Pre-adjustment process of real retail trade series in Croatia
Josip Arnerić, Anita Čeh Časni
University of Zagreb, Faculty of Economics and Business Zagreb, Croatia
Ante Rozga
University of Split, Faculty of Economics Split, Croatia

Keywords
Calendar effects, time series, pre-adjustment process, country specific regressors.

Abstract
Economic activity fluctuations are often influenced by various factors related to the calendar. These factors include non-working (non-trading) days, leap years, public holidays etc. In analysis of many economic variables in which time series are seasonally adjusted it is necessary to identify and correct calendar effects using suitable method, but there is no general or unique procedure for correcting these effects in pre-adjustment process. If these effects are not well adjusted, the identification of the ARIMA model may not be correct, and the quality of the seasonal adjustment is poor. Therefore, in these paper different methods of correction of the calendar effects are compared and applied to time series of real retail trade turnover (RRT) in Croatia (monthly data observed from January 2001 to December 2013). The most common method used is regression model with different types of explanatory variables which take into account calendar effects. The contribution of this paper is to define new explanatory variables (regressors that include not only different number of working and non-working days of the month but also country specific calendar effects) which will give most accurate correction of RRTT time series.

1. Introduction
Short-term statistics are often characterized by seasonal fluctuations and other calendar and trading-day effects, which can camouflage relevant short and long-term movements of the series, and obstruct a clear understanding of analyzed economic phenomena. Therefore, the main aim of seasonal adjustment is to remove changes that are due to seasonal or calendar influences to produce a clearer picture of the underlying behavior of the analyzed series according to Cleveland and Devlin (1980).

Even particular adjustment methods are similar; for example method that uses U.S. Census Bureau compared to the Eurostat method, it is not clear which method is better. The choice of method usually depends on experience and the policy of entire organization such as Croatian Bureau of Statistics. Calendar adjustment as well as seasonal adjustment is a source of constant debate, due to the different methods that can be used, and the different tools and computer programs that exist.

According to INE (2013), seasonal fluctuations are movements that occur with a similar intensity each month (quarter) or each season of the year, and which are expected to continue occurring. Furthermore, they define the calendar effect as the impact produced in the time series, due to the different structure that the months (quarters) present in the different years (in both length and composition), even if the remaining factors influencing the variable of interest remain constant.

Most of economic series are observed on a monthly or quarterly basis, but months (quarters) are not comparable due to different number of working and non-working days (not the same number of Mondays, Tuesdays, etc.). This may have an impact on the observed variables such as
retail sales, industrial productions and transportation. For example, retail sale turnover is likely to be more important on Saturdays than on other days of the week. Hence, there is a need for revision of the data adjusted for seasonal effects and calendar effects with two main reasons: they are a result of the revision of the gross data, due to an improvement in the information (in terms of coverage and/or reliability) or they are a result of a better estimation of the seasonality pattern, due to new information provided by new gross data, and due to the characteristics of the filters and procedures that eliminate the calendar and seasonal components by Palma and Marini (2004).

Even so, despite being a frequent practice, seasonal adjustment is still a source of constant debate resulting from different methods applied, different tools and computer programs that exist. Moreover, the manipulation of the original data that might occur through seasonal adjustment is also questioned. According to Arteche et al. (2011), calendar effects can be divided into two groups. The first group includes the effects of working (or trading) days and the second group deals with special calendar effects, such as Christmas, Easter or other (national) holidays. During the data processing, the abovementioned effects have to be taken into account and various methods of seasonal adjustment related to them must be applied. An inappropriate or poor quality of seasonal and calendar adjustment can generate false signals and can negatively affect the interpretation of the adjusted data. Having all this in mind, the aim of this paper is to apply different methods of correction of the calendar effects of real retail trade turnover (RRTT) in Croatia and define new explanatory variables which will give most accurate correction of RRTT time series. Therefore, new regressors of calendar effects adjustment will be defined.

The reminder of this paper is as follows. After the introduction in section 2 the different methods and alternative models for calendar effect correction are presented. Section 3 will give the results of the empirical analysis. Finally, in section 4 concluding remarks will be given.

