

Correction before comparison : same rationale as for seasonal adjustment

Calendar effects must be removed in order to do :

- comparisons across time
  - example : between months of different lengths
- comparisons across different branches in the economy
  - example : between trade and industry
- geographical comparisons (example : across countries)
  - example : France and Germany do not have the same number of trading days each month

Calendar adjustment as expressed in the guidelines

2 types of effects are highlighted :

- length of period (month/quarter)
- composition of period (type of day)

> *"The aim of calendar adjustment is to* ***obtain a seasonally adjusted series*** *whose values are* ***independent of the length and the composition in days*** *(number of Mondays, Tuesdays, etc. or number of working days and weekend days)* ***of the month/quarter."***
>
> *ESS guidelines, item 2.3*

Table of Contents

Calendar effects are partly seasonal

Calendar effects are seasonal :

- the number of trading days is nearly always smaller in February than in March
- some months have more bank holidays (thus fewer trading days) than others
  - example : May in France

Decomposition of calendar effects

Calendar effects = Seasonal effects + Structural effects

↓ ↓

Properties of the calendar that ...

... recur every year	... vary from year to year
(e.g. difference in no. of working days in June and July)	(e.g. weekday composition of months)
↓	↓
Seasonal	**Calendar**

Method : Guidelines recommendations

Alternative A):

"RegARIMA approach, with all pre-tests for number of regressors, length and composition of month, national and religious holiday effects, check of plausibility of effects (sign and size of estimated coefficients), etc.

*The calendar adjustment should be done for those time series for which there is an **economic rationale** for the existence of calendar effects **and statistical evidence**."*

ESS guidelines, item 2.3

The Reg-Arima model

Hypothesis : the effect of a month's length or of a type of day for a given month (or quarter) is constant on the period under review. We will use an econometric model to estimate an average effect via an Reg-Arima algorithm.

A more simple method was historically used before the Reg-Arima approach became ubiquitous. A coefficient was applied to each month (quarter) of the raw series :

$$\text{coefficient} = \frac{\text{average number of days for this type of month over a "long period"}}{\text{number of the month's trading days}}$$

Important : in algorithms presented in this course, seasonality can evolve (via the use of moving averages or filters) but calendar effects CAN'T. This might have consequences when the series is very long : calendar effects may be incorrectly estimated at the end of the series, which usually is the period of interest.

Building trading day regressors : assumptions on the types of days

Construction of a set of “calendar effects” regressors, fitted to each series

To identify “significant” effects and to reduce the number of parameters to estimate with the regression, we must try to establish similarities between types of days.

- Worked Mondays
- Worked Tuesdays
- ...
- Worked Sundays
- Bank holiday Mondays
- Bank holiday Tuesdays
- ...
- Bank holiday Sundays

Bank holiday days are usually treated as Sundays, and the other types of days are grouped or separated, depending on the phenomenon measured by the series at hand.

Building trading day regressors : characteristics

Regressors characteristics :

- seasonally adjusted

in order not to weaken the seasonal signal before decomposing the series

- economically relevant : hypothesis are made on the “type of day” effects

The econometric model (1/4)

We'll use a linear multivariate model. The explanatory variables are the number of days of a given day type : 7 variables.

$$X_t = \sum_{i=1}^{7} \alpha_i N_{it} + \varepsilon_t$$

- N_{it} is the number of Mondays ($i = 1$), ..., Sundays ($i = 7$)
- Example : $N_{3,t=jan2007}$ is the number of Wednesdays in January 2007 (ie. 4 or 5)
- α_i is the effect of a type of day i (coefficient of the variable)

As $N_t = \sum_{i=1}^{7} N_{it}$, the regressors are co-linear. That hampers the estimation of the coefficients α_i.

The econometric model (2/4)

Let's re-write the model, splitting α_i in two, the total effect of a type of day i is the sum of :

- the effect specific to the type i : β_i
- the average effect of a random day $\bar{\alpha}$, with $\bar{\alpha} = \frac{1}{7}\sum_{i=1}^{7} \alpha_i$

We have : $\alpha_i = \beta_i + \bar{\alpha}$

When replacing α_i, the model becomes :

$$X_t = \sum_{i=1}^{7} \beta_i N_{it} + \bar{\alpha} \sum_{i=1}^{7} N_{it} + \varepsilon_t$$

And with $N_t = \sum_{i=1}^{7} N_{it}$ the number of days of the month t :

$$X_t = \sum_{i=1}^{7} \beta_i N_{it} + \bar{\alpha} N_t + \varepsilon_t$$

The econometric model (3/4)

By construction, $\sum_{i=1}^{7} \beta_i = 0$, meaning that the day-specific effect is null over a week, just as seasonality is null over a year.

This allows to write one β_i as a function of the other coefficients and to get rid of the co-linearity issue.

