Seasonal Adjustment of the Spanish Sales Data

Keywords: Daily Time Series, Seasonal Adjustment, Unobserved Components, Kalman Filter, Calendar Effects.

1. Introduction

Daily economic time series pose several and difficult challenges due to the coexistence of multiple seasonal components, the complex structure of their calendar effects, the strength of their irregular component and their sensitivity to exogenous factors (e.g. outliers) that distort its usual behavior.

We use a structural model of unobserved components to perform modeling and seasonal adjustment of the Spanish sales time series. This model, complemented with a suitable dynamic-regression pre-processing, provides a flexible although parsimonious way to handle the complex nature of daily time series.

We will briefly present the modeling approach as well as the main results when applied to the total sales data as an example that represents its application to the breakdown of daily sales data (~25 supply-side components).

2. Methods

We have followed the approach proposed by [1], called TBATS: Trigonometric seasonality, Box-Cox transformation, ARMA innovations, Trend and Seasonality. The TBATS approach is based on the representation of the unobserved components (trend, seasonality, and irregularity) by means of explicit dynamic models, see [2].

This approach is very flexible, since it allows the simultaneous consideration of several seasonal components (weekly, monthly, and annual) and, thanks to its trigonometric representation by means of Fourier expansions, it enables the treatment of fractional periodicities.

Following the structural approach, the model incorporates a parsimonious but rather general representation of the trend. It also includes an explicit ARMA model for the irregular component that acts as a sort of “safety valve”, accommodating elements that, for whatever reason, did not find a proper fit within the basic systematic components (trend and seasonality). In this way, the plain representation of these two components does not compromise neither the fit of the model to the sample nor its forecasting performance.

TBATS modeling may include a preliminary intervention analysis by means of exogenous deterministic variables designed to control for the presence of outliers and specific calendar effects that, due to their moving or fractional periodicity nature, do not fit into the structural representation considered by TBATS, see [3] and [4].

An interesting feature of TBATS is that it can handle complex seasonal patterns, comprising both multiple periodicities (weekly, monthly and yearly) and fractional periodicities (e.g. 30.4375 days for monthly seasonality or 365.25 for annual seasonality).

The TBATS procedure sets the model in state space form, computes its likelihood and maximizes it using the model parameters as instruments. It also determines the most appropriate Box-Cox transformation and, once applied, the proper number of harmonics for the seasonal subcomponents, starting with j=1. In all the cases, the different combinations are ranked according to the Akaike information criterion (AIC) and the one that minimizes AIC is chosen.

Finally, TBATS performs a search for the most adequate ARMA(p,q) model for the innovation, starting with a white noise (p=q=0). If the innovation fails to be considered as a white noise, a search along p and q is implemented, selecting the combination that minimizes the AIC.

3. RESULTS

3.1. Data

Sales data, valued at current prices are derived from the Vax Added Tax (VAT) system, administered and implemented by the Spanish Tax Agency.

This system (*Sistema Inmediato de Información*, SII ~ Instantaneous Information System) is a fully digital procedure that allows Spanish firms to comply their tax duties on a daily basis, according to their invoicing practices.

The current, fully operational version of the system started in July, 1th 2017. The previous system (called “Registry Book”) dates back to January 1th 2010 and it was less timely and complete.

The system operates as a reservoir, so it has (usually upward) revisions that tend to stabilize around 10 days after the reference date, covering around 80% of the Spanish total firms turnover.

3.2. Empircal results

Preliminary data analysis suggests the convenience to preprocess the time series from working-days and calendar effects. The holiday variable is based on the official working calendar, including national and regional holidays. In this way, we get 20 daily time series (1 national, 17 regional and 2 for the autonomous cities). We have built a single regression variable by combining the 20 time series according to its weight on the distribution of interior sales as reported by the 56 offices of the Tax Agency. The role of this variable is enhanced if we restrict it to be binary, when it has a national coverage of at least 2/3.

The effects linked to the deterministic part of the monthly seasonality are collected using three binary variables that separately consider whether the day is the beginning of the month, the 15th day, or the end of month, adopting the value 1 in this case and 0 otherwise. These three effects interact with the weekly seasonality, for which three additional binary variables are considered that adopt the value of 1 if, in addition to being the beginning of the month, the 15th day or the end of the month, the day is also a weekend (Saturday or Sunday). The following table shows the results of the estimation of the deterministic effects, by means of an extended regression model, see [1]. The following table shows the results of the estimation of the deterministic effects.

Table 1. Estimation of deterministic effects

		Monthly component					
		Basic effect			Interaction with the weekends		
	Holiday	End of month	Beginning of month	15th day	End of month	Beginning of month	15th day
β	-1.40	1.67	0.50	0.40	1.24	0.65	0.67
t(β)	-34.00	31.44	9.17	8.11	16.30	8.63	8.84

Note: The number of harmonics associated with each seasonal component (weekly, monthly and annual) is 3, 9 and 5, respectively. These values are derived from the preliminary estimate of a TBATS model applied to the original time series.

