Trading-Day Adjustment as a Practical Problem

Attal-Toubert, Ketty
INSEE, Département des statistiques de court terme
18, Boulevard Adolphe Pinard
75675 PARIS CEDEX 14, FRANCE
E-mail: ketty.attal-toubert@insee.fr

Ladiray, Dominique
INSEE, Département des statistiques de court terme
18, Boulevard Adolphe Pinard
75675 PARIS CEDEX 14, FRANCE
E-mail: dominique.ladiray@insee.fr

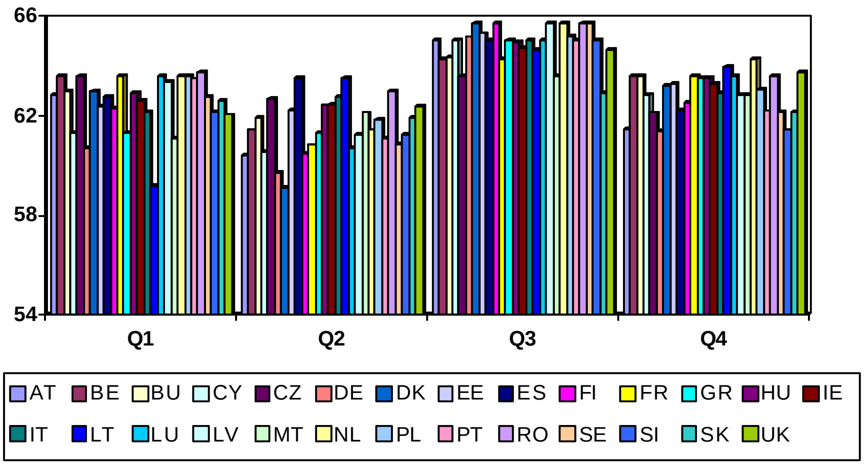
1 Introduction

Trading day effects are present in many economic time series. They occur when the intensity of activity is not the same according to the day of the week. In that case monthly or quarterly time series can contain fluctuations linked to the calendar structure. One more Saturday in a month for example drastically impacts the retail trade turnover in European countries. In the widely used seasonal adjustement softwares Tramo-Seats and X12-ARIMA, trading-day adjustement is done using Reg-ARIMA modeling. For the regression part of the model, the user can specify pre-defined regressors and/or user-defined regressors. In this article, we propose a way to build a user-defined set of trading-day regressors taking into account specificities of the National calendar and the sector of activity under review.

1.1 Why should we take into account National calendars in trading-day adjustment?

Figure 1 represents the average quarterly number of working days across european countries. Working days are here defined as the number of Mondays, Tuesdays, Wednesdays, Thursdays, Fridays which are not public holidays. We can notice big differences between countries, in particular in first and second quarter. Therefore there is an important potential impact on the economic time series. That's why Eurostat asks European countries to do their own trading-day correction before performing a direct seasonal adjustement.

Figure 1: Average quarterly number of working days across European countries



Even for a specific country, the number of working days for a given quarter can vary a lot from a year to another. In France, the second quarter of 2016 will have three working days more than the same quarter in 2015 (see Table 1). This may lead to fluctuations in time series that are not corrected by seasonal adjustment.