2. Overview of different methods of calendar effects correction

There are many different methods for correction of calendar effects. The empirical analysis in this paper is focused on the linear regression method with different regressors for calendar adjustment of time series. Namely, corrections of working (or trading) days are carried out from the estimation of the linear regression. So, in a manner of Arteche et al. (2011) the following model of calendar structure using the explanatory variables is defined:

\[ y_t = z_t'\gamma + \epsilon_t \]

\[ \phi(L)\delta(L)\epsilon_t = \theta(L)v_t, \quad t = 1, 2, \ldots, T \]

where \( y_t \) is the time series of interest, \( \epsilon_t \) are error terms that follow an ARIMA process, \( \phi(L) \) and \( \theta(L) \) are finite polynomials of the lag operator \( L \), \( z_t' \) is a vector of \((K \times 1)\) of \( K \) relevant regressors, \( \gamma \) is a \((K \times 1)\) vector of unknown parameters and \( v_t \) is an error term defined as white noise.

Component \( z_t'\gamma \) represents nonstochastic effects which are subtracted from the original series before applying the ARIMA methodology for decomposition of the time series into trend/cycle, seasonal and irregular component. The simplest nonstochastic effect is the average, in this case the regression constant. More complex effects are for instance the intervention variables, atypical observations or calendar effects. Each time period is characterized by different number of Mondays, Tuesdays, ..., Sundays, therefore economic activity (in our case the real retail trade turnover) can be affected by this fact.
The working days effect differentiates working days from weekend days, so, in relevant literature, variable \( w_e \) is usually used to express the weighted difference between the number of working days \( w \) and non-working days \( n_w \) during the analyzed period \( t \). This is defined as:

\[
w_e = \left( w - \frac{5}{2}n_w \right)
\]

where the number of non-working days is multiplied by \( 5/2 \) so that the average of the newly created variable was zero. The coefficient of the variable \( w_e \) includes the effect of additional working days in the period \( t \).

It must be taken into consideration that the length and composition of each month or quarter has a seasonal part which has to be captured in the seasonal component, and must not be eliminated. For example, March always has 31 days, and on average, it has more Mondays than February. Moreover, the working-day adjustment must only be associated with the non-seasonal part of the effect. Namely, the non-seasonal part of the composition of the working days of the month (quarter) may be estimated by the deviation of this number of days from its long-term average. In order not to eliminate seasonality, the regressors must be calculated in deviations from the average for each month/quarter.

According to INE (2013), in order to eliminate the seasonal part of this effect the following working day regressor has to be calculated:

\[
w_{(m/t)T} - \bar{w}_{m/t} = \frac{\bar{w}}{\bar{n}} \cdot (n_{(m/t)T} - \bar{n}_{m/t})
\]

where:

- \( w_{(m/t)T} \) is the number of working days in month \( m \) or quarter \( t \) from year \( T \),
- \( \bar{w}_{m/t} \) is the average of the number of working days for each month \( m = \) January, February, ... December, or quarter = I, II, III, IV, calculated over a 28-year calendar (in shorter series, this average of working days each month or quarter is obtained with the calendar of the complete series, and is recalculated every year),
- \( n_{(m/t)T} \) is the number of non-working days each month \( m \) or quarter \( t \) in year \( T \),
- \( \bar{n}_{m/t} \) is the average of the number of non-working days in each month \( m = \) January, February, ... December, or quarter = I, II, III or IV, calculated over a 28-year calendar (in shorter series this average is recalculated every year),
- \( \bar{w} \) is the quotient between the average of the number of working days and the average of the non-working days, calculated over a 28-year calendar. In shorter series, this quotient is obtained with the calendar of the complete series, and is recalculated every year; though it may stabilize near the same value in long series.

According to equation two the quotient between the average number of the working days and the average number of the non-working days equals \( 5/2 \) and it is assumed to be constant over time.

Furthermore, Eurostat and the European Central Bank emphasize the importance of leap year correction, which can be modelled using the following zero mean variable:

\[
l_{py} = \begin{cases} 
0.75 & \text{if } t= \text{February of leap year} \\
-0.25 & \text{if } t= \text{February of non-leap year}
\end{cases}
\]
The adjustment to moving holidays aims to eliminate those values that are affected by events following a complex pattern over the years from the series. In this paper, Easter is moving holiday that most affects analyzed series. This effect is partially seasonal, since on average, it is celebrated more often in April than in March. Given that the seasonal part must be captured in the seasonal component, it must not also be eliminated with the correction of the Easter holiday effect. Namely, the regressor(s) for the Easter holiday must be built in such a way that they intend to capture the effect which this moving holiday can have on the economic series of interest. It is not possible to specify a standard way of preparing this regressor, since the repercussion that this holiday may have is very different from one series to another. For more details on this regressor please refer to INE (2013).