The decomposition of the "linearized" time series by means of TBATS is shown in the following figure.

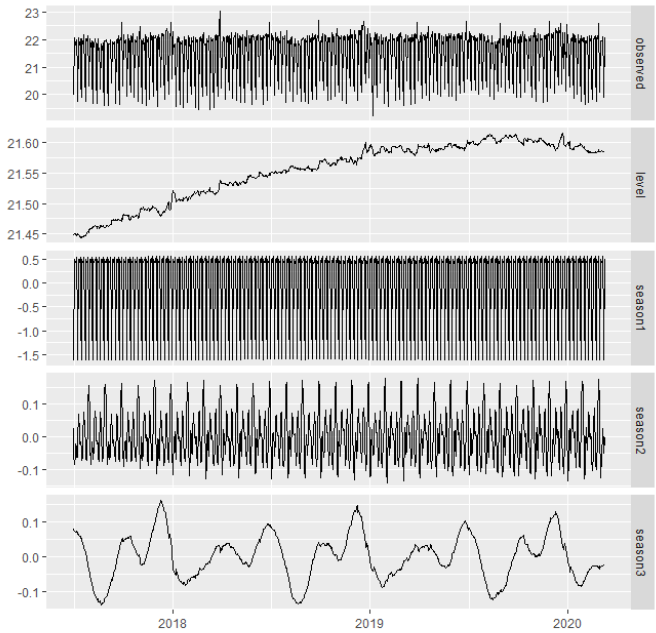

Figure 1. Decomposition of the linearized Spanish daily sales time series

The complete monthly seasonal component includes a (dominant) deterministic effect, as can be seen in the next figure.

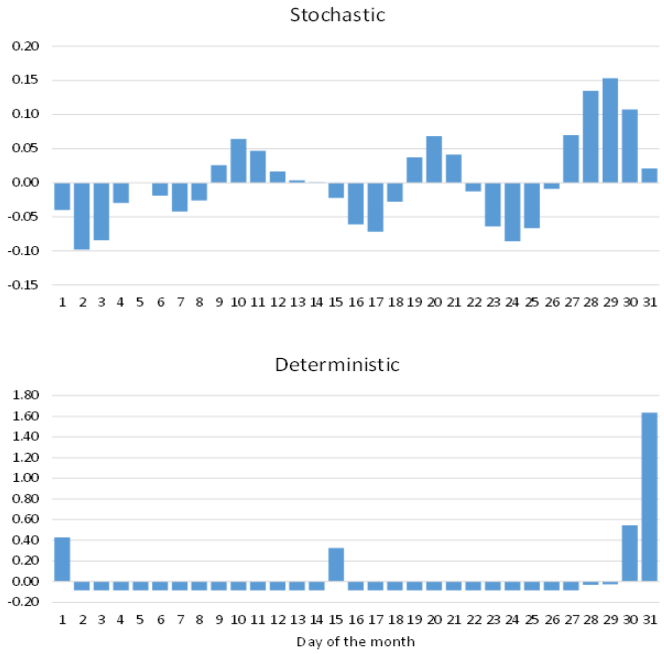

Figure 2. Estimation of the monthly seasonal component (mean profile).

4. CONCLUSIONS

The approach based on TBATS allows us to perform a seasonal adjustment of the Spanish daily sales data. This approach requires a preliminary step to treat effects with moving periodicity (notably, "end of the month" effect). Currently, we apply it to a set of 30 daily time series that comprise a sectorial breakdown of sales and several intermediate aggregates.

In addition to dealing with special adjustments required by the outlying nature of the COVID-19 shock, we are preparing a robustness check, following [5] and [6].

REFERENCES

[1] A. De Livera, R.J. Hyndman and R.D. Snyder, "Forecasting Time Series with Complex Seasonal Patterns Using Exponential Smoothing", *Journal of the American Statistical Association*, 106(496), (2011), 1513-1527.

[2] A.C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, (1989).

[3] S.C. Hillmer and W.R. Bell, "Modeling Time Series with Calendar Variation", *Journal of the American Statistical Association*, 78(383), (1983), 526-534.

[4] S.C. Hillmer, W.R. Bell and G.C. Tiao, "Modeling Considerations in the Seasonal Adjustment of Economic Time Series". In: Zellner, A. (Ed.) *Applied Time Series Analysis of Economic Data*, p. 74-100. US Bureau of the Census, (1983).

[5] D. Ladiray, J. Palate, G.L. Mazzi and T. Proietti, "Seasonal adjustment of daily and weekly data". In: D. Ladiray and G.L. Mazzi (Eds.) *Handbook on Seasonal Adjustment*, Eurostat, (2018).

[6] D. Ollech, "Seasonal adjustment of daily time series", Deutsche Bundesbank, Discussion Paper n. 41/2018, (2018).