Instituto Nacional de Estadística

Working Papers

01/2012

Analysis of the calendar effects on the Industry Turnover and New Orders Received Indices

Silvia Rama, Ignacio Arbués, María Mancebo, Luis Andrés de las Mozas and Eva María Vicente

First draft: August 2011

This draft: June 2012

Analysis of the calendar effects on the Industry Turnover and New Orders Received Indices

Abstract

Most economic monthly time series contain calendar effects. It is important to remove the calendar variation to allow an effective assessment of the variation due to other factors. Several methods exist which can adjust for trading-day and holiday effects in monthly economic time series. This paper reviews these methods and shows the procedure for determining the calendar adjustment carried out on the Industrial Turnover and New Orders Received Indices.

Keywords

Working-day adjustment, dynamic regression, ARIMA, model identification.

Authors and Affiliations

Silvia Rama and Ignacio Arbués

D. G. of Methodology, Quality and Information and Communications Technology

María Mancebo and Luis Andrés de las Mozas

S. G. for Short-Term Statistics and Prices

Eva María Vicente

S. G. for Population Statistics

Spain National Statistics Institute

Analysis of the calendar effects on the Industry Turnover and New Orders Received Indices

Silvia Rama, Ignacio Arbués, María Mancebo, Luis Andrés de las Mozas and Eva María Vicente

National Statistics Institute, Spain

June 22, 2012

Abstract

Most monthly time series contain calendar effects due to the fact that some levels of economic activity change over the days of the week and that the composition of the calendar changes over year (so a particular month contains a different configuration of days of the week each year). It is important to remove the calendar variation to allow an effective assessment of the variation due to other factors. Several methods exist which can adjust for trading-day and holiday effects in monthly economic time series. This paper reviews these methods and shows the procedure for determining the calendar adjustment carried out on the Industrial Turnover and New Orders Received Indices.

Contents

1 Introduction

The industrial turnover comprises the value of the invoicing of the establishment in the reference month, for sales of industrial goods and provision of industrial services, considering both those carried out by the establishment itself, and those performed through subcontracting with third parties.

According to this definition, turnover generally includes sales of finished products, semi-finished products, subproducts, waste and recovered materials, packages and packaging, merchandise (goods acquired for resale in the same state as that in which they were acquired), as well as the income from the provision of services related to the ordinary activity of the establishment. It also includes all other charges (transport, packaging, etc.) passed on to the customer, even if these charges are listed separately in the invoice. Subsidies received from public authorities or the institutions of European Union are not included. Turnover excludes VAT and other similar deductible taxes directly linked to turnover, as well as all duties and taxes on the goods or services invoiced by the unit. Reductions in prices, rebates and discounts as well as the value of returned packing must be deducted. Price reductions, rebates and bonuses that are applied afterwards, for example at the end of the year, are not taken into account.

An Order is defined as the value of the agreement, regardless of the form in which it is adopted (verbal, written, etc.), by which the producer agrees to supply some goods or to provide some industrial services to a third party, whether they have been performed by the producer or through subcontracting. The order is accepted if, by the judgment of the producer, there is sufficient proof that it is a valid agreement. As the order implies a future sale of goods and services, the headings to be considered in its definition are the same as those of turnover. Industrial New Orders Received are considered as the value of the orders received and firmly accepted by the observation unit during the reference period.

The Industry Turnover Index (ITI) and Industrial New Orders Received Index (INORI) are value indices, in other words, they measure the joint development of quantities, qualities and prices. Unlike what happens with the Industrial Production Index (IPI) or Industrial Price Index (IPRI), these indices are not based on a basket of representative products, but rather, the fundamental variable for their compilation is the main activity of the company. It is the objective of the Turnover Index to show the evolution of companies that are part of the industrial sector. The Industrial New Orders Received Index has the objective of measuring the evolution of the future demand aimed at the industrial branches.

The starting-point to calculate these indices are the fixed-base Laspeyres formula (currently base 2005), in which, each month, the current month is compared with the average of the twelve months of the year 2005.

An elementary aggregate is the lowest grouping component for which indices are obtained. In their calculation no weights are involved. Ratios of these aggregates are called elementary indices, and their definition is according to the formula

$$_{2005}I_i^{mt} = \frac{\sum_j f_{ji}^{mt}}{\frac{\sum_{m=1}^{12} \sum_j f_{ji}^{m2005}}{12}} \times 100,$$

where

- The index i ranges across certain aggrupations of activities according the CNAE classification[1] (originally, CNAE-93, but the series have been recalculated since January 2002, in base 2005 and with new CNAE-09),
- $_{2005}I_i^{mt}$ is the index, referred to year 2005, to the elementary aggregate i, in the month m of the year t,
- f_{ji}^{mt} is the invoice (or new orders received) value of the establishment j corresponding to elementary aggregate i,
- $\frac{\sum_{m=1}^{12} \sum_j f_{ji}^{m2005}}{12}$ is the average invoice (or new orders received) value in 2005 for all establishments j corresponding to the activity (elementary aggregate) i.

However, in practice, the process of calculating the basic indices should consider the possibility of non-response. In order to make uniform temporal comparisons, the information used in the calculation of elementary index is that provided by establishments that have collaborated in two consecutive months. Therefore, the elementary index is obtained by applying the monthly rate of change in the invoice (or new orders received) of establishments that have collaborated to the current and previous months

$$_{2005}I_i^{mt} = {}_{2005}\, I_i^{m-1t} \frac{\sum_j f_{ji}^{mt}}{\sum_j f_{ji}^{m-1t}},$$

[1]The Spanish version of the EU classification NACE.

where the sums in numerator and denominator are across all units common to periods t and $t-1$ and

- ${}_{2005}I_i^{mt}$ is the index, referred to year 2005, to the elementary aggregate i, in the month m of the year t,
- ${}_{2005}I_i^{m-1t}$ is the index, referred to year 2005, to the elementary aggregate i, in the month $m-1$ of the year t,
- f_{ji}^{mt} is the invoice (or new orders received) value of the establishment j, in the month m corresponding to elementary aggregate i,
- f_{ji}^{m-1t} is the invoice (or new orders received) value of the establishment j, in the month $m-1$ corresponding to elementary aggregate i.

For each elementary aggregate the weight used is the ratio between the value of the turnover of the industrial activity corresponding to this elementary aggregate and the total of the invoicing of the industries that comprise the population scope of those indices (sections B and C of the CNAE-09). These weights are computed with Structural Business Statistics (SBS) estimates as

$${}_{2005}W_i = \frac{\text{Turnover of the CNAE's activities of the elementary aggregate } i \text{ in 2005}}{\text{Turnover of all manufactured products, mining and quarrying (sections B and C) in 2005}}.$$

We need to take into account that, because we have not annual statistics that provide data regarding new orders received, the weights used to obtain the aggregate indices of the INORI are also calculated as the relation of the value of the invoicing of that basic aggregate over the total invoicing.

For other functional aggregations the weights are obtained as the sum of the weights of the elementary aggregates that compose them,

$$w_A = \sum_{i \in A} W_i.$$

The aggregate index, base 2005, for any functional aggregation or industrial sectors by economic destination is obtained as the aggregation of elementary indices belonging to that aggregation, with their corresponding weights. The index for the aggregate A is showed by the formula

$$ {}_{2005}I_A^{mt} = \sum_{i \in A} {}_{2005}I_i^{mt} \; {}_{2005}W_i^{CNAE}, $$

where

- ${}_{2005}I_A^{mt}$ is the index, referred to year 2005, to the aggregate A, in the month m of the year t,
- ${}_{2005}I_i^{mt}$ is the index, referred to year 2005, to the elementary aggregate i, in the month m of the year t,
- ${}_{2005}W_i^{CNAE}$ is the weight of the elementary aggregate i,
- ${}_{2005}W_i^{CNAE} = \frac{\text{Turnover of the CNAE's activities of the elementary aggregate } i \text{ in 2005}}{\text{Turnover of all CNAE's activities of the aggregate A in 2005}}$.

These indices are frequently influenced by the structure and composition of the calendar and by other fluctuations which can difficult the identification of relevant movements in the series, cause model misspecification and compromise the quality of the analysis. Hence, if Industrial Turnover and New Orders Received Indices will be inputs for economic analysis, it is important to remove the calendar effects from series to allow an effective interpretation.

In particular, monthly time series are affected by the length of the month and the composition of the month. In the first place, we should notice that the length of the month is not completely absorbed into the seasonal component. Because of the existence of Leap Years, the length of February is not the same every year. Thus, we have to remove the month-length effect (or, more precisely, the part of it that is purely non-seasonal). On the other hand, the month composition includes two other non-seasonal components, namely (i) the trading-day variation[2], that is, the variation in the monthly time series that is due to the changing number of times each day of the week occurs in a month and (ii) the Holiday variation. This last effect refers to the changes from year to year in the composition of the calendar with respect to holidays.

The most frequently used calendar adjustment procedures are the method of the Bureau of Census' X-12 ARIMA and the TRAMO/SEATS method. The methodology used in these programs is based on the discussion of calendar effects by Findley et al. (1998), using regARIMA models (which are regression models with seasonal ARIMA errors) to adjust the

[2]Some authors include leap year effect into trading-day variation.

series. However, almost all of the previous research on calendar effects has dealt with it in relation to the seasonal adjustment. Specifically, X-11 ARIMA (Young, 1965) uses a method based on modeling the residual component identified after removing the trend and seasonality. Cleveland and Devlin (1980, 1981) also apply regression methods to the series once removed the trend and seasonal component as X-11 ARIMA. This procedure has many drawbacks, so we prefer to use a model with ARIMA noise structure where we simultaneously estimate the regression and ARIMA parameters.

In the next section, we review the different methods present in the literature to remove the calendar effects. In section 3 we describe the expression of the calendar component that we use to adjust the series of the Industrial Turnover and New Orders Received Indices, whereas in section 4 we describe the methodology used to build models. The focus in section 5 and 6 is to show the models that we finally choose to adjust the series of the Industrial Turnover and New Orders Received Indices.

2 Different methods of adjustment

In this section we review the different methods in the literature to remove the calendar effects from economic time series, taking into account the distinction between flow and stock data. We also need to make a distinction between models where these effects are estimated from regARIMA model, and the alternative approach of indirect estimation from a regression model of the irregular component from a preliminary seasonal adjustment.

2.1 Methods of adjustment on flow series

The most used specification to describe trading-day effect is due to Hillmer (1982). He assumes that this effect can be approximated by a deterministic model. The model he uses is a sum of ARIMA and a regression model like the X-12 method. The effect attributable to the trading day is

$$TD_t = \sum_{i=1}^{7} \xi_i X_{it}, \tag{2.1.1}$$

where

X_{it}, $i = 1, ..., 7$, are respectively, the number of Mondays, Tuesdays, etc. in the month t, and

ξ_i $i = 1, ..., 7$, represent the average rates of activity on Mondays, Tuesdays, etc.

In order to reduce the correlations between the estimates, he performs a reparameterization of this formula as

$$TD_t = \sum_{i=1}^{7} \beta_i T_{it}, \tag{2.1.2}$$

where

$\beta_i = \xi_i - \bar{\xi}$, $i = 1, ..., 6$ and $\beta_7 = \bar{\xi}$ with $\bar{\xi} = \frac{1}{7} \sum_{i=1}^{7} \xi_i$,

$T_{it} = X_{it} - X_{7t}$, $i = 1, ..., 6$ and $T_{7t} = \sum_{i=1}^{7} X_{it}$.

Inference about the parameters can be made using asymptotic theory. To test if the daily effects are different for the different days of the week we must test

$$\begin{aligned} H_0 &: \quad \xi_1 = ... = \xi_7 \\ H_1 &: \quad \text{not all } \xi_i \text{ are equal.} \end{aligned}$$

This is equivalent to testing

$$\begin{aligned} H_0 &: \quad \beta_1 = ... = \beta_6 = 0 \\ H_1 &: \quad \text{at least one } \beta_i \neq 0, \text{ where } i \in \{1, ..., 6\}. \end{aligned}$$

Bell and Hillmer (1983) also propose a model for the holiday effects that accounts for the daily impact of Easter on the level of the series. Then, the Easter effect in t, E_t, is

$$E_t = \sum_{i=1}^{K} \tilde{\alpha}_i n_{i,t}, \tag{2.1.3}$$

where

$\tilde{\alpha}_i$ denotes the effect on the ith day before the Easter,

$$n_{i,t} = \begin{cases} 1 & \text{when the } i\text{th day before Easter falls in month } t \\ 0 & \text{otherwise} \end{cases},$$

K denote some suitable upper bound on the length of the effect in days.

The relationship (2.1.3) is expressed in terms of daily levels. Unfortunately, in most situations, only monthly values are available, so it is difficult to estimate the coefficients. Hence, they define a similar model where they group some terms using a behavior pattern. Any grouping of the $n_{i,t}$ can be used as long as it produces explanatory variables that are linearly independent. The new model with N groups is defined as

$$E_t = \sum_{\iota=1}^{N} \alpha_\iota n_{\iota,t}, \tag{2.1.4}$$

where

$n_{\iota,t} = \sum_{i \in \iota} n_{i,t}$, $\iota = 1, ..., N$, can be defined as the number of days of the time period ι that fall in month t,

α_ι denote the effect on the series of the ι period of days before the Easter.

Later, William Bell (1984) proposes for the model (2.1.1) the decomposition of the trading-day effect

$$\begin{aligned} TD_t &= \sum_{i=1}^{7} \xi_i X_{it} = \sum_{i=1}^{6} \beta_i (X_{it} - X_{7t}) + \beta_7 \sum_{i=1}^{7} X_{it} \\ &= \sum_{i=1}^{6} \beta_i (X_{it} - X_{7t}) + \beta_7 LF_t + \beta_7 \omega_t + \beta_7 (30.4375), \end{aligned} \tag{2.1.5}$$

where

the value $30.4375 = \frac{365.25}{12}$ is the average month length over a 4-year cycle,

$$LF_t = \begin{cases} -0.25 & \text{in a non-leap year February} \\ 0.75 & \text{in a leap year February} \\ 0 & \text{otherwise} \end{cases},$$

$$\omega_t = \begin{cases} 0.5625 & \text{in a month with 31 days} \\ -0.4375 & \text{in a month with 30 days} \\ -2.1875 & \text{in February} \end{cases}.$$

The first two terms of the last equation belong to the trading-day effect and the third term is purely seasonal.

Cleveland and Grupe (1981) also propose a regression model for the calendar effects and an ARIMA model for the residuals. The basic equation for the effect of composition and length of the month is the same that Hillmer defines (2.1.1). However, they rewrite this expression to show a new decomposition to the trading-day effect as

$$\begin{aligned} TD_t &= \sum_{i=1}^{7} X_{it}\xi_i - \frac{N_t}{7}\sum_{i=1}^{7}\xi_i + \frac{N_t-\bar{N}}{7}\sum_{i=1}^{7}\xi_i + \frac{\bar{N}}{7}\sum_{i=1}^{7}\xi_i \qquad (2.1.6)\\ &= \sum_{i=1}^{7} X_{it}(\xi_i-\bar{\xi}) + (N_t-\bar{N})\bar{\xi} + \bar{N}\bar{\xi}\\ &= \sum_{i=1}^{7}\big(X_{it}-\frac{X_{it}}{7}\big)(\xi_i-\bar{\xi}) + (N_t-\bar{N})\bar{\xi} + \bar{N}\bar{\xi}, \end{aligned}$$

where

$N_t = \sum_{i=1}^{7} X_{it}$ is the length in days of month t, equal to the variable T_{7t} defined above,

$\bar{N} = \frac{365.25}{12}$ is the average month length over a 4-year cycle,

$\frac{N_t}{7}$ is the average number of days of each type in month t.

The first term of (2.1.6) corresponds to the effect of the composition of the month and $(N_t - \bar{N})$ represents the effect of the length. This effect can be expressed as a function $f = X\upsilon$ where X is a design matrix with T rows and υ is a vector of estimated coefficients. The first seven columns of X are $(X_{it} - N_t/7)$, $i = 1, 2, ..., 7$ and the eighth column is $(N_t - \bar{N})$. The first seven υ_i of vector υ are the differences $\xi_i - \bar{\xi}$ and the last, υ_8, is equal to $\bar{\xi}$.

For U.S Bureau of Census' X-12 ARIMA program we draw the information from Findley et all (1998), Monsell (2010), Findley and Soukup (2000). X-12 ARIMA program uses a regARIMA model where the calendar effects is accounted for by the (similar to (2.1.1)) formula

$$\begin{aligned} TD_t &= \sum_{i=1}^{7}\xi_i X_{it} = \bar{\xi}N_t + \sum_{i=1}^{7}(\xi_i-\bar{\xi})X_{it} \qquad (2.1.7)\\ &= \bar{\xi}N_t + \sum_{i=1}^{6}(\xi_i-\bar{\xi})(X_{it}-X_{7t}). \end{aligned}$$

The equation (2.1.7) has a seasonal component that resides in the calendar month means of $\bar{\xi}N_t$. Hence, X-12 ARIMA program uses a deseasonalized formula obtained removing calendar month means from (2.1.7). The deseasonalized calendar effect depends on the kind of model specified (additive, multiplicative...). For the additive model, the new formula is obtained by subtracting the calendar month means, $\bar{\xi}N_t^*$, as

$$\begin{aligned} TD_t &= \bar{\xi}N_t - \bar{\xi}N_t^* + \sum_{j=1}^{6}(\xi_i - \bar{\xi})(X_{it} - X_{7t}) \qquad (2.1.8) \\ &= \bar{\xi}(N_t - N_t^*) + \sum_{i=1}^{6}(\xi_i - \bar{\xi})(X_{it} - X_{7t}) \\ &= \bar{\xi}(N_t - N_t^*) + \sum_{i=1}^{6}\beta_i(X_{it} - X_{7t}), \end{aligned}$$

where N_t^* is the average length of the month of the year corresponding to t (that is, if t is January, then N_t^* is the average length of January across years).