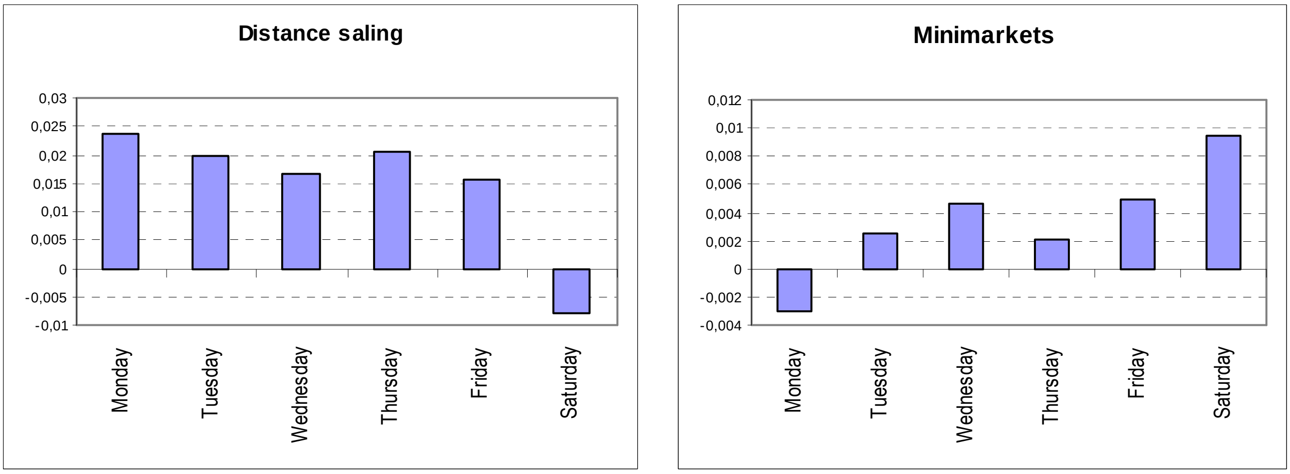
Table 1: Number of working days in France

	Q1	Q2	Q3	Q4
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
2016	63	63	64	63

1.2 Why should we take into account the sector of activity under review?

Trading day effects are very different depending on the sector of activity. The following graphs (Figure 2) show the coefficients of trading-day regressors for French turnover indexes in two retail trade sectors: Minimarkets and Distance saling.

Figure 2: Coefficients of trading-day regressors for Minimarkets and Distance saling turnover indexes



For Minimarkets, Monday's coefficient is negative, because in France many shops are closed on Monday. On the opposite Saturday's coefficient is very high as many people do shopping on Saturday. Other coefficients are positive, but smaller. For Distance saling, coefficients from Monday to Friday are quite similar, when Saturday's coefficient is negative.

Predefined regressors in Tramo-Seats and X12 ARIMA consider only two cases:

- the first one is that each day of the week has a distinct effect, then six trading-day regressors (Monday, …, Saturday), are introduced in the Reg-ARIMA model,
- the second case is founded on the hypothesis that weekdays, from Monday to Friday, are quite similar, whereas Saturday is similar to Sunday, which leads to one contrast regressor.

Distance saling corresponds to the second case. Minimarkets does not correspond to any of the two cases as the middle week coefficients are quite similar.

That's why it could be useful to propose trading-day regressors which permit to consider more various situations than those two cases.

2 The classic models

2.1 The basic model

Trading-day effects are linked to the composition in days of the month (quarter).
The basic model for fixed trading-day effects of flow time series is therefore:

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

Where X_t is the raw series and N_{it} denotes the number of Mondays (i=1), Tuesdays (i=2) … at date t and ε_t follows an ARIMA model:

$$(1-B)^d (1-B^{12})^D \varphi(B)\phi(B^{12})\varepsilon_t = \vartheta(B)\Theta(B^{12})a_t$$ for a monthly time series.

d and D are the differencing orders, $\varphi, \Phi, \vartheta, \Theta$ are polynomials, a_t is a white noise.
α_i is the average effect of an i-type day on variable X.

The model presents at least two problems (see Bell, Hillmer (1983) and Bell (1984 and 1995)):

- α_i's estimates tend to be highly correlated
- N_{it}'s regressors are seasonal. For "pure" trading-day effect estimation, regressors should be seasonally adjusted.

Let us note $\overline{\alpha} = \frac{1}{7}\sum_{i=1}^{7}\alpha_i$ and $\beta_i = \alpha_i - \overline{\alpha}$ for i=1, …, 7.