When building the regressor for working days, Croatian national holidays have to be taken into account. In the calculation of both working and non-working days, this matter considers the number of each of these days in each month in Croatia weighted by the specific weight that Croatia has in the economic series that is being adjusted for this effect (in our case the real retail trade turnover).

3. Alternative models for the calendar effect correction

Since the aim of this paper is to apply different methods of correction of the calendar effects of real retail trade turnover (RRTT) in Croatia and to define new explanatory variables which will give most accurate correction of RRTT time series, in this section five models which allow for different corrections of the calendar effects will be presented. Namely, the model called A1 is the basic alternative since it is highly similar to the model which some European NSI use for the correction of the calendar effect. Other alternative models generalize A1 by adding new variables for other effects (A2 and A3) or simplify it (A4 and A5) performing goodness of fit of the model and its prediction power. In table 1, abovementioned alternative models are given.

However, alternative A2 is most popular model which combines standard regression analysis with ARIMA modeling before seasonally adjustment of the series. This model is known as Reg-ARIMA model and it is a part of X-12-ARIMA technique.

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>$y_t = \beta_0 + \beta_1 PHC_t + \beta_2 WE_t + \beta_3 PLD_t + \varepsilon_t$</td>
</tr>
<tr>
<td></td>
<td>$PHC_t = PH_t - DF$</td>
</tr>
<tr>
<td></td>
<td>$WE_t$ - defined in (2)</td>
</tr>
<tr>
<td></td>
<td>$PLD_t$ - defined in (4)</td>
</tr>
<tr>
<td></td>
<td>$PH_t$ - the number of non-working days corresponding to Monday, Tuesday, ..., Friday in the month $t$</td>
</tr>
<tr>
<td></td>
<td>$DF$ - the long run average of non-working days in the month excluding weekends</td>
</tr>
<tr>
<td>A2</td>
<td>$y_t = \beta_0 + \beta_1 PHC_t + \sum_{i=1}^{t} \beta_{i+3} W_t^i + \varepsilon_t$</td>
</tr>
<tr>
<td></td>
<td>$W_t$ - defined in (2)</td>
</tr>
<tr>
<td></td>
<td>$PH_t$ - the number of non-working days corresponding to Monday, Tuesday, ..., Friday in the month $t$</td>
</tr>
<tr>
<td></td>
<td>$W_t$ - the effect of working days, trading days, Easter effect and implicitly the effect of leap year</td>
</tr>
<tr>
<td></td>
<td>$DL$ - the average number of working days in period $t$</td>
</tr>
<tr>
<td>A3</td>
<td>$y_t = \beta_0 + \beta_1 PHC_t + \beta_2 WE_t + \sum_{i=1}^{t} \beta_{i+3} W_t^i + \varepsilon_t$</td>
</tr>
<tr>
<td></td>
<td>$WE_t$ - the number of Saturdays and Sundays during the period $t$, similar to the variable $WE_t$ defined in (2) but it includes the effect of non-working days which are not weekends</td>
</tr>
<tr>
<td></td>
<td>$RDF$ - the ratio between working days and non-working days</td>
</tr>
</tbody>
</table>

Table 1: Alternative models for calendar effect correction
Model A2 analyzes the effect of trading days. Model A3 includes the leap year correction with the variable $t_{ly}$, defined by Eurostat.

According to abovementioned alternatives the one which is a parsimonious (model with one regressor) and includes effect of working days, Easter effect and implicitly the effect of leap year is alternative A5. Therefore, alternative A5 will be applied to real retail trade turnover series with modification according to equation (3). Proposed modification is based on two following assumptions:

- The ratio between average number of working days and average number of non-working days $R_{DF}$ is not constant (it should be recalculated every year).
- Saturdays and Sundays are working days of the week.

Furthermore, proposed modification of alternative A5 according to equation (3) will be compared to model defined in equation (2). The empirical results are presented in section 3.

It is rational to assume that Saturdays and Sundays are working days of the week due to retail sales in Croatia. Working on weekends, especially on Sundays became an important marketing element and usual practice in Croatia, in accordance with arrival of foreign retail chains and the formation of new national retail chains which mainly trade in nutrition products and consumer goods. Large shopping centers are becoming a destination for family trip on Sundays and skilled traders constantly organize actions and promotions in order to attract more customers.