For the case of a multiplicative decomposition we deseasonalize and detrend the trading-day effect by dividing (2.1.7) by $\bar{\xi}N_t^*$. Setting $\tilde{\beta}_i = \xi_i/\bar{\xi} - 1$, we have

$$\frac{N_t}{N_t^*} + \sum_{i=1}^{6}\tilde{\beta}_i\left(\frac{X_{it} - X_{7t}}{N_t^*}\right) = \frac{\sum_{i=1}^{7}(\tilde{\beta}_i + 1)X_{it}}{N_t^*}. \qquad (2.1.9)$$

Normally we work with log-additive models. For this kind of models, the deseasonalized calendar effects can be obtained by taking the logarithm on the LHS of (2.1.9) and using the approximation $\log(1 + x) \approx x$ as

$$\log\left\{1 + \frac{N_t - N_t^*}{N_t^*} + \sum_{i=1}^{6}\frac{(\xi_i - \bar{\xi})}{\bar{\xi}}\left(\frac{X_{it} - X_{7t}}{N_t^*}\right)\right\} \approx \frac{N_t - N_t^*}{N_t^*} + \sum_{i=1}^{6}\tilde{\beta}_i\left(\frac{X_{it} - X_{7t}}{N_t^*}\right). \qquad (2.1.10)$$

From this equation we can obtain the four types of trading-day models used in X-12 ARIMA for log-additive models. The first model is given by

$$\log Y_t = \lambda_0 LY_t + \sum_{i=1}^{6}\lambda_i(X_{it} - X_{7t}) + Z_t, \qquad (2.1.11)$$

where

$LY_t = N_t - N_t^*$ and,

Z_t denotes a process with a user-specified ARIMA structure.

The second model is similar but it has only six estimated trading-day coefficients because the leap year coefficient, λ_0, can be replaced by a constant equal to $\frac{1}{N_t^*}$, where N_t^* is the February average, so the constant is $1/28.25 \approx 0.0354$.

A more parsimonious model, which was originally suggested by TRAMO (Gómez and Maravall, 1996), arises from reducing the number of trading-day regressors from six to one by assuming the daily effect of weekdays (Monday through Friday) is the same, and the daily effect of weekend days (Saturday and Sunday) is the same. Thus, the number of weekend days is subtracted from the number of weekdays, providing a single regressor. Noticing that $\sum_{i=1}^{7} \lambda_i = 0$ with the constraints $\lambda_1 = ... = \lambda_5$ and $\lambda_6 = \lambda_7$ we can derive the regressor as

$$\sum_{i=1}^{5} \lambda_i + \sum_{i=6}^{7} \lambda_i = 0 \qquad 5\lambda_{M-F} + 2\lambda_{S-S} = 0 \qquad \lambda_{S-S} = -\frac{5}{2}\lambda_{M-F}$$

$$\begin{aligned}
TD_t &= \sum_{i=1}^{6} \lambda_i (X_{it} - X_{7t}) + \lambda_0 LY_t \\
&= \sum_{i=1}^{5} \lambda_i (X_{it} - X_{7t}) + \lambda_6 (X_{6t} - X_{7t}) + \lambda_0 LY_t \\
&= \lambda_{M-F} \sum_{i=1}^{5} X_{it} - 5\lambda_{M-F} X_{7t} + \lambda_{S-S} X_{7t} - \lambda_{S-S} X_{6t} + \lambda_0 LY_t \\
&= \lambda_{M-F} \sum_{i=1}^{5} X_{it} - \frac{5}{2}\lambda_{M-F} (X_{7t} + X_{6t}) + \lambda_0 LY_t \\
&= \lambda_{M-F} (\sum_{i=1}^{5} X_{it} - \frac{5}{2} \sum_{i=6}^{7} X_{it}) + \lambda_0 LY_t.
\end{aligned}$$

Thus, the third model can be expressed as

$$\log Y_t = \lambda_0 LY_t + \lambda_{M-F} \left(\sum_{i=1}^{5} X_{it} - \frac{5}{2} \sum_{i=6}^{7} X_{it} \right) + Z_t, \qquad (2.1.12)$$

where if we replace the leap year coefficient by a constant, $\frac{1}{N_t^*}$, we will have the fourth model.

The X-12 ARIMA program also includes Easter effect. The regressor for this effect assumes that the fundamental structure of the time series changes for a fixed number of days before Easter and remains at the new level until the day before Easter. For a given effect τ

the Easter regressor is generated as

$$E(\tau,t) = \frac{n_{\tau,t}}{\tau} - \mu_{\tau,t}, \tag{2.1.13}$$

where

$n_{\tau,t}$ is the number of days before Easter of the total τ days of Easter that fall in the month t,

$\mu_{\tau,t}$ is the long-run monthly mean.

Monsell (2007) proposes some variations to the X-12 ARIMA Easter effect. The main critic that Monsell makes of this Easter regressor is that assuming that the level of activity is raised by a constant level for the τ days before Easter is unrealistic.

The first alternative form assumes that we can break down the Easter effect in two parts: a pre-holiday effect from the τth days before Easter Sunday to the day before Good Friday, and an effect of the period starting on Good Friday and lasting until Easter Sunday. Both regressors are generated as

$$BE(\tau,t) = \frac{n_t^{BE}}{\tau - 3} - \mu_{\tau,t}^{BE} \tag{2.1.14}$$

$$DE(t) = \frac{n_t^{DE}}{3} - \mu_t^{DE}, \tag{2.1.15}$$

where

n_t^{BE} is the number of days from the τth day before Easter to the day before Good Friday that fall in month t,

n_t^{DE} is the number of days between Good Friday and Easter that fall in month t,

$\mu_{\tau,t}^{BE}$ and $\mu_{\tau,t}^{DE}$ are the long run monthly means used to center the regressors.

The next alternative regressor assumes that the level of activity before Easter increases linearly before the holiday. The linear regressors for March and April are

$$LE(\tau, \text{March}, y) = \frac{n^2_{\text{March},y}}{\tau^2} - \mu^{LE}_{\tau,\text{March}} \tag{2.1.16}$$

$$LE(\tau, \text{April}, y) = \left(1 - \left(\frac{n^2_{\text{April},y}}{\tau^2}\right)\right) - \mu^{LE}_{\tau,\text{April}}, \tag{2.1.17}$$

where

$n_{\text{March},y}$ and $n_{\text{April},y}$ are the number of days before Easter falling in March and April respectively for year y,

$\mu^{LE}_{\tau,\text{March}}$ and $\mu^{LE}_{\tau,\text{April}}$ are the long run monthly means of the first part of the equations (2.1.16) and (2.1.17) respectively.

These two regressors allow to distinguish between pre-holiday and during holiday effect too. The final proposed regressor assumes that the change in the level of activity for the weekend days is different than the change for the weekdays leading up to Easter.

$$WE(t) = \frac{n_{we,t}}{8} - \mu_{we} \tag{2.1.18}$$

$$WD(t) = \frac{n_{wd,t}}{8} - \mu_{wd}, \tag{2.1.19}$$

where

$n_{we,t}$ is the number of days in month t that fall on a Friday, Saturday and Sunday in the period of 16 days before Easter inclusive,

$n_{wd,t}$ is the number of days in month t that fall on a Monday, Tuesday, Wednesday or Thursday in the period of 16 days before Easter,

μ_{we} and μ_{wd} are the long-run monthly means of the first part of the equations (2.1.18) and (2.1.19) respectively.

On the other hand, there are some methods of adjustment based on a regression model for the irregular component. They are motivated by their historical success, by practical considerations and by certain requests of statistical agencies and central banks in different countries. These methods consist of deseasonalizing and detrending the original series to obtain the irregular component and then, identifying the calendar effects over this component. This implies that the regression models for the irregular component should also be deseasonalized and detrended.

In this context, Cleveland and Devlin (1980) propose the following procedure for fitting the calendar components:

a) calculate the month-length corrected monthly series,

b) choose a power transformation,

c) remove trend and seasonal components,

d) estimate the calendar parameters.

Therefore, the first step is to calculate the average of the daily values of a series for each month

$$\bar{x}(t) = \frac{\sum_i X(i)}{\text{number of days in month } t},$$

where $X(i)$ is the daily value of the series for the ith day. From $\bar{x}(t)$, they calculate the month-length-corrected series $x(t) = \bar{N}\bar{x}(t) = 30.4375\bar{x}(t)$.

The second step is to decide the value of the parameter p of the power transformation (Box and Cox, 1964; Tukey, 1957) defined by

$$u^{(p)} = \begin{cases} u^p & \text{if } p > 0 \\ \log u & \text{if } p = 0 \\ -u^p & \text{if } p < 0 \end{cases}.$$

They suppose that the power-transformed daily data, $X^{(p)}(i)$, has four additive components $X^{(p)}(i) = T(i) + S(i) + C(i) + I(i)$, where $T(i)$ is the trend component in the daily series, $S(i)$ is the seasonal component with a period of one year, $I(i)$ is the irregular component and $C(i)$ accounts for a day of the week effect in series so that $C(i) = \pi_k$, if i is the kth day of the week, where $\sum_{k=1}^{7} \pi_k = 0$. For the transformed monthly series, $x^{(p)}(t)$, depending on the power transformation that we use the model can be exactly or approximately equal to

$$x(t)^{(p)} = t(t) + s(t) + c(t) + i(t), \tag{2.1.20}$$

where

$s(t)$, $t(t)$ and $i(t)$ are similarly defined as $x(t)$, in other words, they are the aggregate values of the respective components divided by the month length and multiplied by the average month length,

$c(t) = \sum_{k=1}^{7} \pi_k d_k(t)$, $k = 1, ..., 7$, represents the calendar effects,

$d_k(t) = (30.4375)\bar{d}_k(t)$ being $\bar{d}_k(t)$ the fraction of times that the kth day of the week occurs in the month t.

They now apply a procedure, to remove the trend and seasonality, assuming that it can be represented by a linear function L. Then, they estimate the coefficients π_k regressing $Lx^{(p)}(t)$ on the seven explanatory variables $Ld_k(t)$ $k = 1, ..., 7$ subject to constraint that $\sum_{k=1}^{7} \pi_k = 0$.

Young (1965) uses a similar method based on two main assumptions: (i) all residual trading-day variation appears in the irregular component and (ii) the variation may be expressed in terms of 7 daily weights. First, he makes a preliminary seasonal adjustment to obtain the irregular component, I, and deletes the extreme values from it. To the new irregular component, $\tilde{I}$, the transformation (2.1.21) is applied to ensure that the estimated weights sum zero, obtaining $\bar{I}$ as

$$\bar{I}_t = \left(\frac{\tilde{I}_t}{100} - 1 \right) N_t^*, \tag{2.1.21}$$

where N_t^* is the average number of days in that month. Usually, it is 31, 30 or 28.25 depending upon whether month t is a 31-day, 30-day month or February. However he proposes to use the average length of month $\bar{N} = (30.4575 = \frac{365.25}{12})$ for all t if no allowance for the length of the month is desired.

This variable is regressed as

$$\bar{I} = Xb + E, \tag{2.1.22}$$

where

$\bar{I} = [\bar{I}_1, \bar{I}_2, ..., \bar{I}_n]$ is the vector of transformed irregular components,

$E = [E_1, E_2, ..., E_n]$ is the vector of the true irregular series,

$b = [b_1, b_2, ..., b_7]$ is the vector of daily weights to be estimated,

$X = [X_{1t}, X_{2t}, ..., X_{7t}]$ is the matrix of independent variables where $X_{1t}, X_{2t}, ..., X_{7t}$ is the number of Mondays, Tuesdays,...,Sundays in a giving month t,

n is the number of months included in the regression.

Since the weights add up to $(\sum_{i=1}^{7} b_i = 0)$, by definition $b_7 = -\sum_{i=1}^{6} b_i$, and so $\bar{I} = Xb+E$ becomes $\bar{I} = T\hat{b} + E$ where $\hat{b} = [b_1, b_2, ..., b_6]$ and $T = [T_{1t}, T_{2t}, ..., T_{6t}]$ where $T_{it} = X_{it} - X_{7t}$ or equivalently,

$$\bar{I}_t = \sum_{i=1}^{6} T_{it} b_i + E_t \qquad (t = 1, 2, ..., n). \tag{2.1.23}$$

2.2 Methods of adjustment for stock series

Many authors have developed models for calendar effects for series that represent stock at the end of the month. Cleveland an Grupe (1981) define the stock at the end of the month t, S_t, as $S_t = S_{t-1} + \sum_i F_t(i)$ where $F_t(i)$ is the flow on the day i of the month t. Hence, they propose that since (2.1.6) is intended to describe flows, it is related to the first difference of stock data, so the same matrix X can be used. To transform the effect back to the non-differenced data, the columns of X must be integrated to form W. The first row of W is set to zero and the rest are formed recursively as $w_t' = w_{t-1}' + x_{t-1}'$ where w_t' is the row t of W.

William Bell (1984) develops the model proposed by Cleveland and Grupe. Let F_t be a monthly flow series and let $I_t = I_0 + \sum_{j=1}^{t} F_j$ be the end of the month stock series, starting at $t = 0$. Assume the trading-day effect in I_0 is zero. Then, if the trading-day effect in F_j is $\sum_{i=1}^{7} \xi_i X_{ij}$, that in I_t is

$$\sum_{j=1}^{t} \sum_{i=1}^{7} \xi_i X_{ij} = \sum_{j=1}^{t} \sum_{i=1}^{7} [(\xi_i - \bar{\xi}) X_{ij} + \bar{\xi} X_{ij}] = \sum_{j=1}^{t} \sum_{i=1}^{7} \beta_i X_{ij} + \bar{\xi} \sum_{j=1}^{t} N_j \tag{2.2.1}$$

$$= \sum_{k=1}^{7} \gamma_k I_t(k) + \bar{\xi} \sum_{j=1}^{t} N_j = \sum_{k=1}^{7} (\gamma_k - \bar{\gamma}) I_t(k) + \bar{\gamma} + \bar{\xi} \sum_{j=1}^{t} N_j,$$

where β_i is as in (2.1.2),

$$I_t(k) = \begin{cases} 1 & \text{if the month } t \text{ ends on a } k\text{-th day} \\ 0 & \text{otherwise} \end{cases},$$

k_0 is the type of day of the week just before the start of month $t = 1$,

$N_j = \sum_{i=1}^{7} X_{ij}$, $j = 1, ..., t$,

$\gamma_k = \sum_{i=1}^{k} \beta_i + \gamma_k$ for $k = 1, ..., 6$ and $\gamma_7 = -\sum_{i=1}^{k_0} \beta_i$,

$\bar{\gamma} = \frac{1}{7} \sum_{k=1}^{7} \gamma_k$.

Bell (1995) defines $\tilde{\gamma}_k = \gamma_k - \bar{\gamma}$ where $\sum_{k=1}^{7} \tilde{\gamma}_k = 0$, and rewrites the first term of (2.2.1) as

$$\sum_{k=1}^{7}(\gamma_k - \bar{\gamma})I_t(k) = \sum_{k=1}^{7}\tilde{\gamma}_k I_t(k) = \sum_{k=1}^{7}\tilde{\gamma}_k(I_t(k) - I_t(7)) + I_t(7)\sum_{k=1}^{7}\tilde{\gamma}_k = \sum_{k=1}^{6}\tilde{\gamma}_k I_t^*(k), \quad (2.2.2)$$

where $I_t^*(k) = \begin{cases} 1 & \text{if the month } t \text{ ends on a } k\text{-th day} \\ -1 & \text{if the month } t \text{ ends on Sunday} \\ 0 & \text{otherwise} \end{cases}$, $k = 1, ..., 6$.

Findley and Monsell (2007) consider the one-coefficient weekly-weekend contrast that arises from equality constraints between the weekday coefficients, $\xi_1 = \xi_2 = ... = \xi_5$, and between ξ_6 and ξ_7 in the model (2.2.1) (or equivalently $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$ and $\beta_6 = \beta_7$). If these restrictions, together with $\sum_{i=1}^{7}\beta_i = 0$, are imposed to the model for the flow series, we can derive a single regressor for the trading-day effect (as X-12 ARIMA program (2.1.12)). However, in stock series it is not so easy. The restriction may be obtained in terms of the parameters $\tilde{\gamma}_k$. Through the relationship between the coefficients β_i and γ_k (explained below) the authors obtain the restricted model

$$\frac{3}{5}I_{1t} - \frac{1}{5}I_{2t} + \frac{1}{5}I_{3t} + \frac{3}{5}I_{4t} + I_{5t}. \quad (2.2.3)$$

Now we explain the relationship between the flow and stock coefficients (β_i and γ_k respectively). We have the equation system

$$\begin{aligned} \gamma_1 &= \beta_1 + \gamma_7 \\ \gamma_2 &= \beta_1 + \beta_2 + \gamma_7 \\ &\cdots \\ \gamma_6 &= \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \gamma_7, \end{aligned}$$

or in matrix notation

$$\begin{bmatrix} \gamma_1 - \gamma_7 \\ \gamma_2 - \gamma_7 \\ \gamma_3 - \gamma_7 \\ \cdots \\ \gamma_6 - \gamma_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ \cdots \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \cdots \\ \beta_6 \end{bmatrix},$$

$$[\gamma_k - \gamma_7] = L\beta \qquad k = 1, ..., 6.$$

Using $\sum_{k=1}^{7} \gamma_k = 0$, we observe that $\gamma_k - \gamma_7 = \tilde{\gamma}_k - \tilde{\gamma}_7 = \tilde{\gamma}_k + \sum_{j=1}^{6} \tilde{\gamma}_j = 2\tilde{\gamma}_k + \sum_{j \neq k} \tilde{\gamma}_j$. Thus $[\gamma_k - \gamma_7] = M\tilde{\gamma}_k$, $k = 1, ..., 6$ where

$$M = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 2 \end{bmatrix}.$$

Therefore we have $L\beta = M\tilde{\gamma}$, so $\tilde{\gamma} = M^{-1}L\beta = N\beta$. Taking into account the restriction where $\beta_1 = ... = \beta_5$ and $\beta_6 = \beta_7$ ($\beta = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad -\frac{5}{2}]'\beta_5$) we obtain

$$\tilde{\gamma} = \frac{1}{7} \begin{bmatrix} 1 & -5 & -4 & -3 & -2 & -1 \\ 1 & 2 & -4 & -3 & -2 & -1 \\ 1 & 2 & 3 & -3 & -2 & -1 \\ 1 & 2 & 3 & 4 & -2 & -1 \\ 1 & 2 & 3 & 4 & 5 & -1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -\frac{5}{2} \end{bmatrix} \beta_5 = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \\ 5 \\ 0 \end{bmatrix} \frac{1}{2}\beta_5.$$

Since $\tilde{\gamma}_5 = -\frac{5}{2}\beta_5$, the relationship that describes the new model is

$$\tilde{\gamma} = [\ -\tfrac{3}{5} \quad -\tfrac{1}{5} \quad \tfrac{1}{5} \quad \tfrac{3}{5} \quad 1 \quad 0\ ]'\tilde{\gamma}_5.$$

3 A model for calendar effects

In this section we describe the model that was finally chosen for the calendar effects in the Industrial Turnover and New Orders Received Indices. This choice was largely a trade-off between two kinds of requirements, namely (a) related to the quality of the adjustment and (b) related to policies on consistency, comparability and standardization. Because of the second kind of requirements, the adjustment of TRAMO/SEATS was considered (since its use is a practice recommended by Eurostat). Consistency and comparability also suggested to use a similar adjustment to that of the Industrial Production Index (IPI).