Model (1) can be rewritten as $X_t = \sum_{i=1}^{7}\beta_i N_{it} + \overline{\alpha}\sum_{i=1}^{7} N_{it} + \varepsilon_t$ (2), where:

$\overline{\alpha}$ is the average effect of any day;

β_i is the specific effect of a i-type day and $\sum_{i=1}^{7}\beta_i = 0$;

and $\sum_{i=1}^{7} N_{it} = N_t$ is the length of month t.

Using the constraint $\sum_{i=1}^{7}\beta_i = 0$, we can write: $\beta_7 = -\sum_{i=1}^{6}\beta_i$

Replacing β_7 in model (2) we obtain:

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

Contrast regressors $(N_{it} - N_{7t})$ are not seasonal because for each i-type day, the seasonal part of N_{it}, which is its long-term monthly mean, is almost equal to N_{7t}'s one. As N_t is seasonal, it is replaced by it's deviation to its long-term monthly mean (Leap Year variable). Then the model becomes:

$$X_t = \sum_{i=1}^{6}\beta_i (N_{it} - N_{7t}) + \overline{\alpha} LY_t + \varepsilon_t \qquad (3)$$

With : $$LY_t = \begin{cases} 0.7575 & \text{if February and Leap Year} \\ -0.2425 & \text{if February and not Leap Year} \\ 0 & \text{Otherwise} \end{cases}$$

2.2 A more parsimonious model

In model (3), six trading-day regressors coefficients have to be estimated, as well as ARIMA coefficients and possibly outliers or other types of regressors. To get a more robust estimation of the coefficients, it would be better to find a more parsimonious model.

Tramo and X12-ARIMA propose a simplified form of the model, introducing a "weekday" regressor. The underlying hypothesis is that in some economic activities, weekdays (Monday, Tuesday,…, Friday) are similar and in the other hand that Saturday is similar to Sunday. You then have:

$\beta_6 = \beta_7$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$

$$\sum_{i=1}^{i=7} \beta_i = 0 \Rightarrow \beta_6 = \beta_7 = -\frac{5}{2}\beta_1$$

and the final model is: $X_t = \beta_1 \left[\sum_{i=1}^{i=5} N_{it} - \frac{5}{2}\left(N_{6t} + N_{7t}\right) \right] + \overline{\alpha} LY_t + \varepsilon_t$

3 General modeling

3.1 Introduction of National calendars

Now we propose a more general form of the model in order to introduce National calendar. Intuitively, the effect of a Monday on the activity is not the same if this Monday is ordinary or if it is a public Holiday. Then, instead of a partition in seven days, we write a partition in fourteen types of days.

$$X_t = \sum_{i=1}^{14} \beta_i N_{it} + \overline{\alpha} LY_t + \varepsilon_t \qquad (4)$$

- *i*=1 to 7 indexes denote Mondays "in" , Tuesdays "in",…, Fridays "in" and *i*=8 to 14, Mondays "off", Tuesdays "off",…, Fridays "off".
- $\overline{\alpha} = \frac{1}{14}\sum_{i=1}^{14} \alpha_i$ and $\beta_i = \alpha_i - \overline{\alpha}$ for i=1, …, 14.

As in the classic model, the number of regressors can be reduced under some hypotheses.

For example, if we assume that a Sunday "off" is equivalent to a Sunday "in" and that any day "off" is equivalent to a Sunday, then we have: $\beta_7 = \beta_8 = \beta_9 = ... = \beta_{13} = \beta_{14}$ and the model becomes:

$$X_t = \sum_{i=1}^{6} \beta_i \left(N_{it} - \frac{1}{8}\sum_{i=7}^{14} N_{it} \right) + \overline{\alpha} LY_t + \varepsilon_t$$

Note that, here, using a contrast does not remove the seasonal component anymore as "Sundays + days off" do not have the same long-term monthly means as the other days. We need to remove seasonal part of those regressors by replacing them by their deviation to their long-term monthly mean.