4. Empirical analysis

The data set analyzed in this paper is series of real retail trade turnover indices in Croatia. The sample period is from January 2001 to December 2013 for a total of 156 monthly observations. According to parsimony principle two regression models are estimated using different single regressor which includes all calendar effects.

First regression model uses regressor as defined in previous equation. Results of this regression model are obtained with OLS method (Figure 1).

![Figure 1: Estimated OLS regression model with REG_1 as regressor](image)

According to figure 1 effect of additional working days (estimated coefficient of REG_1) have positive sign which means that on average retail trade turnover series should be upward corrected for 0.1277 when taking calendar effects into account.

Non-significance of this coefficient is expected due small variations that can be explained by calendar effects (Figure 3). However, the significance tests as well as other regression diagnostics are not as important as subtraction fitted values from original series to compute calendar adjusted values.
Therefore, calendar adjusted values are residuals plus 100 which is exactly the mean of retail trade turnover indices. Due to lack of space actual values (RTI), fitted values and residuals only for year 2002 are presented as a part of Figure 2, while actual values and calendar adjusted values are presented on Figure 3.

Second regression model uses newly proposed regressor defined as modification of alternative A5 according to equation (3). Results of this regression model are obtained with OLS method (Figure 4).
According to figure 4, effect of additional working days (estimated coefficient of REG_2) have negative sign which means that on average retail trade turnover series should be downward corrected for -0.096 when taking calendar effects into account. Non-significance of this coefficient is expected due small variations that can be explained by calendar effects (Figure 6). Calendar adjusted values are residuals plus 100 which is exactly the mean of retail trade turnover indices. Due to lack of space actual values (RTI), fitted values and residuals only for year 2002 are presented as a part of Figure 5, while actual values and calendar adjusted values are presented on Figure 6.

In figure 7, two series of calendar adjusted values are compared with original series of retail trade turnover indices in Croatia (due to lack of space only results for 2001 – 2006 are presented). The first series of calendar adjusted values (CA_1) are obtained according to residuals of first regression model with REG_1 regressor. The second series of calendar adjusted values (CA_2) are obtained according to residuals of second regression model with REG_2 regressor. Regressor REG_2 is newly proposed as modification of alternative A5.
Table 2: Two series of calendar adjusted values (CA_1 and CA_2) compared with original series of retail trade turnover indices for first six years of observation

For example, significant deviation between calendar adjusted values 1 and 2 is evident for October 2002. At this month the value of original series of retail trade turnover index is 99,03; the calendar adjusted value 1 is 97,61; the calendar adjusted value 2 is 100,65; and the difference between calendar adjusted values is 3,1%.

This difference can be explained that in October 2002 there is no non-working days (public holidays), the number of working days in October are maximal (31 days), which included Saturdays and Sundays (CA_2).

5. Conclusion

Economic activity fluctuations are often influenced by various factors related to the calendar. These factors include non-working (non-trading) days, leap years, public holidays etc. In analysis of many economic variables in which time series are seasonally adjusted it is necessary to identify and correct calendar effects using suitable method, but there is no general or unique procedure for correcting these effects in pre-adjustment process. In this paper different methods of correction of the calendar effects are compared and applied to time series of real retail trade turnover (RRT) in Croatia (monthly data observed from January 2001 to December 2013). The most common method used is regression model with different types of explanatory variables which take into account calendar effects. The contribution of this paper is to define new explanatory variables (regressors that include not only different number of working and non-
working days of the month but also country specific calendar effects) which will give most accurate correction of RRTT time series.

Estimation results show that newly proposed regressor REG_2, as modification of alternative A5 according to equation (3) is better alternative for calendar adjustment of retail trade turnover series. This can be justified by taking into account the two important assumptions:
(a) the ratio between average number of working days and average number of non-working days RDF is not constant (it should be recalculated every year) and
(b) Saturdays and Sundays are working days of the week.

This analysis also indicates that the detection and correction of the calendar effect should consider other factors which are not included in the models of the five analyzed alternatives.

References
Di palma, f., marini, m.: the working/trading day adjustment of italian quarterly National accounts: methodology and presentation of the main result. In proceeding of joint Unece/eurostat/oecd meeting on national accounts, geneva, ces/ac.68/2004/12, 2004.
Eurostat commission, theme 4, industry trade and services, 2002b.