There are important reasons for not using the TRAMO/SEATS adjustment. The program deals with the trading-day effect as the variation in a monthly time series that is due to the changing number of time each day of the week occurs in a month. However, we define the trading-day effect as the variation due to the changes in the number of working and non-working days in a month, taking into account not only the day of the week but also the holidays.

Some adjustment methods used at the Institute are discussed in 3.1 and we describe the model chosen for Turnover and New Orders Received in 3.2.

Since the form of the regressor used in the adjustment of IPI is different from that of the finally chosen regressor, we considered that a theoretical analysis of the consequences of that difference was in order. This matter is discussed in 3.3.

3.1 Some methods of adjustment used in the Spain National Statistics Institute

In the Spain National Statistics Institute, different ways to correct the calendar effects of time series are used, but all of them within the regARIMA framework. They are similar to the regressors reviewed in section 2, but accounting for the holidays. For the treatment of a series of Industrial Production Index, the model used for working day effect is

$$TD_t = MF_t - HOL_t, \tag{3.1.1}$$

where

$MF_t = \sum_{i=1}^{5} X_{it}$,

X_{it}, $i = 1, ..., 7$ is the number of Mondays, Tuesdays, etc. respectively in the month t,

$HOL_t = \sum_{\delta=1}^{17} \left(w_\delta \sum_{i=1}^{N_t} I_{i\delta t} \right)$,

w_δ, $\delta = 1, ..., 17$ is the weight of the Autonomous Community δ in industry value added for 2005,

$$I_{i\delta t} = \begin{cases} 1 & \text{if the } i\text{-th day is holiday in the Autonomous Community } \delta \\ 0 & \text{otherwise} \end{cases} .$$

For the Holy Week effect, they base the model in two main assumptions. One of them is that Easter has a global duration of 8 days and the second is that this effect is not constant along that period. Consequently, they identify three different periods of time that have specific weights. The first period is composed by Monday, Tuesday and Wednesday before Good Friday. The second extends from Holy Thursday to Easter Sunday. The last period only consists on Easter Monday.

In the Retail Trade Index the methodology used to model the trading-day effect is similar to the X-12-ARIMA program, taking into account the moving holidays, which change the number of working days in a month and the characteristics of this sector. The proposed model is

$$TD_t = H_t - 5.59 I_t \tag{3.1.2}$$

$$I_t = X_{7t} + NH_t + C_t \qquad H_t = \sum_{i=1}^{7} X_{it} - I_t,$$

where

H_t and I_t are the number of working and non-working days respectively,

X_{it}, $i = 1, ..., 7$ is the number of Mondays,....,Sundays in the month t respectively,

NH_t is the number of national holidays (except Good Friday and the holiday that fall in Sunday) in the month t,

$C_t = \sum_{\delta=1}^{17} w_\delta I_{\delta it}$ where $I_{\delta it}$ is 1 if the i-th day is holiday in the Autonomous Community δ and 0 otherwise (except Easter Thursday and Monday), and w_δ are the weights of each Autonomous Community δ in the Retail Trade Index,

$\sum_{i=1}^{7} X_{it}$ is the total number of days in the month t,

5.59 is the average of the relationship between working/non-working days calculated in the sample period.

They consider that to the Easter effect is necessary to take as duration a period of five days (from Easter Thursday to Easter Monday) taking into account that some Communities celebrate only Easter Thursday, or Easter Monday or both. So the Easter effect is represented as

$$E_t = \sum_{\delta=1}^{17} h_{\delta,t} w_\delta, \tag{3.1.3}$$

where $h_{\delta,t}$ is the proportion of Easter days that fall in the month t in the Autonomous Community δ.

3.2 Modeling calendar effects

A. MODELING TRADING-DAY VARIATION

We assume that trading day effects can be approximated by a linear deterministic model. Initially we part of the idea that each day of the month has a different effect on industrial activity. If D_t denotes the accumulate effect of the days of month t then

$$TD_t = \sum_{i=1}^{N_t} \sum_{j=1}^{d} X_{tij} \xi_j \tag{3.2.1}$$

where N_t is the length of the month, we distinguish among d day groups (for example, $d = 7$ for the days of the week) and

$$X_{tij} = \begin{cases} 1 & \text{if the } i\text{th day of month } t \text{ belongs to group } j \\ 0 & \text{otherwise} \end{cases}$$

ξ_j is the average effect of a day from the jth group

We may rewrite (3.2.1) first as

$$TD_t = \sum_{j=1}^{d} N_{tj} \xi_j, \tag{3.2.2}$$

where $N_{tj} = \sum_{i=1}^{N_t} X_{tij}$ is the number of days of group j in month t and then

$$TD_t = \sum_{j=1}^{d} N_{tj} \beta_j + \bar{\xi} N_t \tag{3.2.3}$$

where $\beta_j = \xi_j - \bar{\xi}$ and $\bar{\xi}$ represents the activity of an average day. This average can be obtained either as a group mean

$$\bar{\xi} = \frac{1}{d} \sum_{j=1}^{d} \xi_j \tag{3.2.4}$$

or as a long-term average

$$\bar{\xi} = \frac{\sum_{t=1}^{T} \sum_{j=1}^{d} N_{tj} \xi_j}{\sum_{t=1}^{T} N_t}, \tag{3.2.5}$$

where T is a long enough number of months. If we consider only the effect of the days of the week, this choice is of not great consequence, but when we take into account the effect of holidays, then there is a difference and (3.2.4) makes no much sense. On the other hand, (3.2.4) and (3.2.5) imply respectively either

$$\sum_{j=1}^{d} \beta_j = 0 \qquad \text{or} \qquad \sum_{t=1}^{T} \sum_{j=1}^{d} N_{tj} \beta_j = 0, \tag{3.2.6}$$

which we will use later.

Let us assume now that the level of industrial activity is the same for all working day and for all non-working day, in other words, $\forall j \in \Omega$ (where Ω is the set of working days) $\beta_j = \beta_{WD}$ and $\forall j \in \Theta$ (where Θ is the set of non-working day) $\beta_j = \beta_{ND}$. On the other hand the β_j coefficients have to satisfy the either part of (3.2.6) depending on the definition of $\bar{\xi}$. Let us assume that $\bar{\xi}$ is the long-term average. Then

$$\sum_{t=1}^{T} \sum_{j=1}^{d} N_{tj} \beta_j = \beta_{WD} \sum_{t=1}^{T} \sum_{j \in \Omega} N_{tj} + \beta_{ND} \sum_{t=1}^{T} \sum_{j \in \Theta} N_{tj} = 0 \tag{3.2.7}$$

so

$$\beta_{ND} = -\frac{\sum_{t=1}^{T} WD_t}{\sum_{t=1}^{T} ND_t} \beta_{WD} = -\frac{\overline{WD}}{\overline{ND}} \beta_{WD}$$

where WD_t and ND_t are the number of working and non-working days in month t respectively[3], and $\overline{WD} = (1/T) \sum_{t=1}^{T} WD_t$, $\overline{ND} = (1/T) \sum_{t=1}^{T} ND_t$ are their long-term means.

Therefore,

$$TD_t = \sum_{j=1}^{d} N_{tj} \beta_j + \bar{\xi} N_t = \beta_{WD} \sum_{j \in \Omega} N_{tj} + \beta_{ND} \sum_{j \in \Theta} N_{tj} + \bar{\xi} N_t =$$

$$\beta_{WD} WD_t - \frac{\overline{WD}}{\overline{ND}} \beta_{WD} ND_t + \bar{\xi} N_t = \beta_{WD} \left\{ WD_t - \frac{\overline{WD}}{\overline{ND}} ND_t \right\} + \bar{\xi} N_t. \tag{3.2.8}$$

We shall obtain a deseasonalized and level neutral version of (3.2.8) by removing calendar month means. The monthly calendar repeats itself over any 400-year cycle or 28-year cycle

[3]If we distinguish only the days of the week, $\bar{\xi} = (1/7) \sum_{j=1}^{7} \xi_j$ and we consider Monday to Friday as working days, then we arrive at $\beta_{ND} = -(5/2)\beta_{WD}$

if we omit the fact that only one of each four secular year is leap year. Consequently, the variables N_{tj} (the number of ith days in a month t for $i = 1, ..., 7$) are periodic with period $336 (= 12 \times 28)$, and the calendar month means $(1/28) \sum_{k=1}^{28} N_{t+12k,j}$ have the same value for all t and j. This implies that the calendar month means of the difference variables N_{tj} are zero and so, the seasonal component of this part annihilates. However, if we take into account the number of working or non-working days its means are not the same. So we need to discount the calendar month mean of this variables of the calendar effects. There is another seasonal component that resides in $\bar{\xi} N_t$. Because the length of the month is repeated every four years, $N_{t+48} = N_t$, the calendar month means of $\bar{\xi} N_t$ is $N_t^* = (1/4) \sum_{k=1}^{4} N_{t+12k}$. The final model where these components are removed is

$$TD_t = \beta_{WD} \left[(WD_t - \overline{WD}_t) - \frac{\overline{WD}}{\overline{ND}} (ND_t - \overline{ND}_t) \right] + \bar{\xi} LY_t \qquad (3.2.9)$$

where $\overline{WD}_t$ and $\overline{ND}_t$ are the calendar month means of working and non-working days respectively.

B. MODELING EASTER EFFECT

To the current Easter regressor we initially use a regressor similar to used by X-12 ARIMA program (2.2.3) with special proportions based on the behavior of the Industrial Turnover. We assume that Easter has a five days global duration, and there are three different periods of time. The first period consists of Holy Thursday, the second runs from Good Friday to Holy Sunday and the last is composed by Easter Monday. Each one has its own weight depending on whether the day is working or non-working in the different Autonomous Communities. The Easter regressor is generated as

$$E_t = \sum_{\delta=1}^{17} w_\delta h_{\delta,t} - \mu_t, \qquad (3.2.10)$$

where $h_{\delta,t}$ is the proportion of Easter days that fall in the moth t and μ_t is the long run monthly means. It is important to point out that Easter Sunday and Saturday would not be taken into account in this proportion because they are already included as non-working days.

3.3 Relationship between different types of calendar adjustment

The form of the regressor used in IPI is very different from the one used in Turnover and New Orders Received –or the TRAMO/SEATS regressor. This difference can make one believe that the adjustment is necessarily very different. In this section, we will compare theoretically a simplified version of the Turnover regressor, that is, the one in 3.2.8, without removing the long-term averages, with the IPI regressor. We will not take into account the details of holidays, but we focus on what are the consequences of subtracting a multiple of the number of non-working days to the number of working days.

When a model is used to adjust the calendar effects we follow two steps:

STEP 1:

We build a model for the variable of interest which includes as an independent variable one that is a function of composition of days of the month. For example:

$$y_t = f(c_t, \varepsilon_t), \tag{3.3.1}$$

where c_t represents the composition in days and ε_t is the random part.

STEP 2:

This step (which is sometimes overlooked in describing the method of adjustment) consists of using the model of the previous step to estimate **the value that would have taken the variable of interest if the composition of the month had been a certain reference one**. This estimate is what we call **adjusted for calendar**.

The result of the adjustment depends therefore on: (i) the model and (ii) the reference composition of days.

We will say that two models are equivalent when applied to the same period of time (or actually, to two periods which months have exactly the same composition of days) and varying its parameters, yield the same family of distributions for the dependent variable.

As an example, suppose that we have the following models:

$$y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \qquad \varepsilon_t \sim \text{ARIMA}(p,d,q;P,D,Q) \tag{3.3.2}$$

$$y_t = \gamma_0 + \gamma_1 Z_t + \varepsilon_t \qquad \varepsilon_t \sim \text{ARIMA}(p,d,q;P,D,Q) \tag{3.3.3}$$

Since the random part is common, the two models are equivalent if for any values β_0 and β_1, there are values of γ_0 and γ_1 such that $\gamma_0 + \gamma_1 Z_t = \beta_0 + \beta_1 X_t$.

The main idea that we need to take into account to understand the relationship between different methods of adjustment is: if the models are equivalent and the reference composition of days used is the same for every t, then the result of the adjustment will be the same.

Let us compare in this terms two ways to make the calendar adjustment.

Model 1:

$$\delta(B)\{y_t - \beta TD_t^1\} = \varepsilon_t \qquad \varepsilon_t \sim \mathrm{ARMA}(p, q; P, Q)$$

where TD_t^1 is the number of working days of the month t.

Model 2:

$$\delta(B)\{y_t - \gamma TD_t^2\} = \varepsilon_t \qquad \varepsilon_t \sim \mathrm{ARMA}(p, q; P, Q)$$

where TD_t^2 is the number of working days of the month t minus 5/2 times the number of non-working days.

It holds that

$$TD_t^2 = TD_t^1\{1 + \frac{5}{2}\} - \frac{5}{2}D_t,$$

where D_t is the total number of working and non-working days of the month. Assuming no leap years, D_t is periodic of period 12. So, if $\delta(B)$ contains a seasonal difference, the model 2 can be written as

$$\delta(B)y_t - \gamma\delta(B)TD_t^2 = \delta(B)y_t - \gamma\{1 + \frac{5}{2}\}\delta(B)TD_t^1 = \varepsilon_t.$$

Therefore, the two models are equivalent and the results of the estimation will be

$$\hat{\beta} = \hat{\gamma}\{1 + \frac{5}{2}\}.$$

Now, to make the adjustment, we have to perform step 2: to set a reference composition of days. We consider the following two options:

(a) Set a unique composition.

(b) Set different compositions for each month.

Adjustment (a):

If we consider a unique type of month and we replace TD_t^2 and TD_t^1 by the value of the reference month, the model 1 changes to

$$y_t = \beta TD_t^1 + \frac{\theta(B)}{\delta(B)\varphi(B)} a_t,$$

replacing TD_t^1 by a constant $\overline{TD^1}$ we obtain

$$\tilde{y}_t^{TD^1} = \hat{\beta}\overline{TD^1} + \frac{\theta(B)}{\delta(B)\varphi(B)} a_t = y_t - \hat{\beta}\big(TD_t^1 - \overline{DL}\big).$$

In a similar fashion to model 2 we obtain

$$\tilde{y}_t^{(2)} = y_t - \hat{\gamma}\big(TD_t^2 - \overline{TD^2}\big).$$

Since we know the relationship between TD^1 and TD^2, we can substitute it and get

$$\tilde{y}_t^{(2)} = y_t - \hat{\gamma}\big(TD_t^2 - \overline{TD^2}\big) = y_t - \hat{\beta}\frac{TD_t^2 - \overline{TD^2}}{1 + 5/2} = y_t - \hat{\beta}\Big\{TD_t^1 - \frac{5}{7}D_t + \frac{2}{7}\overline{TD^2}\Big\}. \quad (3.3.4)$$

Hence, since $\overline{TD^2} \approx 0$,

$$\tilde{y}_t^{(2)} \approx y_t - \hat{\beta}\Big\{TD_t^1 - \frac{5}{7}D_t\Big\}. \quad (3.3.5)$$

Comparing (3.3.4) and (3.3.5), we can observe that the difference is that in the adjustment $\tilde{y}_t^{(2)}$ the raw data are replaced by an estimation of the value that we have get if the month t would have had the average proportion of working days and the total number of days (working and non-working days) that in fact had month t.

By contrast, the adjusted data $\tilde{y}_t^{(1)}$ equals the value that the model would have yielded if the number of working days of month t had been that of the reference month. Consequently, we are also discounting the effect of different lengths of the months. For a logarithmic model, the annual rate would be equal, but the index would have a deterministic seasonal component in $\tilde{y}_t^{(2)}$ that would not be present in $\tilde{y}_t^{(1)}$.

However, if we want to preserve the seasonal component, we can do it using the regressor TD_t^1 if we choose the adjustment (b). Suppose we choose for each month t, a reference composition having the average proportion of working days, but the actual length of the month, D_t

$$\tilde{y}_t^{(1)} = \hat{\beta}\frac{5}{2}D_t + \frac{\theta(B)}{\delta(B)\varphi(B)}a_t = y_t - \hat{\beta}\Big\{DL_t - \frac{5}{2}D_t\Big\} = \tilde{y}_t^{(2)}.$$

Thus, by doing the adjustment with a variable composition of days, we get the same effect both in $\tilde{y}_t^{(2)}$ as $\tilde{y}_t^{(1)}$.