3.2 The very complete model

It is possible to generalize the preceding model by writing a partition of the month in p types of days:

$$X_t = \sum_{i=1}^{p} \alpha_i N_{it} + \varepsilon_t \quad (5)$$

Let us note $\overline{\alpha} = \frac{1}{p}\sum_{i=1}^{p} \alpha_i$ and $\beta_i = \alpha_i - \overline{\alpha}$ for *i*=1, …, *p*.

As before, the model is changed into: $X_t = \sum_{i=1}^{p} \beta_i N_{it} + \overline{\alpha} LY_t + \varepsilon_t$ and as we have $\beta_j = -\sum_{\substack{i=1, \\ i \neq j}}^{p} \beta_i$, the model can be written with contrast regressors, in order to solve collinearity problems:

$$X_t = \sum_{\substack{i=1, \\ i \neq j}}^{p} \beta_i (N_{it} - N_{jt}) + \overline{\alpha} N_t + \varepsilon_t$$

We then try to formulate hypotheses in order to simplify the model. As it is preferable to have contrast regressors, we define two types of hypotheses:

- equality of coefficients of contrast variables
- equality of coefficients of other day-variables

The contrast variable is a combination of different days that are supposed to have the same behaviour (and then the same coefficient):

$$C = \{i_1, i_2, ..., i_c\} \subset \{1,2,..., p\} \text{ with } \beta_{i_1} = \beta_{i_2} = ... = \beta_{i_c}$$

The number of elements of set C is noted c.

Let $E_1, E_2, \cdots, E_m$ be *m* sets of «equivalent day-variables»:

$$E_1 = \{j_1, j_2, ..., j_{e_1}\} \subset \{1,2,..., p\} \iff \beta_{j_1} = \beta_{j_2} = \cdots = \beta_{j_{e_1}} \text{ and } \# E_1 = e_1$$

$$E_2 = \{k_1, k_2, ..., k_{e_2}\} \subset \{1,2,..., p\} \iff \beta_{k_1} = \beta_{k_2} = \cdots = \beta_{k_{e_2}} \text{ and } \# E_2 = e_2$$

$$\vdots$$

$$E_m = \{l_1, l_2, ..., l_{e_m}\} \subset \{1,2,..., p\} \iff \beta_{l_1} = \beta_{l_2} = \cdots = \beta_{l_{e_m}} \text{ and } \# E_m = e_m$$

Note that C and $E_1, E_2, \cdots, E_m$ make a partition of the days of the months. We use the same method as before to transform the model:

$$\sum_{i=1}^{p} \beta_i = 0$$

$$\Rightarrow \sum_{i \in E_1} \beta_i + ... + \sum_{i \in E_m} \beta_i + \sum_{i \in C} \beta_i + \sum_{\substack{i=1 \\ i \notin \{C \cup E_1 \cup ... \cup E_m\}}}^{p} \beta_i = 0$$

$$\Rightarrow e_1 \beta_{j_1} + ... + e_m \beta_{k_1} + c \beta_{i_1} + \sum_{\substack{i=1 \\ i \notin \{C \cup E_1 \cup ... \cup E_m\}}}^{p} \beta_i = 0$$

And $\beta_{i_1} = ... = \beta_{i_c} = -\frac{1}{c} \left(e_1 \beta_{j_1} + ... + e_m \beta_{k_1} + \sum_{\substack{i=1 \\ i \notin \{C \cup E_1 \cup ... \cup E_m\}}}^{p} \beta_i \right)$

The final form of the very complete model is therefore:

$$X_t = \beta_{j_1} \left(\sum_{i \in E_1} N_{it} - \frac{e_1}{c} \sum_{i \in C} N_{it} \right) + ... + \beta_{k_1} \left(\sum_{i \in E_m} N_{it} - \frac{e_m}{c} \sum_{i \in C} N_{it} \right) + \sum_{\substack{i=1 \\ i \notin \{C \cup E_1 \cup ... \cup E_m\}}}^{p} \beta_i \left(N_{it} - \frac{1}{c} \sum_{i \in C} N_{it} \right) + \overline{\alpha} N_t + \varepsilon_t \qquad (6)$$