β_7 usually is the Sunday coefficient. We re-write it as :
$\beta_7 = -\sum_{i=1}^{6} \beta_i$

After replacing the new expression into the equation, the model becomes :

$$X_t = \sum_{i=1}^{6} \beta_i N_{it} - \sum_{i=1}^{6} \beta_i N_{7t} + \bar{\alpha} N_t + \varepsilon_t$$

The econometric model (4/4)

We seasonally adjust $\bar{\alpha} N_t$ by subtracting its long term average : it becomes$\bar{\alpha}(N_t - \bar{N}_t)$

The final model is :

$$X_t = \sum_{i=1}^{6} \beta_i (N_{it} - N_{7t}) + \bar{\alpha}(N_t - \bar{N}_t) + \varepsilon_t$$

Two terms appear :

- the specific effect of a type of day i as a contrast between the number of days i and the number of Sundays and bank holidays
- the effect of the month's (or period's) length.

Once seasonally adjusted, this term comes down to the leap year effect :

- for all months except Februaries $\bar{N}_t = N_t$
- for Februaries $\bar{N}_t = 28.25$ and $N_t = 28$ or $N_t = 29$

Building trading day regressors in practice

JDemetra+ offers two sets of built-in regressors :

- Trading Days (6 regressors)
  - All days from Monday to Saturday are considered different
- Working Days (1 regressor)
  - The difference is only made between week days (Monday = Tuesday = ... = Friday) and week-ends
- Leap year(1 regressor)

Warning :

- your national bank holidays are not taken into account by default in JDemetra+
- can be fixed that by creating a new calendar in the GUI (use JD+ calendar customizing features) (cf. GUI and Practice session)

Building your own regressors and using the “Calendar user-defined variables” : more flexible

Examples of regressors

Grouping together different types of days leads to several sets of regressors (explanatory variables). The user or the algorithm will select the one that best corrects each series.

Here all sets used at Insee. The variables are formulated in "contrasts" which corresponds to the final expression of the model above.

Sets of regressors	Hypotheses	Reference (contrast)	Number of regressors
REG1	(Worked Monday = ...= Worked Friday) and (Saturday = Sunday = Bank holidays)	Saturday, Sunday and Bank holidays	1 + LPY
REG2	(Worked Monday = ...= Worked Friday), (Saturday) and (Sunday = Bank holidays)	Sunday and Bank holidays	2 + LPY
REG3	(Worked Monday), (Worked Tuesday = ...= Worked Friday) and (Saturday = Sunday = Bank holidays)	Saturday, dimanches et Bank holidays	3 + LPY
REG5	(Worked Monday), ..., (Worked Friday) and (Saturday = Sunday = Bank holidays)	Saturday, Sunday and Bank holidays	5 + LPY
REG6	(Worked Monday), ..., (Worked Friday), (Saturday) and (Sunday = Bank holidays)	Sunday and Bank holidays	6 + LPY
Leap Year (LPY)	Every day has the same effect	All days	1

A few questions to answer

Before choosing a set of regressors, one must answer the following questions :

- can the series present calendar effects (economic rationale) ?
- which days do we expect to have an impact on the series values ?
  - Trading days from Monday to Saturday ?
  - Trading days from Monday to Friday ?
  - Can all trading days have the same impact on the series values ?

Two possible tests

In the diagnostics panel ("Pre-processing"), JD+ offers significance tests to evaluate the used set of regressors :

- a Fisher test of joint nullity of the coefficients (H_0 all coefficients are null and H_1 at least one of the coefficients isn't null)
  - can lead to removing the set of regressors from the Reg-ARIMA model
- for each regressor, a Student test of coefficient nullity (H_0 the coefficient is null and H_1 the coefficient isn't null)
  - can lead to changing the set of regressors

Warning : regressors must be used in whole coherent sets, if one variable is not significant do not remove it alone, change sets or not.

Is there a gradual Easter effect ?

Does the series present a gradual Easter effect ?

- If so, an extra regressor is added to remove the gradual effect
  - Variable "Easter" in JDemetra+
- If not, no variable is added
- JDemetra+ can "pre-test" the presence of a gradual Easter effect. If the test is positive, JD+ adds the regressor "Easter [n]" (n=1,. . ., 20) to the model.

In the diagnostics panel ("Pre-processing"), JDemetra+ offers a significance test of the coefficient regressor "Easter [n]" (Student type t-test : H_0 coeff(Easter [n]) = 0 et H_1 coeff(Easter [n]) $\neq$ 0)

Table of Contents

Take home message

- It is essential to remove calendar effects before making temporal and spatial comparisons (identical philosophy to seasonal adjustment, with the estimated effects being fixed for the whole estimation period)
- There are two types of calendar effects : period length (LY) and type of day, to which in some rare cases a gradual Easter effect can be added
- We only remove non-seasonal calendar effects in order to preserve the raw series seasonality for the decomposition step.
- Custom sets of regressors can be calculated after making assumptions on the similarity of day types
- JDemetra+ offers functionalities to build regressors (2 types) with customized national calendars
-
  - In a later sequence we will see how to build customized regressors with {rjd3modelling} R package
- JDemetra+ tests for the presence of residual calendar effects (called residual trading days effects in the software)