If the leap year is taken into account the relationship between TD^1 and TD^2 can be rewritten as

$$TD_t^2 = TD_t^1\{1 + \frac{5}{2}\} - \frac{5}{2}D_t^* - \frac{5}{2}LY_t, \tag{3.3.6}$$

where D_t^* is the total number of days in the month t including all Februaries with 28 days, in other words, is a periodic function with period 12.

There are different cases according to the inclusion of the variable LY_t or not.

CASE A:

If we include the variable in both models:

$$\begin{aligned} \delta(B)\{y_t - \beta_1 TD_t^1 - \beta_2 LY_t\} &= \varepsilon_t \qquad \varepsilon_t \sim \text{ARMA}(p,q;P,Q) \\ \delta(B)\{y_t - \gamma_1 TD_t^2 - \gamma_2 LY_t\} &= \varepsilon_t \qquad \varepsilon_t \sim \text{ARMA}(p,q;P,Q) \end{aligned} \tag{3.3.7}$$

Since we have the relationship (3.3.6), then

$$\begin{bmatrix} TD_t^2 \\ LY_t \end{bmatrix} = \begin{bmatrix} 1+5/2 & -5/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} TD_t^1 \\ LY_t \end{bmatrix} - \frac{5}{2}\begin{bmatrix} D_t^* \\ 0 \end{bmatrix}.$$

Thus, both models are equivalent if there is a seasonal difference.

CASE B:

If we include the variable LY only in one model, then this one nests the other one.

CASE C:

If we do not include the variable LY in any of the models, then we will have not the same models even when applying a seasonal difference, since

$$\nabla_{12} TD_t^2 = \nabla_{12} TD_t^1 \{1 + \frac{5}{2}\} - \frac{5}{2} \nabla_{12} D_t.$$

To decide which method is more appropriate in this case, we would have to make some assumptions about the true behavior of the series. The analysis of this case remains for future research.

4 ARIMA model-building procedure

In this section we describe the methodology used to build models for the different series of Industrial Turnover and New Orders Received Indices. We describe in 4.1 the class of models we are considering and in 4.2 the modeling procedure, that is an extended form of the methodology originally proposed by Box and Jenkins, stated in its latest edition in Box et al. (1994). The choice of the software used for the identification, estimation and diagnosis of the models is discussed in 4.3.

4.1 Structure of the Model

The variable Z_t can be decomposed in two components, one of them a deterministic component, and other purely stochastic, as

$$Z_t = \delta_t + N_t,$$

where the deterministic component, δ_t, includes the variables of the calendar effects defined in the previous sections as well as the intervention variables that we consider necessary, and the stochastic component N_t is represented by an ARIMA$(p, d, q)(P, D, Q)_s$ model.

The stochastic part could be modeled in different ways. We decided to use the family of ARIMA processes, specifically multiplicative seasonal ARIMA models, because it allows modeling the seasonal dependence (which is associated with observations separated by s periods, where s is the number of observations per year), modeling the regular dependence (which is associated with consecutive observations) and treating non-stationarity in half. The ARIMA$(p,d,q)(P,D,Q)_s$ representation of the purely stochastic component is a generalization of the ARIMA(p,d,q) models and is given by the expression

$$\begin{aligned} w_t &= \nabla^d \nabla_s^D N_t \\ \Phi_P(B^s)\phi_p(B)w_t &= \Theta_Q(B^s)\theta_q(B)a_t, \end{aligned} \tag{4.1.1}$$

where $\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}$ is the annual autoregressive polynomial $(AR(P)_s)$, $\Theta_Q(B^s) = 1 + \Theta_1 B^s + \ldots + \Theta_Q B^{Qs}$ is the annual moving average polynomial $(MA(Q)_s)$, $\phi_p = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ is the autoregressive polynomial $(AR(p))$, $\theta_q = 1 + \theta_1 B + \ldots + \theta_q B^q$ is the moving average polynomial $(MA(q))$, w_t is a stationary variable and P, D, Q, p, d, q are non-negative integers.

4.2 Modeling procedure

As stated before, the search for an ARIMA model compatible with the data, is done in a sequential way using the Box-Jenkins methodology that can be described in three main phases:

1. Model identification and model selection.
2. Parameters estimation.
3. Statistical diagnosis of the estimated models (formal and informal methods). If the estimation is inadequate, we have to return to step one and attempt to build a better model.

Following this methodology, the first step must be to determine if the time series is stationary and, otherwise determine the transformation of the variable and the values of d, D that make the transformed and differenced series stationary.

Before modeling the time series data, a power transformation may be used, such as the Box-Cox Transformation, that is a particular way of parameterizing a power transform that has advantageous properties. Let Y_t the value of the original time series in t,

$$Z_t = Y_t(\lambda, \alpha) = \begin{cases} \frac{(Y_t+\alpha)^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log Y_t & \text{if } \lambda = 0 \end{cases} .$$

Thus, first of all, we need to determine the value of the parameter λ of the Box-Cox transformation that finally we will apply to the original series.

In our application, the choice of λ is limited to the cases $\lambda = 0$, that is, to apply the natural logarithm or $\lambda = 1$, that is, no transformation. The tools used in this choice are, on the one hand, the time series plot of the original and transformed series and, on the other hand, the range-mean graph. To apply the natural logarithm is justified when the graph of the original series shows that the local standard deviation of the data increases with increasing local mean or when the range-mean graph shows a positive and approximately linear relationship between local mean and local standard deviation. If the log transformation is appropriate, it is expected that the local deviation of the transformed series is stable. To apply the logarithm is also justified to induce linearity when the original data graph has an exponential form.

Second, we need to determine the value of d. It can be assessed from a run sequence plot. We plot the original time series ($d = 0$) and observe his parameters of location. If they are not constant or the series wanders around an average, then we apply a regular difference to the series and we asses the plot of the differenced series ($d = 1$). Again we check if the location parameters are constant in the differenced series and so on until the plot shows constant parameters. Then we conclude that the d times differenced series is stationary. Another way to detect non-stationarity is through the autocorrelation function. If we observe a slow decay of the coefficients it is necessary to apply a difference to the series. The last method is the most widely used to detect seasonal non-stationarity, so if the seasonal coefficients decrease slowly we apply a seasonal difference. The final number of seasonal differences taken is D. Seasonality can also be detected when the observations of the same month are systematically above (below) the overall sample average.

Also, these graphics can help to detect abnormal incidents that can be treated with intervention analysis, considering that can distortion specification tools. Abnormal incidents are depicted with intervention analysis, according to Box and Tiao (1975). In this article two types of interventions are used: a pulse function

$$P_t = \begin{cases} 1 & \text{if } t = t_0 \\ 0 & \text{if } t \neq t_0 \end{cases},$$

that indicates that the intervention only occurs in the single time index t_0, and a step function

$$S_t = \begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases},$$

that shows that the intervention continues to exist starting with the time index t_0.

At the model identification stage, we need to detect seasonality, if it exists, and identify the order of the seasonal autoregressive and seasonal moving average terms. The primary tools for doing this are the autocorrelation function plot (ACF) and the partial autocorrelation function plot (PACF). We determine the orders P, Q through the observed structure in the seasonal lags of the ACF and PACF of the transformed series, which is stationary, and their comparison with the theoretical behavior of these plots when the order is known.

Once stationarity and seasonality have been addressed, the next step is to identify the order p, q of the autoregressive and moving average terms respectively. The identification process is similar to the used for seasonality but in this case we observe the structure in the early delays of the ACF y PACF.

In practice, the sample autocorrelation and partial autocorrelation functions are random variables and will not give the same picture as the theoretical functions. This makes model identification more difficult. In addition, many times there are more than one model that seem to fit the series. For this reason we use the Bayesian information criterion (BIC), a criterion for model selection among a class of parametric models with different number of parameters, to choose between the different models for the time series.

For the parameters estimation stage, we use maximum likelihood estimation. In the model diagnostic phase, we test whether the estimated model conforms to the specifications. Specifically, the error term a_t is assumed to follow the assumptions for a stationary univariate process. So, the residuals should be white noise (or independent when their distributions are normal) drawings from a fixed distribution with constant mean and variance. If these assumptions are not satisfied, we need to fit a more appropriate model. That is, go back to the model identification step and try to develop a better model. One way to assess if the residuals from the model follow the assumptions is to generate statistical graphics (including a time series and autocorrelation function plots). We can also look at the value of the Ljung-Box statistic.

4.3 Programs used

There are many programs that allow to estimate regARIMA models (R, Gretl, SAS, Matlab...). We tested many of them and obtained that two or more programs returning different results, it is enough reason to make a pause and think which, if any, of the results is the correct. After checking some programs, because it is impossible to test all of the procedures offered, we finally decided to use Gretl (GNU Regression, Econometrics, and Time-series Library) as main program. It is a cross-platform and open-source econometrics package that is freely available. To take this decision we have evaluated the advantages and disadvantages of this program over other.

As advantages we have:

- its simplicity, with an easy and intuitive interface, and accessibility,
- thanks to the availability of the source code, users are allowed to peek inside the program in order to have insight into how it works. The possibility of peer review also makes it possible for errors within the program to be noticed and fixed quickly, resulting on a very high quality software,
- integrated scripting language: enter commands either via the gui or via script,
- GUI controller for fine-tuning Gnuplot graphs,
- the accuracy of the estimations is high. The program proves to be as good or even better in terms of numerical precision compared to other.

And as disadvantages:

- the possibility to correct program errors quickly can cause the user is using an outdated and obsolete version of the program,
- the program currently does not offer many tools of analysis as some of the widely used alternatives. For example it does not allow to estimate factored models.

5 Analysis of the Industrial Turnover series

In the Industrial Turnover the indices are calculated for all basic aggregates (two digits or set of groups of three digits) and another functional aggregates like other divisions that are not basic aggregates (two digits), sections B and C and general. We calculate also the index for economic sectors by economic destination (durable consumer goods, non-durable consumer goods, capital goods, intermediate goods and energy). The following tables show the composition of the different basic and functional aggregates.

Basic aggregates	
-Extraction of anthracite, coal and lignite	05
-Extraction of crude petroleum and natural gas	06
-Other mining and quarrying	08
-Food industries (except grain mill products and animal feeds)	10.1+10.2+10.3+10.4 +10.5+10.7+10.8
-Manufacture of grain mill products, starches and starch products, and of animal feeds	10.6+10.9
-Manufacture of beverages	11
-Manufacture of tobacco	12
-Preparation and spinning of textile fibres. Manufacture of woven textiles. Textile finishing	13.1+13.2+13.3
-Manufacture of knitted fabrics, carpets, rope, non-woven fabrics, textile products for technical and industrial use and other textile products	13.9
-Manufacture of garments	14
-Leather and footwear industry	15
-Wood and cork industry; except furniture, basketmaking and wickerwork	16
-Paper industry	17
-Graphic arts and reproduction media	18
-Manufacture of coke and refined petroleum products	19
-Manufacture of cleaning articles, perfumes and cosmetics	20.4

Basic aggregates	
-Chemical industry except cleaning articles, perfumes and cosmetics	20.1+20.2+20.3+20.5 +20.6
-Manufacture of pharmaceutical products	21
-Manufacture of rubber and plastic products	22
-Manufacture of other non-metallic ore products	23
-Metallurgy, manufacture of iron, steel and ferro-alloy products	24
-Manufacture of metal elements for construction, containers made of metal, steam generators, weapons and ammunition	25.1+25.2+25.3+ 25.4
-Forging, stamping, embossing and rolling of metals. Manufacture of tools, hardware goods, containers and other metal products	25.5+25.6+25.7+25.9
-Manufacture of electronic components, assembled printed circuits, and magnetic and optical media	26.1+26.8
-Manufacture of computers, peripherals and telecommunications equipment; appliances for measuring and navigation; radiation and electro-medical equipment	26.2+26.3+26.5+ 26.6
-Manufacture of consumer electronics, optical instruments and photographic equipment	26.4+26.7
-Manufacture of household appliances	27.5
-Manufacture of machinery and equipment, n.e.c.	28
-Manufacture of motor vehicles, trailers and semi-trailers	29
-Naval, railway, aircraft and spacecraft construction. Combat vehicles	30.1+30.2+30.3+ 30.4
-Manufacture of motorcycles, bicycles, vehicles for the disabled and others n.e.c.	30.9
-Manufacture of furniture	31
-Manufacture of jewelery, costume jewelery and musical instruments	32.1+32.2
-Manufacture of sporting goods; games and toys. Other manufacturing industries	32.3+32.4+32.9

Basic aggregates	
-Manufacture of medical and dental instruments and supplies	32.5
-Repair and installation of machinery and equipment	33

Other divisions	
-Food Industry	10
-Textile Industry	13
-Manufacture of fabricated metal products, except machinery and equipment	25
-Manufacture of computer, electronic and optical products	26
-Manufacture of electrical equipment	27
-Manufacture of other transport equipment	30
-Other manufacturing	32

Sections	
-Mining and quarrying	05+06+08
-Manufactured products	from 10 to 33

Economic destination	
-Consumer goods	
-Durable consumer goods	26C+27A+30B+31+32A
-Non-durable consumer goods	10A+11+12+13B+14+15+18+20A+21+32B
-Capital goods	25A+26B+28+29+30A+32C+33
-Intermediate products	08+10B+13A+16+17+20B+22+23+24+25A+26A+27B
-Energy	05+06+19

There are some details that need to be clarified before we continue with the analysis of the series. One is the decision criteria to remove variables from the model. In this context, when the coefficient of one variable has a p-value higher than 0.50, in other words, the variable is not significative, or the sign of the coefficient is opposite to that expected, this variable is deleted from the model. A variable is also deleted when its p-value is in the range (0.20;0.50] and the size of the estimated coefficient differs greatly from the value usually taken.

Another aspect to consider is the frequency of re-estimating the model coefficients and the model itself. The final decision is to re-estimate the coefficients each year, and re-identify the model each time that we change the base year. If there are not evidence to change the estimated parameters or the model then we hold them until the next revision.

The last question is whether direct analysis (where all time series, including the aggregates, are adjusted individually) is better than indirect (where the adjusted series of the aggregates are obtained adding the adjusted series of their components). There are favorable opinions in the two ways, but no empirical and neither theoretical evidence that one approach is better than the other in all cases. However, there is agreement that when all series have a similar seasonal component the direct adjustment is better, opposite to where each component has a different seasonality, in which case the indirect adjustment is better. We decide to use direct adjustment because it is more transparent.

In this sense, we have tested how the corrected series change when we use a indirect method. The obtained results show that the indirect corrected series do not present seasonality remains and the differences between the two types of adjustment series are insignificant.

In the analysis of the series we use a single regressor for the trading-day effect that is result of considering the working days and holidays for all the activities as a single labor calendar.

The possibility of considering different regressors for the trading-day effect depending on the elementary aggregate or, in other words, weighting the holidays according to the importance of each Autonomous Community on the total turnover of the aggregate the sector. However, we do not obtain significant differences between the models, not even between the estimated coefficients and variances from the adjustment with regressors for activity versus a single regressor.

In the next subsection (5.1) we describe the different models identified for the basic aggregates, doing a more exhaustive analysis of the series 08 (Other mining and quarrying). In subsections 5.2 and 5.3 we show the final models chosen for other divisions and another functional aggregates by economic destination. Finally in subsections 5.4 and 5.5 we describe the models determined for sections B and C and for the General Index. Monthly data for the period 2002:01-2011:01 are used to specify and estimate the models to the Industrial Turnover Index. However for the analysis of the New Orders Received Index we use monthly data for the period 2002:01-2011:07. The reason for using a different sampling period in each index derived from the different dissemination calendar.

5.1 Analysis of the Industrial Turnover Index for basic aggregates

05 - Mining of coal and lignite

To this basic aggregate the model that we obtain is the following:

Model 5.1.1

$$\log BA_t^{05} = \underset{(.003)}{.005}TD_t - \underset{(.043)}{.024}EE_t - \underset{(.072)}{.135}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.133)}{.440})(1 - \underset{(.173)}{.558}B^{12})a_t$$

$$\hat{\sigma}_a = 12.89\%$$

where TD_t is the trading-day variable, EE_t and LY_t are the Easter and Leap year variables.

In the case of Mining of coal and lignite, the coefficient of Easter is not significant (p-value>0.50), and the leap year is significant at 5% but not at 7%, being the coefficient negative when it would be logical to be positive. So we removed these variables from the model.

08 - Other mining and quarrying

We will perform a complete analysis of the data series to the basic aggregate 08 (Other mining and quarrying), from now on BA^{08}. First we observe the time series plot of the original series (figure 1) and we analyze the range-mean graph (figure 2) to determine if the local deviation of the data increase with increasing local mean.