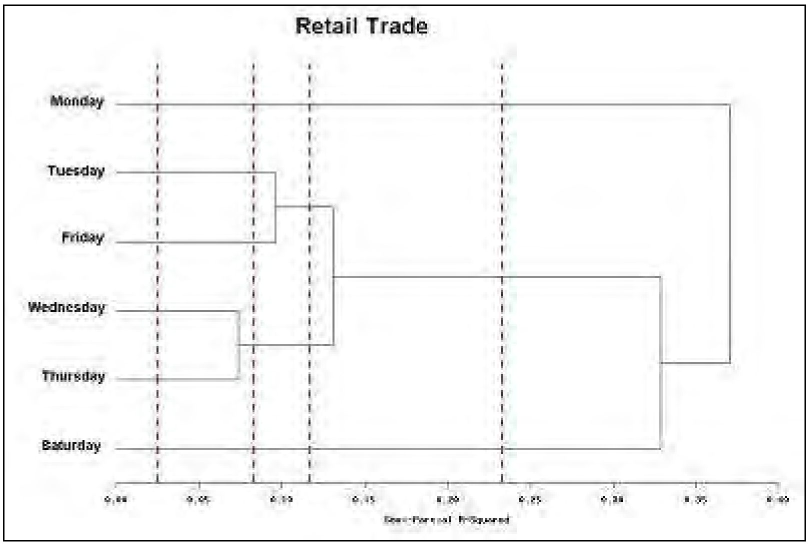
Finally, one must not forget to remove seasonality from those regressors.

4 How to define sets of regressors relevant with sector of activity?

We chose to apply the 15-variable model (see section 3.1 model 4) to the French turnover indexes at NACE 5-digit level, with X12-ARIMA. Among the high number of possible hypotheses to obtain a more parsimonious model, how to choose the more relevant ones, according to the sector of activity?

First of all, we arbitrarily did the hypothesis that Sunday "in", Sunday "off" and all day "off" are similar, which leads to six trading-day regressors in contrasts to Sundays+Days "off". Those regressors were introduced as user-defined variables in Demetra+ and the coefficients were computed with X12-ARIMA. Then the idea was to explore the coefficients of those trading day-regressors using an agglomerating hierarchical cluster analysis. The following graph (in black lines) shows the cluster tree in retail trade sector (57 series) (Figure3):

Figure 3: Cluster tree of trading-day's coefficients of French turnover indexes in retail trade



Cutting the branches at different levels (see vertical dotted lines) leads to retain a selection of relevant sets of regressors. In retail trade, using the 4th red line from the left permits to define a set of three regressors: Monday alone, Tuesday until Friday together, and Saturday alone. It corresponds to the case of Minimarkets, which is frequent in French retail trade. Figure 4 shows the cluster tree corresponding to the industry turnover indexes.

After this exploration, we finally decided to retain five sets of trading-day regressors :