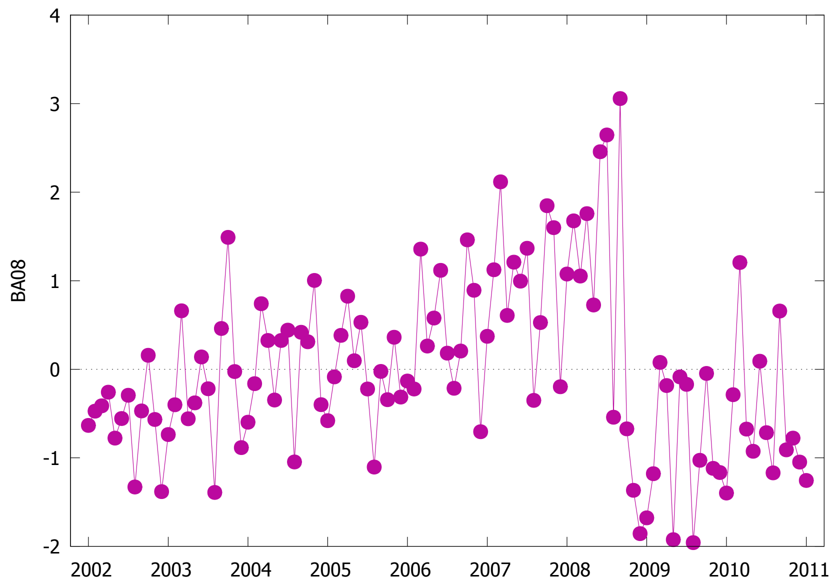
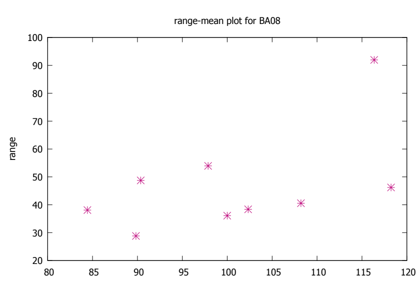
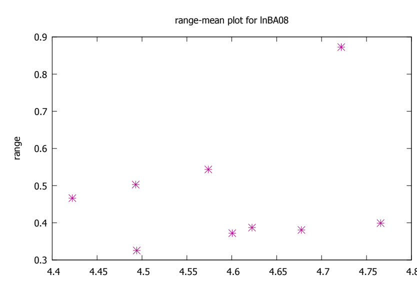


Figure 1: Time series plot of Other mining and quarrying



a) Range-mean plot of Other mining and quarrying series

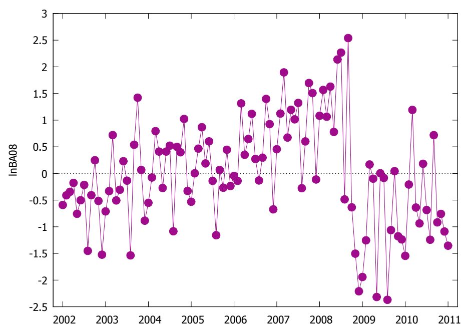


b) Range-mean plot of Other mining and quarrying's logarithm

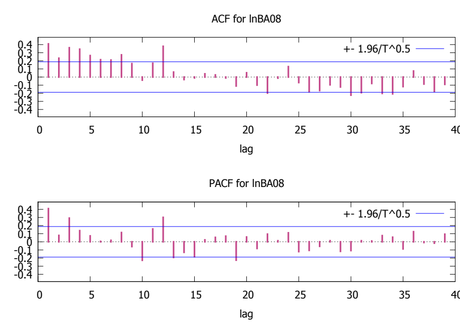
Figure 2: Range-mean plots

Although there is not a clear relationship between the local average and local deviation in the BA^{08} graphs, the slope of the average against the deviation has a p-value that leads us to reject the null hypothesis (H_0 : slope = 0) for levels of significance above 0,10. However the p-value of H_0 for the logarithm is much greater, approximately 0,5. So we decide to apply the natural logarithm to the series.

Concerning to the stationarity of the series, the plot of the series in logs (figure 3) clearly shows a trend and therefore lack of stationary. If we also look at the ACF of the series, we see a high value on the delays which don't decay exponentially. So we need to take differences of the log series. Moreover, if we analyze the time series plot of ∇BA^{08} (figure 4) we observe systematic behavior in some month that are indicative of seasonality together with the no presence of decay in the seasonal coefficients of its simple autocorrelation function. Therefore, we finally work with $\nabla\nabla_{12} \log BA^{08}$.

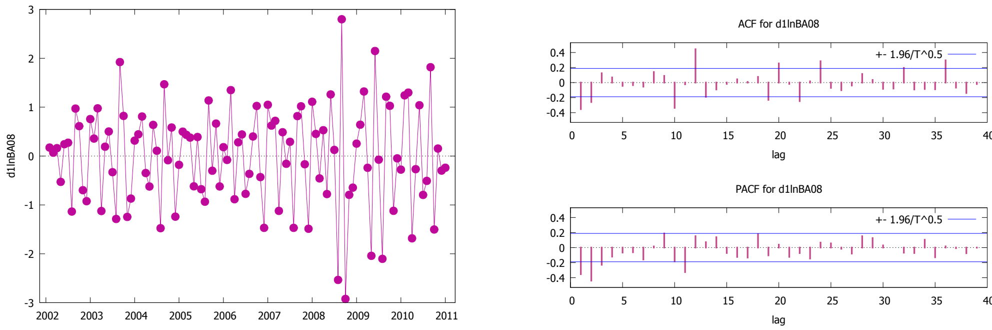


a) Time series plot of the logarithm of other mining and quarrying



b) ACF and PACF of other mining and quarrying logs

Figure 3: Graph of other mining and quarrying logs



a) Time series plot of $\nabla lnBA^{08}$

b) ACF and PACF of $\nabla \log BA^{08}$

Figure 4: Graph of $\nabla \log BA^{08}$

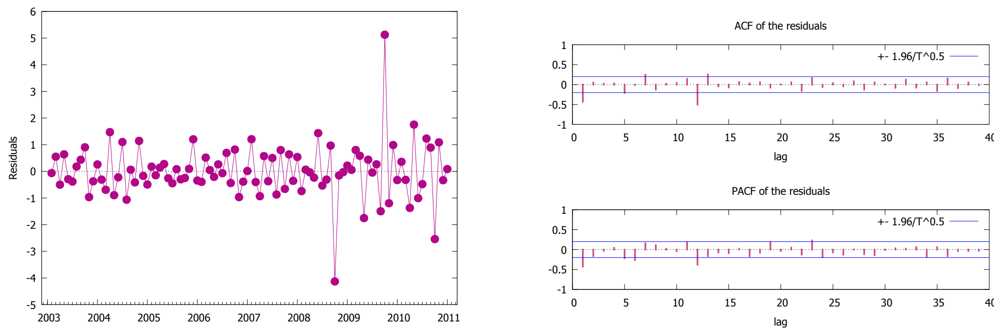
Now that we have a stationary series, we pass to the model identification phase. First we carried out a regression of the series on the calendar effect variables and observe the residuals to appreciate more clearly the structure of the stochastic part. The result of the regression is:

Model 5.1.2

$$\log BA_t^{08} = \underset{(.004)}{.018}TD_t - \underset{(.033)}{.103}EE_t - \underset{(.061)}{.053}LY_t + N_t$$

$$\nabla\nabla_{12} N_t = a_t$$

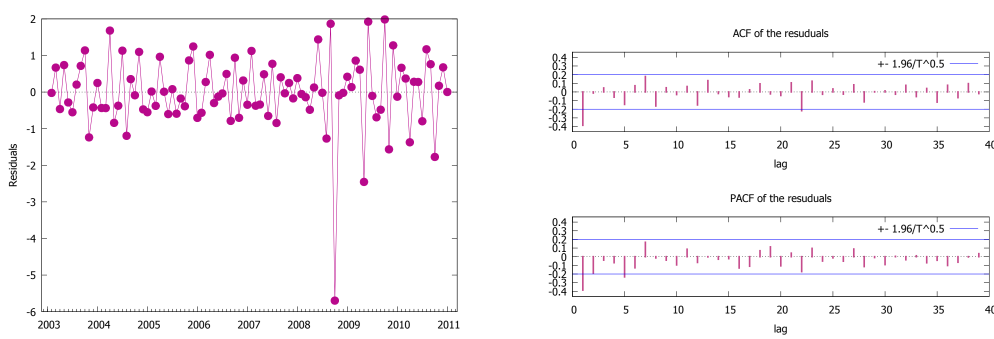
$$\hat{\sigma}_a = 17.16\%$$



a) Time series plot of the residuals b) ACF and PACF of the residuals

Figure 5: Graphics of the residuals of the model 5.1.2

Looking at the graphics of the regression residuals (figure 5) we see that the first seasonal lag of the ACF and PACF have a very high value, while the remaining delays are not significant. So, surely we have a structure of order 1. Since second seasonal delay is not positive in the ACF, and the seasonal delays of the PACF decay quickly, we believe that a moving average model is better than the autoregressive to model the seasonality. Comparing the BIC for these models, we find that $MA(1)$ model is better.



a) Time series plot of the residuals b) ACF and PACF of the residuals

Figure 6: Graphics of the residuals of the model with seasonal structure

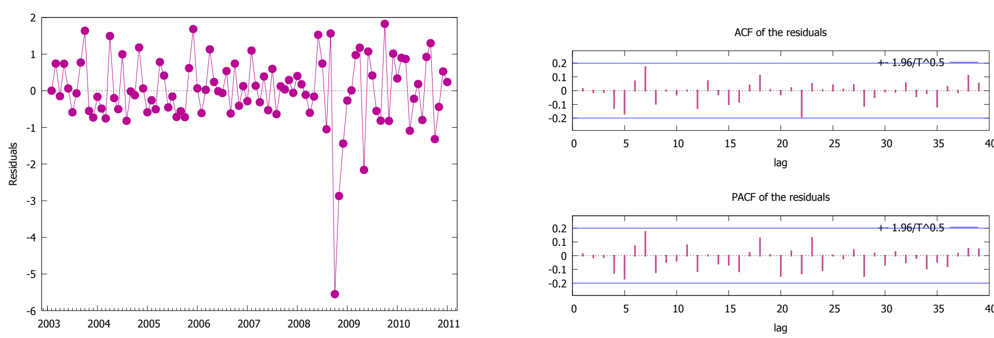
Once decided the model for seasonality, we again regress the series introducing this stochastic structure and observe the new residuals (figure 6) to determine the model of the regular part. In the graphs, we can observe that only the first and seventh coefficients of the ACF are significant. However, the higher value of the seventh delay is due to the distortion introduced by one pair of data, the extreme value in October 2008 with March 2009, and we evaluate its relevance later. So we have a first order regular structure, in particular, a moving average. The complete model is

Model 5.1.3

$$\log BA_t^{08} = \underset{(.002)}{.015} TD_t - \underset{(.043)}{.121} EE_t + \underset{(.071)}{.051} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.098)}{.498})(1 - \underset{(.443)}{.945} B^{12}) a_t$$

$$\hat{\sigma}_a = 10.57\%$$



a)Time series plot of the residuals

b) ACF and PACF of the residuals

Figure 7: Graphics of the residuals to the model 5.1.3

If we observe the time series plot of the original series, we will see a negative permanent change in the level of the series from October 2008, in other words, a step on that date. This effect also can be clearly observed in the graphs of the differenced series, appearing as a pulse in 08-2010. Looking at the plot of the residuals of the last model (figure 7) we still observe an extreme value (-5.55σ) in the date of the incident which is not picked up by the model. We need to add an intervention analysis to find out if the existence of this outliers distorts the estimated coefficients of the model. So we do a new regression including a step variable with $t_0 = 10/2008$ that we call $\xi^{S10:2008}$ and we compare the results with those obtained in the absence of intervention.

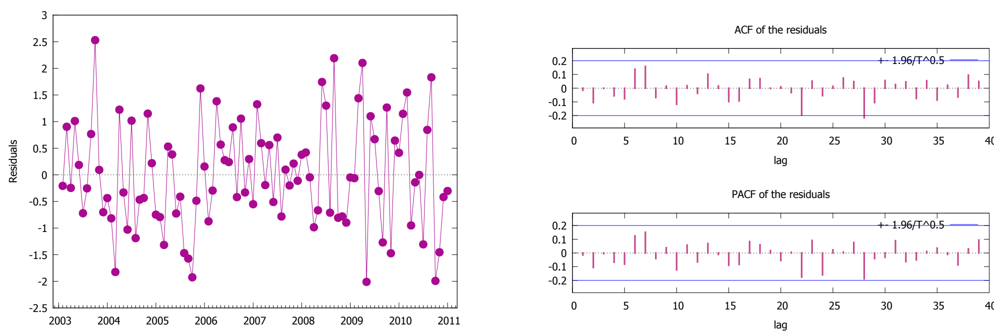
Model 5.1.4

$$\log BA_t^{08} = \underset{(.002)}{.017}TD_t - \underset{(.035)}{.117}EE_t + \underset{(.054)}{.033}LY_t - \underset{(.047)}{.516}\xi^{S10:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.063)}{.873}B)(1 - \underset{(.190)}{.837}B^{12})a_t$$

$$\hat{\sigma}_a = 7.50\%$$

The coefficient of $\xi^{S10:2008}$ is significant and shows the effect discussed by its sign and size. The coefficients of the calendar effects don't change significantly, however, the coefficient of regular and seasonal stochastic structure change notably. The coefficients changes from $\Theta = .945$ to $\Theta = 0.832$ and from $\phi = .498$ to $\phi = .873$. So we finally decide to include the step in the model. Instead, the leap year variable is removed because its p-value is very high.



a) Time series plot of the residuals

b) ACF and PACF of the residuals

Figure 8: Graphics of the residuals to the model 6.1.2

As we can see in the plots of the residuals of the model with intervention analysis (figure 8), no structure remains in the autocorrelation functions and the residuals present a constant average and deviation. So the residuals probably follow a stationary univariate process. But we should check this with a contrast. In particular, we analyze the Ljung-Box statistic. The following table (Table 2) shows the values of the Q-statistic, that, as we can observe, have a p-value high enough to no refuse the null hypothesis of White Noise.

Table 2: Ljung-Box statistic

LAG	ACF	PACF	Q-stat.	p-value
01	-0.0151	-0.0151	0.0225	[0.881]
02	-0.1053	-0.1055	1.1322	[0.568]
03	-0.0020	-0.0054	1.1326	[0.769]
04	-0.0557	-0.0677	1.4495	[0.836]
05	-0.0774	-0.0817	2.0696	[0.839]
06	0.1384	0.1246	4.0721	[0.667]
07	0.1587	0.1508	6.7346	[0.457]
08	-0.0680	-0.0405	7.2293	[0.512]
09	0.0148	0.0381	7.2532	[0.611]
10	-0.1175	-0.1230	8.7632	[0.555]
11	0.0182	0.0567	8.7997	[0.640]
12	-0.0375	-0.0651	8.9575	[0.707]
13	0.1015	0.0683	10.1243	[0.684]
14	0.0162	-0.0108	10.1542	[0.751]
15	-0.0979	-0.0899	11.2681	[0.733]
16	-0.0933	-0.0835	12.2909	[0.724]
17	0.0645	0.0834	12.7869	[0.750]
18	0.0695	0.0599	13.3696	[0.769]
19	-0.0021	0.0168	13.3702	[0.819]
20	0.0101	-0.0555	13.3829	[0.860]
21	-0.0318	0.0050	13.5094	[0.890]
22	-0.1957 *	-0.1755 *	18.3809	[0.683]
23	0.0514	0.0912	18.7212	[0.717]
24	-0.0543	-0.1599	19.1069	[0.746]
25	0.0132	0.0228	19.1300	[0.791]
26	0.0739	0.0058	19.8636	[0.798]
27	0.0483	0.0754	20.1810	[0.823]
28	-0.2149 **	-0.1881 *	26.5725	[0.542]
29	-0.1050	-0.0408	28.1191	[0.512]
30	0.0556	-0.0322	28.5597	[0.541]

LAG	ACF	PACF	Q-stat.	p-value
31	0.0269	0.0894	28.6643	[0.587]
32	0.0458	-0.0629	28.9729	[0.621]
33	-0.0750	-0.0506	29.8121	[0.627]
34	0.0545	0.0109	30.2625	[0.651]
35	-0.0872	0.0354	31.4341	[0.641]
36	0.0210	-0.0107	31.5032	[0.682]
37	-0.0642	-0.0877	32.1599	[0.695]
38	0.0952	0.0297	33.6288	[0.672]
39	0.0492	0.0940	34.0276	[0.696]

10A - Food industries (except grain mill products and animal feeds)

To the basic aggregate 10A the model that we obtain is the following:

Model 5.1.5

$$\log BA_t^{10A} = \underset{(.001)}{.010}TD_t - \underset{(.008)}{.047}EE_t + \underset{(.014)}{.034}LY_t + N_t$$

$$(1 + \underset{(.104)}{.429}B + \underset{(.096)}{.345}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.138)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 2.10\%$$

All coefficients are significant, so in this case none of them is removed from the equation.

10B - Manufacture of grain mill products, starches and starch products, and of animal feeds

Model 5.1.6

$$\log BA_t^{10B} = \underset{(.001)}{.009}TD_t - \underset{(.008)}{.036}EE_t + \underset{(.013)}{.026}LY_t - \underset{(.022)}{.051}\xi^{S10:08} + N_t$$

$$\nabla\nabla_{12}N_t = ((1 - \underset{(.097)}{.186}B + \underset{(.095)}{.097}B^2 + \underset{(.112)}{.444}B^3)(1 - \underset{(.187)}{.589}B^{12} - \underset{(.135)}{.411}B^{24})a_t$$

$$\hat{\sigma}_a = 2.41\%$$

11 - Manufacture of beverages

Model 5.1.7

$$\log BA_t^{11} = \underset{(.001)}{.011}TD_t - \underset{(.015)}{.050}EE_t + \underset{(.024)}{.101}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.093)}{.693}B)(1 - \underset{(.143)}{.403}B^{12})a_t$$

$$\hat{\sigma}_a = 4.30\%$$

12 - Manufacture of tobacco products

Model 5.1.8

$$\log BA_t^{12} = \underset{(.003)}{.020}TD_t - \underset{(.033)}{.140}EE_t + \underset{(.062)}{.027}LY_t + N_t$$

$$(1 + \underset{(.108)}{.256}B - \underset{(.107)}{.070}B^2 + \underset{(.106)}{.106}B^3 + \underset{(.104)}{.205}B^4)\nabla\nabla_{12}N_t = (1 - \underset{(.112)}{.379}B^{12})a_t$$

$$\hat{\sigma}_a = 12.50\%$$

We remove the LY_t coefficient from the model, because its p-value is greater than 0,50.