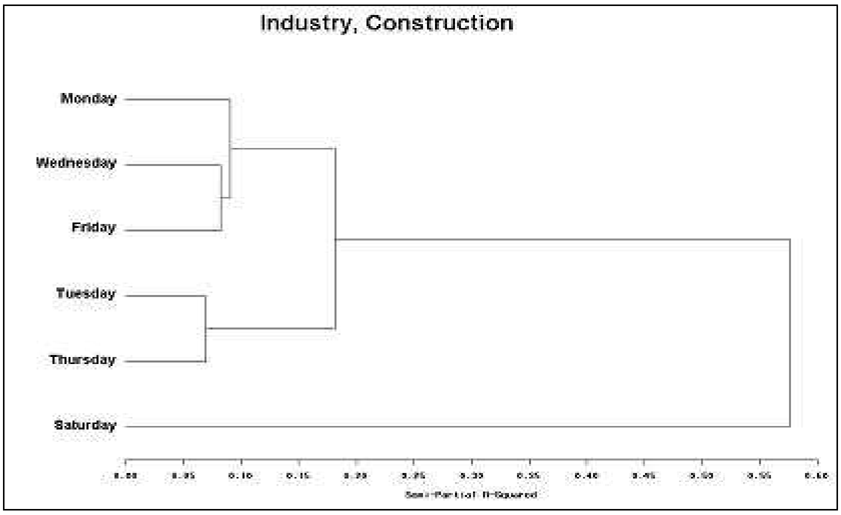
- **S0**: Leap Year alone (hypothesis that all days of the week are similar)
- **S1**: 6 variables: Mon. in, Tue. in, Wed. in, Thu. in, Fri. in, Sat. in; Contrast: Sun. + Days off
- **S2**: 1 variable: (Mon. + Tue. + Wed. + Thu. + Fri.) in; Contrast: Sat.+Sun. +Days off
- **S3**: 5 variables: Mon. in, Tue. in, Wed. in, Thu. in, Fri. in; Contrast : Sat.+Sun. + Days off
- **S4**: 3 variables: Mon. in, (Tue. + Wed. + Thu. + Fri.) in, Sat. in; Contrast : Sun. + Days off

Leap Year is also present in each model with S1, S2, S3, S4.

S0 corresponds to seasonally adjusted length of month effect. We wanted to check if it could be relevant for some activities that have continuous production process like certain chemical industries. S4 was retained for retail trade, wholesale trade and services but not for industry and construction.

The next step is to define a method to select for each serie the best set among those five.

Figure 4: Cluster tree of trading-day's coefficients of French turnover indexes in industry and construction



5 An algorithm for selection of the best set of regressors

5.1 Description of the algorithm

For each series, a Reg-ARIMA model was computed with X12-ARIMA using the more complete set of regressors (S1) as « user-defined » regressors. Automatically identified ARIMA model and outliers (with critical value=5) were then used to compute trading day estimation with each other set of regressors. Then, the selection between the different sets was done according to an algorithm which combines two criteria:

- **Fisher test** of constraints on coefficients: if the hypothesis of equality of several coefficients is not rejected, it is possible to put the corresponding days in the same group (for example Monday, Thuesday,…, Friday together).
- **AICC criterion** gives information about the quality of the Reg-ARIMA model with each set.

Figure 5 describes the complete algorithm of selection, for each series of industry and construction, of the best set of regressors among S0, S1, S2, S3.