13A - Preparation and spinning of textile fibres. Manufacture of woven textiles. Textile finishings

Model 5.1.9

$$\log BA_t^{13A} = \underset{(.001)}{.013}TD_t - \underset{(.014)}{.111}EE_t + \underset{(.021)}{.020}LY_t + N_t$$

$$(1 + \underset{(.104)}{.293}B)\nabla\nabla_{12}N_t = (1 - \underset{(.215)}{.509}B^{12} - \underset{(.151)}{.491}B^{24})a_t$$

$$\hat{\sigma}_a = 4.10\%$$

13B - Manufacture of knitted and crocheted articles, rugs, cordage, non-woven fabrics, textile products for technical and industrial use and other textile products

Model 5.1.10

$$\log BA_t^{13B} = \underset{(.001)}{.012}TD_t - \underset{(.014)}{.103}EE_t + \underset{(.024)}{.049}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.103)}{.478}B + \underset{(.103)}{.203}B^2)(1 - \underset{(.112)}{.285}B^{12})a_t$$

$$\hat{\sigma}_a = 4.99\%$$

14 - Manufacture of wearing apparel

Model 5.1.11

$$\log BA_t^{14} = \underset{(.002)}{.006}TD_t - \underset{(.019)}{.046}EE_t + \underset{(.036)}{.053}LY_t + N_t$$

$$(1 + \underset{(.106)}{.615}B + \underset{(.115)}{.469}B^2 + \underset{(.116)}{.509}B^3 + \underset{(.118)}{.487}B^4 + \underset{(.110)}{.273}B^5)\nabla\nabla_{12}N_t = (1 - \underset{(.111)}{.285}B^{12})a_t$$

$$\hat{\sigma}_a = 5.95\%$$

In determining the seasonal structure we choose to work with $MA(1)_{12}$ existing a model $AR(1)_{12}$ that leaves the ACF and PACF clean seasonal structure and with a similar

BIC. This decision is based on the widespread use of polynomials of first order moving average representation for the seasonal in time series models. The regular part could be represented in a more parsimonious way through a factored model. This seems to be due to the presence of four distinct annual stage in the series.

15 - Manufacture of leather and related products

Model 5.1.12

$$\log BA_t^{15} = \underset{(.001)}{.011}TD_t - \underset{(.018)}{.141}EE_t + \underset{(.031)}{.040}LY_t + N_t$$

$$(1 + \underset{(.100)}{.402}B + \underset{(.109)}{.311}B^2 + \underset{(.112)}{.332}B^3 + \underset{(.119)}{.237}B^4 - \underset{(.109)}{.027}B^5 - \underset{(.101)}{.252}B^6)\nabla\nabla_{12}N_t = (1 - \underset{(.107)}{.525B^{12}})\, a_t$$

$$\hat{\sigma}_a = 5.97\%$$

16 - Manufacture of wood and of products of wood and cork

Model 5.1.13

$$\log BA_t^{16} = \underset{(.001)}{.013}TD_t - \underset{(.012)}{.112}EE_t + \underset{(.020)}{.034}LY_t + N_t$$

$$(1 + \underset{(.104)}{.237}B + \underset{(.101)}{.083}B^2 - \underset{(.102)}{.307}B^3)\nabla\nabla_{12}N_t = (1 - \underset{(.236)}{.612}B^{12} - \underset{(.137)}{.388}B^{24})a_t$$

$$\hat{\sigma}_a = 3.92\%$$

17 - Manufacture of paper and paper products

Model 5.1.14

$$\log BA_t^{17} = \underset{(.001)}{.010}TD_t - \underset{(.009)}{.080}EE_t + \underset{(.016)}{.028}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.089)}{.210}B)(1 - \underset{(.129)}{.758B^{12}})a_t$$

$$\hat{\sigma}_a = 2.78\%$$

18 - Printing and reproduction of recorded media

Model 5.1.15

$$\log BA_t^{18} = \underset{(.001)}{.007}TD_t - \underset{(.013)}{.081}EE_t + \underset{(.021)}{.004}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.072)}{.560}B)(1 - \underset{(.116)}{.436}B^{12})a_t$$

$$\hat{\sigma}_a = 3.87\%$$

We remove the LY_t coefficient from the model because it is not significative.

19 - Manufacture of coke and refined petroleum products

Model 5.1.16

$$\log BA_t^{19} = \underset{(.001)}{.002}TD_t - \underset{(.023)}{.042}EE_t - \underset{(.041)}{.015}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.117)}{1.040}B^{12} + \underset{(.161)}{.270}B^{24})a_t$$

$$\hat{\sigma}_a = 7.40\%$$

In this case, we no observe regular structure. To the seasonal part, a $MA(1)_{12}$ model presents a better BIC, but its ACF and PACF show higher values for some delays. Again, the variable LY_t is removed.

20A - Manufacture of cleaning articles, perfumes and cosmetics

Model 5.1.17

$$\log BA_t^{20A} = \underset{(.001)}{.011}TD_t - \underset{(.015)}{.087}EE_t + \underset{(.027)}{.028}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.106)}{.461}B - \underset{(.106)}{.205}B^2)(1 - \underset{(.107)}{.562}B^{12})a_t$$

$$\hat{\sigma}_a = 4.53\%$$

20B - Chemical industry except cleaning articles, perfumes and cosmetics

Model 5.1.18

$$\log BA_t^{20B} = \underset{(.001)}{.008}TD_t - \underset{(.009)}{.090}EE_t - \underset{(.016)}{.005}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 + \underset{(.095)}{.254}B)(1 - \underset{(.110)}{.700}B^{12})a_t$$

$$\hat{\sigma}_a = 3.83\%$$

Only the coefficient of leap year variable is not significative.

21 - Manufacture of basic pharmaceutical products and pharmaceutical preparations

Model 5.1.19

$$\log BA_t^{21} = \underset{(.001)}{.011}TD_t - \underset{(.019)}{.055}EE_t + \underset{(.027)}{.018}LY_t + N_t$$

$$(1 + \underset{(.116)}{.430}B^{12} + \underset{(.118)}{.336}B^{24})\nabla\nabla_{12}N_t = (1 - \underset{(.071)}{.671}B)a_t$$

$$\hat{\sigma}_a = 4.55\%$$

Based on the series discounting the calendar effects we don't observe the existence of a clear seasonal structure, however, once introduced the regular structure we observed, first, that the coefficients of seasonal delays take more weight in the ACF and PACF and, secondly, the introduction of a moving average structure produces considerable reductions in the variance and focused on the residual series. So we finally decide to include this structure to the model. Also, as in the previous model, the variable of Leap year is removed from de model.

22 - Manufacture of rubber and plastic products

Model 5.1.20

$$\log BA_t^{22} = \underset{(.001)}{.009} TD_t - \underset{(.012)}{.105} EE_t + \underset{(.017)}{.019} LY_t + N_t$$

$$(1 + \underset{(.092)}{.523} B^{12} + \underset{(.103)}{.496} B^{24}) \nabla \nabla_{12} N_t = a_t$$

$$\hat{\sigma}_a = 3.66\%$$

23 - Manufacture of other non-metallic mineral products

Model 5.1.21

$$\log BA_t^{23} = \underset{(.001)}{.013} TD_t - \underset{(.009)}{.120} EE_t + \underset{(.015)}{.014} LY_t + N_t$$

$$\nabla \nabla_{12} N_t = (1 - \underset{(.325)}{.598} B^{12} - \underset{(.158)}{.402} B^{24}) a_t$$

$$\hat{\sigma}_a = 3.22\%$$

In this basic aggregate only some delays in the ACF and PACF are too high and the remaining are not significant. The high value of these delays can be explained by a reduced pair of values that introduce distortion. So if we remove these distortion no regular structure is observed. Part of these distortion disappears if we introduce an intervention analysis in the model to capture the effect of a step that we see in the original series. However, we decide not to introduce the intervention because it does not change any coefficient of the model.

24 - Manufacture of basic metals

Model 5.1.22

$$\log BA_t^{24} = \underset{(.001)}{.010} TD_t - \underset{(.017)}{.106} EE_t - \underset{(.032)}{.028} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.199)}{1.000} B^{12}) a_t$$

$$\hat{\sigma}_a = 5.62\%$$

The leap year variable is again not significative.

25A - Manufacture of metal products for construction, containers made of metal, steam generators, weapons and ammunition

Model 5.1.23

$$\log BA_t^{25A} = \underset{(.002)}{.012} TD_t - \underset{(.017)}{.077} EE_t + \underset{(.029)}{.018} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.099)}{.388} B - \underset{(.097)}{.127} B^2 + \underset{(.099)}{.374} B^3)(1 - \underset{(.111)}{.548} B^{12}) a_t$$

$$\hat{\sigma}_a = 5.69\%$$

There is an extreme value of large magnitude (about -5σ), but we do not enter a step in model since the insert does not change significantly the estimated coefficients of the model. As in the recent models, the variable LY_t must be removed from the model.

25B - Forging, stamping, embossing and rolling of metals. Manufacture of tools, hardware goods, containers and other metal products

Model 5.1.24

$$\log BA_t^{25B} = \underset{(.001)}{.012}TD_t - \underset{(.017)}{.115}EE_t + \underset{(.020)}{.035}LY_t + N_t$$

$$(1 + \underset{(.090)}{.526}B^{12} + \underset{(.094)}{.547}B^{24})\nabla\nabla_{12}N_t = (1 - \underset{(.101)}{.181}B + \underset{(.097)}{.259}B^2)a_t$$

$$\hat{\sigma}_a = 4.23\%$$

There are extreme values that distort the ACF and PACF, so that the coefficient of the fourth delay of the same function is increased by the effect of the pair of values [August-2009; December-2009]. Although the model shows the existence of a possible step, no interventions are added to the model because again the estimated coefficients do not change significantly.

26A - Manufacture of electronic components, assembled printed circuits, and magnetic and optical media

Model 5.1.25

$$\log BA_t^{26A} = -\underset{(.003)}{.005}TD_t - \underset{(.043)}{.090}EE_t + \underset{(.078)}{.007}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.137)}{.731}B^{12})a_t$$

$$\hat{\sigma}_a = 15.67\%$$

In this case, both the leap year and trading-day coefficients are not significative, and must be removed.

26B - Manufacture of computers, peripherals and telecommunications equipment; appliances for measuring and navigation; radiation and medical and therapeutic equipment

Model 5.1.26

$$\log BA_t^{26B} = \underset{(.002)}{.003}TD_t - \underset{(.033)}{.044}EE_t + \underset{(.054)}{.008}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.071)}{.635}B)(1 - \underset{(.138)}{.637}B^{12})a_t$$

$$\hat{\sigma}_a = 8.87\%$$

All of the calendar variables are not significative, however only the leap year variable is removed from the model, following the guidelines previously set.

26C - Manufacture of consumer electronics, optical instruments and photographic equipment

Model 5.1.27

$$\log BA_t^{26C} = \underset{(.003)}{.009}TD_t - \underset{(.038)}{.074}EE_t + \underset{(.070)}{.075}LY_t + \underset{(.151)}{.074}\xi_t^{S02_10} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.099)}{.609}B^{12})a_t$$

$$\hat{\sigma}_a = 14.88\%$$

27A - Manufacture of domestic appliances

Model 5.1.28

$$\log BA_t^{27A} = \underset{(.001)}{.012}TD_t - \underset{(.010)}{.162}EE_t + \underset{(.019)}{.045}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.124)}{.263}B^{12})a_t$$

$$\hat{\sigma}_a = 4.76\%$$

No regular and seasonal structure are present in the series of this basic aggregate. Also all calendar variables are highly significative.

27B - Manufacture of electrical material and equipment except household appliances

Model 5.1.29

$$\log BA_t^{27B} = \underset{(.001)}{.012}TD_t - \underset{(.012)}{.101}EE_t + \underset{(.022)}{.034}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.129)}{.785}B^{12})a_t$$

$$\hat{\sigma}_a = 4.33\%$$

28 - Manufacture of machinery and equipment n.e.c

Model 5.1.30

$$\log BA_t^{28} = \underset{(.002)}{.010}TD_t - \underset{(.021)}{.102}EE_t + \underset{(.035)}{.062}LY_t + N_t$$

$$(\nabla\nabla_{12}N_t = (1 - \underset{(.080)}{.415}B)(1 - \underset{(.106)}{.546}B^{12})a_t$$

$$\hat{\sigma}_a = 6.53\%$$

29 - Manufacture of motor vehicles, trailers and semi-trailers

Model 5.1.31

$$\log BA_t^{29} = \underset{(.002)}{.010}TD_t - \underset{(.027)}{.130}EE_t + \underset{(.047)}{.054}LY_t + N_t$$

$$(1 + \underset{(.099)}{.228}B)\nabla\nabla_{12}N_t = (1 - \underset{(.175)}{1.000}B^{12})a_t$$

$$\hat{\sigma}_a = 7.45\%$$

There are pairs of extreme values that distort the ACF and PACF of the series. Particularly if we look at the robust autocorrelation functions we note as the coefficients of the lags 2, 3, 4 and 6 decrease significantly. It is no necessary to include the step that cause the extreme values because it does not produce a significative change in the estimated coefficients. We can observe that the seasonal moving average is no invertible with a coefficient $\Theta_1 = -1$. An explanation for this is the pile-up phenomenon due to the lack of identifiably of Θ (Breidt et all, 2006). It follows that $\Theta = -1$ is always a critical point of the likelihood function.

30A - Naval, railway, aircraft and spacecraft construction. Combat vehicles

Model 5.1.32

$$\log BA_t^{30A} = \underset{(.005)}{.006}TD_t + \underset{(.063)}{.009}EE_t + \underset{(.105)}{.176}LY_t + N_t$$

$$(1 + \underset{(.087)}{.570}B^{12})\nabla\nabla_{12}N_t = (1 - \underset{(.073)}{.767}B)a_t$$

$$\hat{\sigma}_a = 18.92\%$$

The Easter coefficient presents a higher p-value so we need to remove the variable from the model.

30B - Manufacture of motorcycles, bicycles, vehicles for disabled persons and others n.e.c.

Model 5.1.33

$$\log BA_t^{30B} = \underset{(.002)}{.015}TD_t - \underset{(.025)}{.084}EE_t + \underset{(.046)}{.060}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.249)}{1.000}B^{12})a_t$$

$$\hat{\sigma}_a = 8.10\%$$

31 - Manufacture of furniture

Model 5.1.34

$$\log BA_t^{31} = \underset{(.001)}{.013}TD_t - \underset{(.014)}{.120}EE_t + \underset{(.023)}{.035}LY_t + N_t$$

$$(1 - \underset{(.136)}{.752}B)\nabla\nabla_{12}N_t = (1 - \underset{(.143)}{1.198}B + \underset{(.106)}{.501}B^2)(1 - \underset{(.137)}{.469}B^{12})a_t$$

$$\hat{\sigma}_a = 4.62\%$$

32A - Manufacture of jewelery, costume jewelery and musical instruments

Model 5.1.35

$$\log BA_t^{32A} = \underset{(.002)}{.009}TD_t - \underset{(.026)}{.101}EE_t + \underset{(.042)}{.065}LY_t + N_t$$

$$(1 + \underset{(.095)}{.531}B)\nabla\nabla_{12}N_t = (1 - \underset{(.111)}{.216}B^{12})a_t$$

$$\hat{\sigma}_a = 9.18\%$$

The ACF and PACF of the residuals of the model present higher values of some delays that we can not pick up introducing a regular or seasonal structure, nor introducing a step or pulse as an intervention analysis. The more parsimonious model that presents the autocorrelation function free from significative delays is finally chosen.

32B - Manufacture of sports goods; games and toys. Other manufacturing industries

Model 5.1.36

$$\log BA_t^{32B} = \underset{(.002)}{.018}TD_t - \underset{(.028)}{.102}EE_t + \underset{(.049)}{.042}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.116)}{.433}B - \underset{(.123)}{.244}B^2)(1 - \underset{(.131)}{.502}B^{12})a_t$$

$$\hat{\sigma}_a = 8.94\%$$

32C - Manufacture of medical and dental instruments and supplies

Model 5.1.37

$$\log BA_t^{32C} = \underset{(.001)}{.008}TD_t - \underset{(.020)}{.109}EE_t + \underset{(.033)}{.070}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.066)}{.666}B)(1 - \underset{(.407)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 4.44\%$$

33 - Repair and installation of machinery and equipment

Model 5.1.38

$$\log BA_t^{33} = \underset{(.003)}{.006}TD_t - \underset{(.036)}{.032}EE_t - \underset{(.057)}{.004}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.059)}{.719}B)(1 - \underset{(.108)}{.693}B^{12})a_t$$

$$\hat{\sigma}_a = 8.93\%$$

Although both Easter and leap year coefficients are not significative, only the second is deleted from the model due to its higher p-value.

5.2 Analysis of the Industrial Turnover Index by divisions

10 - Manufacture of food products

Model 5.2.1

$$\log D_t^{10} = \underset{(.001)}{.010}TD_t - \underset{(.008)}{.047}EE_t + \underset{(.014)}{.031}LY_t + N_t$$

$$(1 + \underset{(.105)}{.357}B + \underset{(.098)}{.288}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.128)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 2.08\%$$

13 - Manufacture of textiles

Model 5.2.2

$$\log D_t^{13} = \underset{(.001)}{.013}TD_t - \underset{(.011)}{.105}EE_t + \underset{(.017)}{.032}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.101)}{.342}B + \underset{(.113)}{.258}B^2)a_t$$

$$\hat{\sigma}_a = 4.50\%$$

20 - Manufacture of chemicals and chemical products

Model 5.2.3

$$\log D_t^{20} = \underset{(.001)}{.009}TD_t - \underset{(.008)}{.090}EE_t + \underset{(.016)}{.001}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 + \underset{(.098)}{.178}B)(1 - \underset{(.103)}{.637}B^{12})a_t$$

$$\hat{\sigma}_a = 3.68\%$$

In this division the coefficient of leap year variable is not significative as in the basic aggregates that compose it.