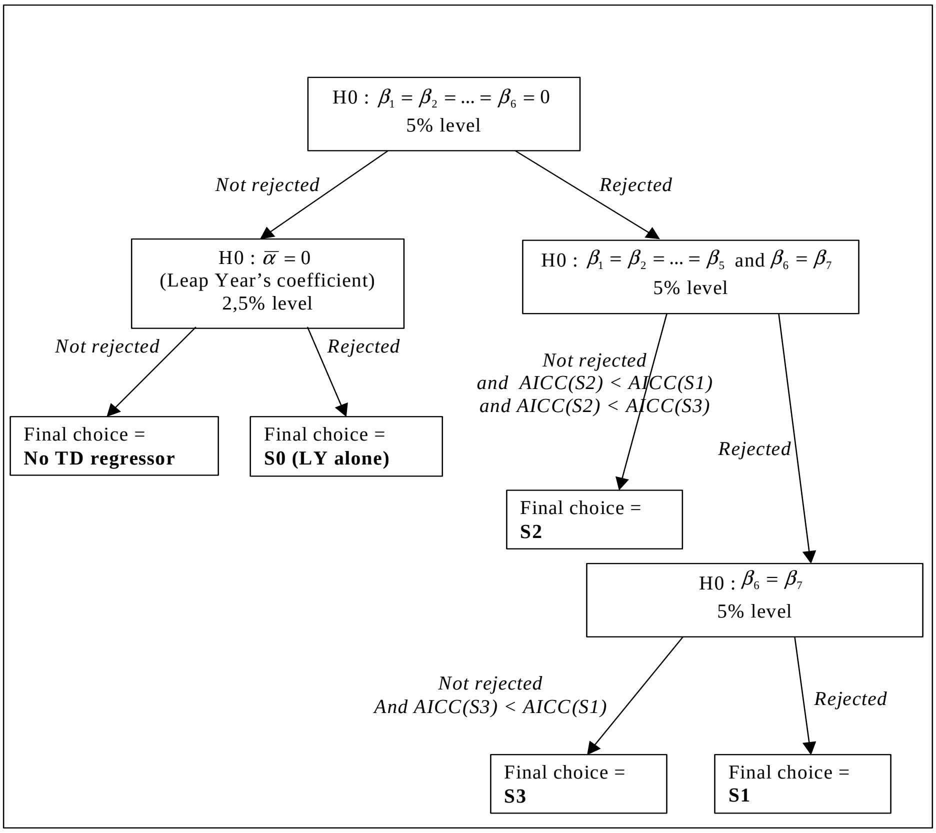
5.2 Results

Algorithms of selection were applied for each turnover indexes: the one of Figure 5 for industry-construction, and a similar one, but including S4 set, for other sectors. We give in Table 2 the results of this selection process.

Table 2: Distribution (in %) of the trading-day sets selected in each sector of activity

	No TD effect	S0 (LY alone)	S1	S2	S3	S4
Retail Trade (57)	5	2	21	4	7	61
Wholesale Trade (89)	17	4	8	27	16	28
Services (138)	26	1	7	32	14	19
Industry (122)	15	0	14	47	25	Not tested

Figure 5: Algorithm of selection, for each industry and construction series, of the best set of regressors



We can notice that in retail trade, for a large majority of series, S4 set was selected as the best one, which confirms clustering results. In wholesale trade, S4 is also the more frequently selected. In services and in industry-construction, S2 is more often selected. Except for retail trade, the percentage of series without trading-day effects is great. Almost no series selected Leap Year alone (S0).

Finally, we can ask if trading-day adjusment with those new sets of regressors is good. The diagnostic chosen for quality check is the residual trading-day spectral peaks diagnostic produced in Demetra+. For each set of regressors, we calculed the percentage of series with bad diagnostic. We also calculated this percentage in case the series were adjusted with the set selected according to our algorithm. Results for retail trade sector are presented in Table 3.

Table 3 Percentage of retail trade turnover indexes with residual trading-day spectral peaks

No TD regressor	S0	S1	S2	S3	S4	Selected set
61%	65%	9%	37%	14%	16%	12%

As it could be expected, this proportion is smaller with the most complete set (S1), and higher with more parsimonious sets. But the proportion obtained with the sets selected according to our algorithm is barely greater than the one with the complete set S1.

6. Conclusion

In this article, we described a way to set up relevant sets of trading-day regressors, in order to take into account of both specificities of National calendar and sector of activity. We also proposed an algorithm for choosing, for each series, the best one, taking into account parsimony criterion but also quality of the Reg-ARIMA model. An application on French turnover indexes led to very satisfying results.

REFERENCES

Bell, W. R. (1984), Seasonal Decomposition of Deterministic Effects, Research Report, Statistical Research Division, U.S. Bureau of the Census, Washington D.C., RR84/01.

Bell, W. R. (1995), Correction to 'Seasonal Decomposition of Deterministic Effects' (n° RR84/01), Research Report, Statistical Research Division, U.S. Bureau of the Census, Washington D.C., RR95/01.

Bell, W. R., Hillmer, S. C. (1983), Modeling Time series with Calendar Variation, Journal of the American Statistical Association, 383, 78, 526-534.