25 - Manufacture of fabricated metal products, except machinery and equipment

Model 5.2.4

$$\log D_t^{25} = \underset{(.001)}{.012}TD_t - \underset{(.011)}{.107}EE_t + \underset{(.017)}{.027}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 + \underset{(.102)}{.007}B - \underset{(.096)}{.003}B^2 + \underset{(.099)}{.282}B^3 + \underset{(.100)}{.348}B^4)(1 - \underset{(.177)}{.858}B^{12})a_t$$

$$\hat{\sigma}_a = 3.40\%$$

26 - Manufacture of computer, electronic and optical products

Model 5.2.5

$$\log D_t^{26} = \underset{(.002)}{.005}TD_t - \underset{(.028)}{.059}EE_t + \underset{(.049)}{.015}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.098)}{.212}B)(1 - \underset{(.108)}{.733}B^{12})a_t$$

$$\hat{\sigma}_a = 8.93\%$$

In this model is necessary to remove the leap year variable due to its small significance.

27 - Manufacture of electrical equipment

Model 5.2.6

$$\log D_t^{27} = \underset{(.001)}{.012} TD_t - \underset{(.011)}{.112} EE_t + \underset{(.019)}{.037} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.105)}{.636} B^{12}) a_t$$

$$\hat{\sigma}_a = 4.04\%$$

30 - Manufacture of other transport equipment

Model 5.2.7

$$\log D_t^{30} = \underset{(.005)}{.010} TD_t - \underset{(.066)}{.015} EE_t + \underset{(.105)}{.118} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.062)}{.799} B)(1 - \underset{(.115)}{.741} B^{12}) a_t$$

$$\hat{\sigma}_a = 15.51\%$$

The variable of Easter effect has a p-value very high and so we delete it.

32 - Other manufacturing

Model 5.2.8

$$\log D_t^{32} = \underset{(.001)}{.014} TD_t - \underset{(.019)}{.112} EE_t + \underset{(.032)}{.043} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.121)}{.388} B)(1 - \underset{(.119)}{.407} B^{12}) a_t$$

$$\hat{\sigma}_a = 6.31\%$$

5.3 Analysis of the Industrial Turnover Index by economic destination

XC - Consumer goods

Model 5.3.1

$$\log XC_t = \underset{(.001)}{.010} TD_t - \underset{(.008)}{.065} EE_t + \underset{(.014)}{.040} LY_t + N_t$$

$$(1 + \underset{(.104)}{.499} B + \underset{(.099)}{.326} B^2)\nabla\nabla_{12} N_t = (1 - \underset{(.111)}{.582} B^{12}) a_t$$

$$\hat{\sigma}_a = 2.36\%$$

XC - CD - Durable consumer goods

Model 5.3.2

$$\log CD_t = \underset{(.001)}{.012} TD_t - \underset{(.011)}{.118} EE_t + \underset{(.021)}{.035} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.101)}{.549} B^{12}) a_t$$

$$\hat{\sigma}_a = 4.56\%$$

XC - CN - Non-durable consumer goods

Model 5.3.3

$$\log CN_t = \underset{(.001)}{.010} TD_t - \underset{(.007)}{.056} EE_t + \underset{(.012)}{.034} LY_t + N_t$$

$$(1 + \underset{(.101)}{.513} B + \underset{(.110)}{.471} B^2 + \underset{(.116)}{.083} B^3 + \underset{(.098)}{.298} B^4)\nabla\nabla_{12} N_t = (1 - \underset{(.115)}{.663} B^{12}) a_t$$

$$\hat{\sigma}_a = 2.10\%$$

EN - Energy

Model 5.3.4

$$\log EN_t = \underset{(.001)}{.002}TD_t - \underset{(.022)}{.038}EE_t - \underset{(.039)}{.019}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.186)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 6.95\%$$

There is not enough evidence to consider necessary to apply ∇_{12}. Looking at the original series, we don't observe a clear systematic behavior in certain months or a high value of seasonal delays followed by a slow decrease of the same. However, due to the increasing values of the seasonal delays in the ACF, we finally decide to apply a seasonal difference. In addition, the main component of this aggregate also has a seasonal and regular difference. The leap year variable is deleted from the model.

IP - Intermediate products

Model 5.3.5

$$\log IP_t = \underset{(.000)}{.011}TD_t - \underset{(.008)}{.105}EE_t + \underset{(.011)}{.011}LY_t + N_t$$

$$(1 - \underset{(.098)}{.175}B - \underset{(.101)}{.265}B^2)(1 + \underset{(.094)}{.413}B^{12} + \underset{(.095)}{.448}B^{24})\nabla\nabla_{12}N_t = a_t$$

$$\hat{\sigma}_a = 2.89\%$$

CG - Capital goods

Model 5.3.6

$$\log CG_t = \underset{(.001)}{.010}TD_t - \underset{(.015)}{.097}EE_t + \underset{(.026)}{.053}LY_t + N_t$$

$$(1 + \underset{(.102)}{.208}B + \underset{(.099)}{.247}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.214)}{.874}B^{12})a_t$$

$$\hat{\sigma}_a = 4.52\%$$

In the original series we can observe a step in January 2009 and a pulse in December 2005, but we do not include these effects with an intervention analysis since no effects in the coefficients are obtained with the inclusion.

5.4 Analysis of the Industrial Turnover Index for sections B and C

B - Section B

Model 5.4.1

$$\log B_t = \underset{(.002)}{.014}TD_t - \underset{(.031)}{.101}EE_t + \underset{(.048)}{.004}LY_t - \underset{(.042)}{.443}\xi^{S10:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.072)}{.850}B)(1 - \underset{(.0215)}{.850}B^{12})a_t$$

$$\hat{\sigma}_a = 6.62\%$$

where the coefficient of the leap year variable is not significative and must be removed from the model.

C - Section C

Model 5.4.2

$$\log C_t = \underset{(.001)}{.010}TD_t - \underset{(.008)}{.086}EE_t + \underset{(.014)}{.030}LY_t - \underset{(.032)}{.122}\xi^{S10:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.117)}{.223}B)(1 - \underset{(.120)}{.698}B^{12})a_t$$

$$\hat{\sigma}_a = 2.63\%$$

5.5 Analysis of the Industrial Turnover General Index

IG - General Index for the Industrial turnover

Model 5.5.1

$$\log IG_t = \underset{(.001)}{.010}TD_t - \underset{(.008)}{.086}EE_t + \underset{(.014)}{.030}LY_t - \underset{(.032)}{.128}\xi^{S10:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.116)}{.232}B)(1 - \underset{(.122)}{.707}B^{12})a_t$$

$$\hat{\sigma}_a = 2.61\%$$

6 Analysis of the Industrial New Orders Received series

As in the Industrial Turnover, the indices are calculated for all basic aggregates (two digits or set of groups of three digits) and another functional aggregates like other divisions that are not basic aggregates (two digits), sections B and C and general. We calculate also the index for economic sectors by economic destination (durable consumer goods, non-durable consumer goods, capital goods, intermediate goods and energy). In the next subsection (6.1) we describe the different models identified for the basic aggregates. In subsections 6.2 and 6.3 we show the final models chosen for other divisions and functional aggregates by economic destination. Finally in subsections 6.4 and 6.5 we describe the determined models for sections B and C and for the General Index.

6.1 Analysis of the Industrial New Orders Received Index for basic aggregates

05 - Mining of coal and lignite

To this basic aggregate the model that we obtain is the following:

Model 6.1.1

$$\log BA_t^{05} = \underset{(.006)}{.003}TD_t - \underset{(.066)}{.005}EE_t - \underset{(.114)}{.106}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.110)}{.508}B - \underset{(.100)}{.180}B^2 - \underset{(.085)}{.312}B^3)a_t$$

$$\hat{\sigma}_a = 25.68\%$$

We need to remove the trading-day and Easter variables because their p-values are higher than 0.5.

08 - Other mining and quarrying

Model 6.1.2

$$\log BA_t^{08} = \underset{(.003)}{.016}TD_t - \underset{(.047)}{.131}EE_t + \underset{(.077)}{.044}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.089)}{.563}B)(1 - \underset{(.306)}{.920}B^{12})a_t$$

$$\hat{\sigma}_a = 11.45\%$$

The variable LY_T is deleted due to its p-value is greater than 0.5.

10A - Food industries (except grain mill products and animal feeds)

To the basic aggregate 10A the model that we obtain is the following:

Model 6.1.3

$$\log BA_t^{10A} = \underset{(.001)}{.010}TD_t - \underset{(.008)}{.038}EE_t + \underset{(.015)}{.052}LY_t + N_t$$

$$(1 + \underset{(.101)}{.357}B + \underset{(.092)}{.379}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.170)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 2.30\%$$

10B - Manufacture of grain mill products, starches and starch products, and of animal feeds

Model 6.1.4

$$\log BA_t^{10B} = \underset{(.001)}{.009}TD_t - \underset{(.010)}{.043}EE_t + \underset{(.017)}{.012}LY_t - \underset{(.031)}{.067}\xi^{S10:2008} + N_t$$

$$\nabla\nabla_{12}N_t = ((1 - \underset{(.105)}{.175}B + \underset{(.103)}{.053}B^2 + \underset{(.094)}{.228}B^3)(1 - \underset{(.165)}{.789}B^{12})a_t$$

$$\hat{\sigma}_a = 3.13\%$$

The leap year variable presents a p-value in the range [0.2, 0.5], but the value of its coefficient is too small, so we remove the variable from the model.

11 - Manufacture of beverages

Model 6.1.5

$$\log BA_t^{11} = \underset{(.001)}{.012}TD_t - \underset{(.017)}{.036}EE_t + \underset{(.027)}{.077}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.082)}{.675}B)(1 - \underset{(.137)}{.700}B^{12})a_t$$

$$\hat{\sigma}_a = 4.25\%$$

12 - Manufacture of tobacco products

Model 6.1.6

$$\log Ta_t = \underset{(.003)}{.020}TD_t - \underset{(.040)}{.140}EE_t + \underset{(.073)}{.089}LY_t + N_t$$

$$(1 + \underset{(.093)}{.404}B^{12})\nabla\nabla_{12}N_t = (1 - \underset{(.107)}{.474}B)a_t$$

$$\hat{\sigma}_a = 14.01\%$$

13A - Preparation and spinning of textile fibres. Manufacture of woven textiles. Textile finishings

Model 6.1.7

$$\log BA_t^{13A} = \underset{(.001)}{.014}TD_t - \underset{(.016)}{.114}EE_t + \underset{(.029)}{.044}LY_t + N_t$$

$$(1 + \underset{(.096)}{.362}B + \underset{(.096)}{.251}B)\nabla\nabla_{12}N_t = (1 - \underset{(.147)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 4.37\%$$

13B - Manufacture of knitted and crocheted articles, rugs, cordage, non-woven fabrics, textile products for technical and industrial use and other textile products

Model 6.1.8

$$\log BA_t^{13B} = \underset{(.001)}{.012}TD_t - \underset{(.016)}{.096}EE_t + \underset{(.028)}{.085}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.079)}{.424}B)(1 - \underset{(.102)}{.485}B^{12})a_t$$

$$\hat{\sigma}_a = 5.30\%$$

14 - Manufacture of wearing apparel

Model 6.1.9

$$\log BA_t^{14} = \underset{(.003)}{.009}TD_t - \underset{(.037)}{.073}EE_t + \underset{(.056)}{.044}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.046)}{.909}B)(1 - \underset{(.102)}{.394}B^{12})a_t$$

$$\hat{\sigma}_a = 9.67\%$$

15 - Manufacture of leather and related products

Model 6.1.10

$$\log BA_t^{15} = \underset{(.002)}{.016}TD_t - \underset{(.033)}{.116}EE_t - \underset{(.056)}{.015}LY_t + N_t$$

$$(1 + \underset{(.096)}{.530}B + \underset{(.088)}{.423}^{B}2 + \underset{(.086)}{.615}B^3 + \underset{(.103)}{.362}B^4)\nabla\nabla_{12}N_t = (1 - \underset{(.121)}{.699B^{12}})\, a_t$$

$$\hat{\sigma}_a = 9.64\%$$

The LY_t variable is not significative and its sign is opposite at normal, so we must remove it from the model.

16 - Manufacture of wood and of products of wood and cork

Model 6.1.11

$$\log BA_t^{16} = \underset{(.001)}{.013}TD_t - \underset{(.017)}{.096}EE_t + \underset{(.030)}{.032}LY_t + N_t$$

$$(1 + \underset{(.101)}{.295}B + \underset{(.099)}{.248}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.117)}{.699}B^{12})a_t$$

$$\hat{\sigma}_a = 5.39\%$$

17 - Manufacture of paper and paper products

Model 6.1.12

$$\log BA_t^{17} = \underset{(.001)}{.010}TD_t - \underset{(.009)}{.079}EE_t + \underset{(.017)}{.038}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.090)}{.287}B)(1 - \underset{(.110)}{.730B^{12}})a_t$$

$$\hat{\sigma}_a = 2.95\%$$

18 - Printing and reproduction of recorded media

Model 6.1.13

$$\log BA_t^{18} = \underset{(.003)}{.005}TD_t - \underset{(.040)}{.019}EE_t + \underset{(.065)}{.013}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.051)}{.852}B)(1 - \underset{(.153)}{.271}B^{12})a_t$$

$$\hat{\sigma}_a = 11.45\%$$

The Easter and leap year variables are not significative presenting p-values very high.

19 - Manufacture of coke and refined petroleum products

Model 6.1.14

$$\log BA_t^{19} = \underset{(.001)}{.003}TD_t - \underset{(.021)}{.021}EE_t - \underset{(.037)}{.023}LY_t - \underset{(.072)}{.248}\xi^{S04:2003} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.140)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 6.63\%$$

Again we remove the Easter and leap year variables. For the LY_t variable the reason is the same that in the previous division, but Easter coefficient is deleted because its value is small compared with the normal value and its p-value is greater than 0.2.

20A - Manufacture of cleaning articles, perfumes and cosmetics

Model 6.1.15

$$\log BA_t^{20A} = \underset{(.001)}{.008}TD_t - \underset{(.019)}{.072}EE_t + \underset{(.031)}{.019}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.082)}{.711}B)(1 - \underset{(.113)}{.569}B^{12})a_t$$

$$\hat{\sigma}_a = 5.16\%$$

Again we need to delete the LY_t variable.

20B - Chemical industry except cleaning articles, perfumes and cosmetics

Model 6.1.16

$$\log BA_t^{20B} = \underset{(.001)}{.007}TD_t - \underset{(.015)}{.095}EE_t + \underset{(.026)}{.022}LY_t - \underset{(.043)}{.296}\xi^{S11:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.099)}{.484}B)(1 - \underset{(.104)}{.701}B^{12})a_t$$

$$\hat{\sigma}_a = 4.27\%$$

21 - Manufacture of basic pharmaceutical products and pharmaceutical preparations

Model 6.1.17

$$\log BA_t^{21} = \underset{(.001)}{.012}TD_t - \underset{(.017)}{.065}EE_t + \underset{(.027)}{.015}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.063)}{.678}B)(1 - \underset{(.504)}{.947}B^{12})a_t$$

$$\hat{\sigma}_a = 3.84\%$$

We need to remove the LY_t variable because its p-value is greater than 0.5.

22 - Manufacture of rubber and plastic products

Model 6.1.18

$$\log BA_t^{22} = \underset{(.001)}{.008}TD_t - \underset{(.014)}{.099}EE_t + \underset{(.018)}{.004}LY_t + N_t$$

$$(1 + \underset{(.093}{.503}B^{12} + \underset{(.097)}{.464}B^{24})\nabla\nabla_{12}N_t = (1 - \underset{(.101}{.209}B + \underset{(.091)}{.298}B^2)a_t$$

$$\hat{\sigma}_a = 3.97\%$$

The leap year variable is removed because it is not significative at 50%.

23 - Manufacture of other non-metallic mineral products

Model 6.1.19

$$\log BA_t^{23} = \underset{(.001)}{.012}TD_t - \underset{(.009)}{.110}EE_t + \underset{(.017)}{.003}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.126)}{.551}B^{12})a_t$$

$$\hat{\sigma}_a = 3.75\%$$

As in the previous model, the leap year variable is removed.

24 - Manufacture of basic metals

Model 6.1.20

$$\log BA_t^{24} = \underset{(.001)}{.008}TD_t - \underset{(.016)}{.089}EE_t + \underset{(.030)}{.010}LY_t - \underset{(.042)}{.141}\xi^{I07:2004} - \underset{(.057)}{.212}\xi^{I10:2008} - \underset{(.058)}{.161}\xi^{S11:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.184)}{.840}B^{12})a_t$$

$$\hat{\sigma}_a = 5.75\%$$

Again we must remove the LY_t variable.

25A - Manufacture of metal products for construction, containers made of metal, steam generators, weapons and ammunition

Model 6.1.21

$$\log BA_t^{25A} = \underset{(.002)}{.009}TD_t - \underset{(.035)}{.058}EE_t + \underset{(.058)}{.005}LY_t + \underset{(.082)}{.460}\xi^{I12:2006} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.074)}{.483}B)(1 - \underset{(.171)}{.868}B^{12})a_t$$

$$\hat{\sigma}_a = 9.07\%$$

The p-value of the leap year variable is close to 0.90, so we remove this variable.

25B - Forging, stamping, embossing and rolling of metals. Manufacture of tools, hardware goods, containers and other metal products

Model 6.1.22

$$\log BA_t^{25B} = \underset{(.001)}{.011}TD_t - \underset{(.013)}{.114}EE_t + \underset{(.023)}{.030}LY_t - \underset{(.047)}{.185}\xi^{S12:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.109)}{.210}B)(1 - \underset{(.104)}{.679}B^{12})a_t$$

$$\hat{\sigma}_a = 4.39\%$$

26A - Manufacture of electronic components, assembled printed circuits, and magnetic and optical media

Model 6.1.23

$$\log BA_t^{26A} = \underset{(.007)}{-.003}TD_t + \underset{(.107)}{.061}EE_t + \underset{(.165)}{.367}LY_t + \underset{(.247)}{1.33}\xi^{I04:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.085)}{.448}B)(1 - \underset{(.357)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 23.90\%$$

In this basic aggregate both trading-day and Easter effect are not significative.

26B - Manufacture of computers, peripherals and telecommunications equipment; appliances for measuring and navigation; radiation and medical and therapeutic equipment

Model 6.1.24

$$\log BA_t^{26B} = \underset{(.005)}{.015}TD_t + \underset{(.068)}{.126}EE_t - \underset{(.109)}{.004}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.076)}{.674}B)(1 - \underset{(.162)}{.831}B^{12})a_t$$

$$\hat{\sigma}_a = 16.32\%$$

We need to remove the leap year variable given its high p-value and the Easter variable due to its positive sign.

26C - Manufacture of consumer electronics, optical instruments and photographic equipment

Model 6.1.25

$$\log BA_t^{26C} = \underset{(.006)}{.008}TD_t - \underset{(.090)}{.086}EE_t - \underset{(.138)}{.008}LY_t + \underset{(.207)}{.672}\xi_t^{S02:2010} - \underset{(.188)}{1.26}\xi_t^{I11:2007} + N_t$$

$$(1 + \underset{(.087)}{.553}B)\nabla\nabla_{12}N_t = (1 - \underset{(.123)}{.764}B^{12})a_t$$

$$\hat{\sigma}_a = 23.81\%$$

According to the guidelines mentioned before, the $LY-t$ variable is removed from the model although there are more variables with p-values greater than 0.2.

27A - Manufacture of domestic appliances

Model 6.1.26

$$\log BA_t^{27A} = \underset{(.001)}{.008}TD_t - \underset{(.016)}{.094}EE_t + \underset{(.027)}{.017}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.091)}{.364}B)(1 - \underset{(.126)}{.624}B^{12})a_t$$

$$\hat{\sigma}_a = 5.02\%$$

Again we remove the LY_t variable.

27B - Manufacture of electrical material and equipment except household appliances

Model 6.1.27

$$\log BA_t^{27B} = \underset{(.002)}{.008}TD_t - \underset{(.030)}{.046}EE_t + \underset{(.051)}{.083}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.089)}{.423}B)(1 - \underset{(.099)}{.712}B^{12})a_t$$

$$\hat{\sigma}_a = 8.68\%$$

28 - Manufacture of machinery and equipment n.e.c

Model 6.1.28

$$\log BA_t^{28} = \underset{(.003)}{.007}TD_t - \underset{(.038)}{.115}EE_t - \underset{(.066)}{.057}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.105)}{.375}B)(1 - \underset{(.309)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 10.03\%$$

Due to the opposite sign of the leap year coefficient, we need to remove this variable from the model.

29 - Manufacture of motor vehicles, trailers and semi-trailers

Model 6.1.29

$$\log BA_t^{29} = \underset{(.002)}{.009}TD_t - \underset{(.025)}{.058}EE_t + \underset{(.042)}{.042}LY_t + N_t$$

$$(1 + \underset{(.096)}{.325}B)\nabla\nabla_{12}N_t = (1 - \underset{(.127)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 6.62\%$$

30A - Naval, railway, aircraft and spacecraft construction. Combat vehicles

Model 6.1.30

$$\log BA_t^{30A} = \underset{(.020)}{.023}TD_t + \underset{(.284)}{.494}EE_t + \underset{(.431)}{.119}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.038)}{.908}B)(1 - \underset{(.144)}{.839}B^{12})a_t$$

$$\hat{\sigma}_a = 60.39\%$$

The Easter and leap year variables are removed from the model, but their causes are differents. The EE_t variable is removed because it present a opposite sign that is normal, while the leap year variable is removed because its p-value is close to 0.80.

30B - Manufacture of motorcycles, bicycles, vehicles for disabled persons and others n.e.c.

Model 6.1.31

$$\log BA_t^{30B} = \underset{(.005)}{.002}TD_t - \underset{(.056)}{.102}EE_t + \underset{(.088)}{.178}LY_t + N_t$$

$$(1 + \underset{(.085)}{.776}B + \underset{(.088)}{.509}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.125)}{.728}B^{12})a_t$$

$$\hat{\sigma}_a = 14.28\%$$

The trading-day variable is removed from the model because its p-value is greater than 0.5.

31 - Manufacture of furniture

Model 6.1.32

$$\log BA_t^{31} = \underset{(.001)}{.013}TD_t - \underset{(.013)}{.112}EE_t + \underset{(.023)}{.010}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.086)}{.230}B)(1 - \underset{(.099)}{.653}B^{12})a_t$$

$$\hat{\sigma}_a = 4.37\%$$

Again the LY_t variable is not significative.

32A - Manufacture of jewelery, costume jewelery and musical instruments

Model 6.1.33

$$\log BA_t^{32A} = \underset{(.003)}{.009} TD_t - \underset{(.036)}{.136} EE_t + \underset{(.059)}{.031} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.071)}{.631} B)(1 - \underset{(.097)}{.500} B^{12}) a_t$$

$$\hat{\sigma}_a = 10.26\%$$

As in the previous aggregate the leap year variable is removed.

32B - Manufacture of sports goods; games and toys. Other manufacturing industries

Model 6.1.34

$$\log BA_t^{32B} = \underset{(.003)}{.012} TD_t - \underset{(.038)}{.121} EE_t - \underset{(.067)}{.058} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.103)}{.580} B - \underset{(.105)}{.340} B^2)(1 - \underset{(.102)}{.687} B^{12}) a_t$$

$$\hat{\sigma}_a = 10.59\%$$

The leap year coefficient is negative when it is usually positive, so we need to remove it from the model.

32C - Manufacture of medical and dental instruments and supplies

Model 6.1.35

$$\log BA_t^{32C} = \underset{(.002)}{.007} TD_t - \underset{(.025)}{.108} EE_t + \underset{(.040)}{.085} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.063)}{.717} B)(1 - \underset{(.366)}{1.00} B^{12}) a_t$$

$$\hat{\sigma}_a = 5.40\%$$

33 - Repair and installation of machinery and equipment

Model 6.1.36

$$\log BA_t^{33} = \underset{(.009)}{.018}TD_t + \underset{(.135)}{.036}EE_t - \underset{(.205)}{.036}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.048)}{.903}B)(1 - \underset{(.189)}{.844}B^{12})a_t$$

$$\hat{\sigma}_a = 28.66\%$$

Easter and leap year variables are deleted from the model because they are not significative.

6.2 Analysis of the Industrial New Orders Received Index by divisions

10 - Manufacture of food products

Model 6.2.1

$$\log D_t^{10} = \underset{(.001)}{.009}TD_t - \underset{(.008)}{.040}EE_t + \underset{(.014)}{.046}LY_t + N_t$$

$$(1 + \underset{(.102)}{.325}B + \underset{(.094)}{.331}B^2)\nabla\nabla_{12}N_t = (1 - \underset{(.155)}{1.00}B^{12})a_t$$

$$\hat{\sigma}_a = 2.19\%$$

13 - Manufacture of textiles

Model 6.2.2

$$\log D_t^{13} = \underset{(.001)}{.011} TD_t - \underset{(.006)}{.107} EE_t + \underset{(.011)}{.037} LY_t + \underset{(.011)}{.187} \xi^{I08:2006} + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.084)}{.236} B + \underset{(.089)}{.204} B^2 - \underset{(.087)}{.239} B^3 - \underset{(.078)}{.267} B + \underset{(.080)}{.629} B^5)(1 - \underset{(.101)}{.430} B^{12}) a_t$$

$$\hat{\sigma}_a = 3.96\%$$

20 - Manufacture of chemicals and chemical products

Model 6.2.3

$$\log D_t^{20} = \underset{(.001)}{.007} TD_t - \underset{(.013)}{.090} EE_t + \underset{(.023)}{.021} LY_t - \underset{(.037)}{.266} \xi^{S11:2008} + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.096)}{.523} B)(1 - \underset{(.105)}{.682} B^{12}) a_t$$

$$\hat{\sigma}_a = 3.77\%$$

25 - Manufacture of fabricated metal products, except machinery and equipment

Model 6.2.4

$$\log D_t^{25} = \underset{(.001)}{.010} TD_t - \underset{(.017)}{.088} EE_t + \underset{(.029)}{.017} LY_t + \underset{(.041)}{.287} \xi^{I12:2006} + \underset{(.040)}{.098} \xi^{I08:2004} + N_t$$

$$(1 + \underset{(.099)}{.229} B) \nabla\nabla_{12} N_t = (1 - \underset{(.208)}{.883} B^{12}) a_t$$

$$\hat{\sigma}_a = 4.96\%$$

Following the instructions discused above, the LY_t variable is removed since its p-value is greater than 0.5.

26 - Manufacture of computer, electronic and optical products

Model 6.2.5

$$\log D_t^{26} = \underset{(.004)}{.012} TD_t - \underset{(.061)}{.082} EE_t + \underset{(.103)}{.033} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.083)}{.467} B)(1 - \underset{(.223)}{1.00} B^{12}) a_t$$

$$\hat{\sigma}_a = 15.05\%$$

As in the previous aggregate the leap year variable is not significative.

27 - Manufacture of electrical equipment

Model 6.2.6

$$\log D_t^{27} = \underset{(.002)}{.007} TD_t - \underset{(.025)}{.054} EE_t + \underset{(.042)}{.069} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.088)}{.397} B)(1 - \underset{(.110)}{.755} B^{12}) a_t$$

$$\hat{\sigma}_a = 7.07\%$$

30 - Manufacture of other transport equipment

Model 6.2.7

$$\log D_t^{30} = \underset{(.015)}{.020} TD_t + \underset{(.221)}{.386} EE_t + \underset{(.337)}{.143} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.040)}{.883} B)(1 - \underset{(.141)}{.834} B^{12}) a_t$$

$$\hat{\sigma}_a = 47.65\%$$

We need to remove the leap year variable due to its high p-value and the Easter variable because its sign is positive.

32 - Other manufacturing

Model 6.2.8

$$\log D_t^{32} = \underset{(.002)}{.010} TD_t - \underset{(.028)}{.120} EE_t - \underset{(.046)}{.008} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.101)}{.610} B)(1 - \underset{(.110)}{.734} B^{12}) a_t$$

$$\hat{\sigma}_a = 7.39\%$$

Again the leap year variable is removed due its high p-value.

6.3 Analysis of the Industrial New Orders Received Index by economic destination

XC - Consumer goods

Model 6.3.1

$$\log XC_t = \underset{(.001)}{.010} TD_t - \underset{(.009)}{.055} EE_t + \underset{(.014)}{.045} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.080)}{.520} B)(1 - \underset{(.184)}{.476} B^{12} - \underset{(.159)}{.361} B^{24}) a_t$$

$$\hat{\sigma}_a = 2.45\%$$

XC - CD - Durable consumer goods

Model 6.3.2

$$\log CD_t = \underset{(.001)}{.010}TD_t - \underset{(.021)}{.104}EE_t + \underset{(.034)}{.007}LY_t + N_t$$

$$(1 + \underset{(.087)}{.500}B)\nabla\nabla_{12}N_t = (1 - \underset{(.291)}{.913}B^{12})a_t$$

$$\hat{\sigma}_a = 5.46\%$$

We need to remove the leap year variable because its p-value is very high.

XC - CN - Non-durable consumer goods

Model 6.3.3

$$\log CN_t = \underset{(.001)}{.010}TD_t - \underset{(.010)}{.049}EE_t + \underset{(.016)}{.045}LY_t + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.071)}{.628}B)(1 - \underset{(.128)}{.724}B^{12})a_t$$

$$\hat{\sigma}_a = 2.55\%$$

EN - Energy

Model 6.3.4

$$\log EN_t = \underset{(.001)}{.003}TD_t - \underset{(.021)}{.020}EE_t - \underset{(.038)}{.023}LY_t - \underset{(.076)}{.249}\xi^{S04:2003} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.120)}{.798}B^{12})a_t$$

$$\hat{\sigma}_a = 7.46\%$$

Both Easter and leap year variables are removed from the model.

IP - Intermediate products

Model 6.3.5

$$\log IP_t = \underset{(.001)}{.010} TD_t - \underset{(.009)}{.097} EE_t + \underset{(.015)}{.025} LY_t - \underset{(.026)}{.152} \xi^{S10:2008} - \underset{(.027)}{.091} \xi^{S11:2008} - \underset{(.026)}{.104} \xi^{S12:2008} + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.100)}{.328} B)(1 - \underset{(.141)}{.779} B^{12}) a_t$$

$$\hat{\sigma}_a = 2.58\%$$

CG - Capital goods

Model 6.3.6

$$\log CG_t = \underset{(.002)}{.011} TD_t - \underset{(.029)}{.062} EE_t + \underset{(.045)}{.034} LY_t + \underset{(.068)}{.287} \xi^{I03:2008} + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.080)}{.441} B)(1 - \underset{(.161)}{1.00} B^{12}) a_t$$

$$\hat{\sigma}_a = 6.58\%$$

6.4 Analysis of the Industrial New Orders Received Index for section B and C

B - Section B

Model 6.4.1

$$\log B_t = \underset{(.003)}{.014} TD_t - \underset{(.042)}{.107} EE_t + \underset{(.069)}{.064} LY_t + N_t$$

$$\nabla\nabla_{12} N_t = (1 - \underset{(.095)}{.560} B)(1 - \underset{(.089)}{.466} B^{12}) a_t$$

$$\hat{\sigma}_a = 12.57\%$$

C - Section C

Model 6.4.2

$$\log C_t = \underset{(.001)}{.009}TD_t - \underset{(.011)}{.057}EE_t + \underset{(.018)}{.017}LY_t - \underset{(.035)}{.197}\xi^{S11:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.096)}{.458}B)(1 - \underset{(.136)}{.800}B^{12})a_t$$

$$\hat{\sigma}_a = 2.92\%$$

The LY_t variable presents a high p-value and the value of its coefficients is too small, so we remove it from the model.

6.5 Analysis of the Industrial New Orders Received General Index

IG - General Index for the Industrial turnover

Model 6.5.1

$$\log IG_t = \underset{(.001)}{.009}TD_t - \underset{(.011)}{.058}EE_t + \underset{(.018)}{.017}LY_t - \underset{(.035)}{.199}\xi^{S11:2008} + N_t$$

$$\nabla\nabla_{12}N_t = (1 - \underset{(.098)}{.460}B)(1 - \underset{(.145)}{.815}B^{12})a_t$$

$$\hat{\sigma}_a = 2.91\%$$

As in the seccion C, the leap year variable need to be removed.

References

[1] Eurostat (2009), ESS Guidlines on the seasonal adjustment, Methodologies and Workings Papers.

[2] Eurostat (2006), Methodology in short-term business statistics, Official Publications of the European Communities.

[3] Instituto Nacional de Estadística (2010), Índices de Cifra de Negocios. Índices de Entrada de Pedidos. Base 2005. CNAE-09 Manual Metodológico.

[4] Allan Young (1965), Estimating Trading-day Variation in Monthly Economic Time Series, Technical Paper No. 12, Bureau of the Census.

[5] William S. Cleveland and Susan J. Devlin (1980), Calendar Effects in Monthly Time Series: Modeling and Adjustment, Working Paper, Bell Laboratories.

[6] William S. Cleveland and M. R. Grupe (1981), Modeling Time Series when Calendar Effects are Present, Special studies papers No. 162, Board of Governors of the Federal Reserve System (U.S.).

[7] W. R. Bell and S. C. Hillmer (1983), Modeling Time Series With Calendar Variation, Journal of American Statistical Association Vol. 78, No. 383, pp. 526-534.

[8] S. C. Hillmer (1982) Forecasting Time Series with Trading Day Variation, Journal of Forecasting, Vol. 1, No. 4, pp. 385-395.

[9] William R. Bell (1984), Seasonal Descomposition of Deterministic Effects, Research Report 84/01, Statistical Research Division, Bureau of the Census.

[10] William R. Bell (1995), Correction to "Seasonal Descomposition of Deterministic Effects", Research Report 95/01, Statistical Research Division, Bureau of the Census.

[11] Brian C. Monsell (2007), Issues in Modeling and Adjusting Calendar Effects in Economic Time Series, Proceedings of the Third International Conference on Establishment Surveys, Alexandria, American Statistical Association.

[12] David F. Findley and Brian C. Monsell (2007), Modeling Stock Trading Day Effects Under Flow Day-of-Week Effect Constraints, Proceedings of the 2007 ISI Meetings.

[13] David F. Findley, Brian C. Monsell, William R. Bell, Mark C. Otto and Bor-Chung Chen (1998), New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program, Journal of Busines and Economic Statistic, Vol. 16, pp. 127-177.

[14] Brian C. Monsell, Christipher G. Roberts and Scott H. Holan (2010), Comparison of X-12-ARIMA Trading Day and Holyday Regressors With Country Specific Regressors, Journal of Official Statistics, Vol. 26, No. 2, pp. 371394.

[15] David F. Findley and Raymond J. Soukup (2000a), Modeling and Model Selection for Moving Holidays, ASA proceedings.

[16] David F. Findley and Raymond J. Soukup (2000b), Detection and modeling of Trading Day Effects, ICES proceedings.

[17] F. Jay Breidt, Richard A. Davis, Nan-Juang Hsu and Murray Rosenblatt (2006), Pile-up probabilities for the Laplace likelihood estimator of a non-invertible first order moving average, IMS Lecture Notes-Monograph Series, Time Series and Related Topics, Vol. 52, pp. 1-19.