Bessa, M., Dhifalli, R., Ladiray, D.,Lassoued, A., Maghrabi, B. (2009), Les effets de calendrier dans les séries tunisiennes, STATECO N°103.

Chen B-C., Findley D. F. (1993), Multiplicative trading day adjustments: X-11 and RegARIMA compared, American Statistical Association Annual Meetings, Alexandria.

Cholette P. (2003), A Note on the Treatment of Februaries in Trading-Day Adjustment, Working Paper, Time Series Research and Analysis Center, Statistics Canada.

Cleveland W. P. (1987), Calendar Adjustment of Time Series, Discussion Paper, Special Studies Section, Board of Governors of the Federal Reserve System, 198.

Cleveland W. P., Grupe M. R. (1983), Modeling time series when calendar effects are present, Applied Time Series Analysis of Economic Data, Zellner A. (editor), U.S. Department of Commerce, U.S. Bureau of the Census, Washington D.C., pp 57-67.

Dagum, E. B., (1980), The X11-ARIMA Seasonal Adjustment Method, Statistics Canada, Catalogue N° 12- 64E

Dagum, E. B., Quenneville, B., Sutradhar, B. (1992), Trading-day variations multiple regression model with random parameters, International Statistical Review, 1, 60, 57-73.

Fournier, J-M., Ladiray, D. (1995), Les effets de calendrier dans l'analyse conjoncturelle de la production, Economie et Statistique, 285-286, 115-126.

Findley D. F., Soukup R. J. (2000), Detection and Modeling of Trading Day Effects, Proceedings of the International Conference on Establishment Surveys.

Gómez, V. and Maravall, A. (1996), Programs TRAMO (Time series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instructions for the User, Working Paper 9628, Research Department, Banco de España.

Ladiray, D., Quenneville B. (2001), Désaisonnaliser avec la Méthode X11, Methodologica, numéro spécial 8- , Université Libre de Bruxelles

Lytras, Demetra P., Feldpausch, Roxanne M., and Bell, William R., Determining Seasonality: A Comparison of Diagnostics From X-12-ARIMA, U.S. Census Bureau

Maillard, V. (1994), Théorie et pratique de la correction des effets de jours ouvrables, document de travail, Direction des Etudes et Synthèses Economiques, INSEE, Paris, G9405.

Maravall, A. (2008), Notes on Programs Tramo and Seats, Bank of Spain

Soukup R. J., Findley D. F. (1999), On the Spectrum Diagnostics Used by X-12-ARIMA to Indicate the Presence of Trading Day Effects after Modeling or Adjustment, ASA Proceedings of the Joint Statistical Meetings (Business and Economic Statistics Section).

Tondering C. (2005), Frequently Asked Questions about Calendars, Version 2.8, available on www.tondering.dk/claus/calendar.html

U.S. Census Bureau (2009), X-12-ARIMA Reference Manual, Version 0.3, Time Series Staff, Statistical Research Division, Washington, DC

Young A. H. (1965), Estimating trading-day variations in monthly economic series, Technical Paper, U.S. Department of Commerce, U.S. Bureau of the Census, Washington D.C.

Young A. H. (1983), Comment on « Modeling Time Series when Calendar Effects are Present », Applied Time Series Analysis of Economic Data, Zellner, A. (editor), U.S. Department of Commerce, U.S. Bureau of the Census, pp 68-71.

ABSTRACT

The number of working days can explain some short-term movements in the time series. One more Saturday in a month for example drastically impacts the retail trade turnover in European countries. Apart the day composition of the month, other calendars effects such as public holidays or religious events may also affect the series. These periodic fluctuations, as well as the seasonality, are usually detected and eliminated in order to exhibit the irregular or non-periodic movements which are probably of most interest and importance.

In this presentation, we focus on two important practical problems: the design of adequate regressors taking into account the specificities of national calendars and the choice of a correct set of regressors, trading off between the parsimony of the model and the quality of the estimation.