PAPELES DE TRABAJO

5/2021

Modeling daily sales data(*)

ÁNGEL CUEVAS

RAMIRO LEDO

ENRIQUE M. QUILIS

Servicio de Estudios Tributarios y Estadísticas. Agencia Tributaria

(*) We thank Javier Abelaira, Rafael Frutos, Francisco Melis, Francisco Pizarro, Jorge Polo and Silvia Relloso for their input at different stages of this Project and seminar participants at the Tax Agency, Spanish National Statistical Institute, Bank of Spain, Eurostat and European Central Bank. The opinions are those of the authors and do not necessarily reflect the views of the Spanish Tax Agency.

Avda. Cardenal Herrera Oria, 378, 28035-Madrid

I.S.S.N.: 1578-0252

N.I.P.O.: 188-21-012-7

INDEX

Resumen

Presentamos un procedimiento para la modelización y el ajuste estacional de datos de frecuencia diaria, aplicado a las series españolas de ventas. El modelo ajusta la información del Sistema de Suministro Inmediato de Información (SII) basado en los formularios del Impuesto sobre el Valor Añadido (IVA) elaborados por la Agencia Tributaria. El procedimiento realiza la extracción de señales en frecuencia diaria mediante un modelo estructural de componentes no observados. La información diaria, debidamente procesada, permite un seguimiento permanentemente actualizado de las condiciones económicas a corto plazo de la economía española.

Palabras clave: Información tributaria, series temporales diarias, análisis de coyuntura, ajuste estacional, modelo estructural de componentes.

Códigos JEL: C22, C32, C53.

Abstract

We present a procedure to perform seasonal adjustment using daily sales data. The model adjusts daily information from the Immediate Supply of Information System for Value Added Tax declaration forms compiled by the Spanish Tax Agency. The procedure performs signal extraction at the daily frequency by means of an unobserved structural components model. The daily information, properly processed, allows a real-time monitoring of the short-term conditions of the Spanish economy.

Keywords: Tax Data, Daily Time Series, Short-Term Monitoring, Structural Components Model.

JEL Codes: C22, C32, C53.

1. INTRODUCTION

Economic data provided by tax sources are gaining popularity among economic analysts and forecasters due to its timely availability, reliability, coverage and direct economic meaning. In this way, tax-based data have a clear function in the design of a model aimed at nowcasting and short-term forecasting.

Updated and reliable forecasts play a critical role for budgetary planning and for the anticipation of risky situations due to adverse shocks. In particular, the daily information provided by the new Immediate Supply of Information system (SII, *Sistema Inmediato de Información*) based on the Value Added Tax (VAT) forms and developed by the Tax Agency, opens new perspectives for the integration, on a real-time basis, of tax-based reliable quantitative information with macro data observed at lower frequencies.

This paper presents a system for seasonal adjustment of very high frequency data derived from tax sources. Specifically, the model treats the daily sales time series provided by the SII. Seasonal adjustment (SA) of daily data is necessary, in the same way as happens to monthly and quarterly data, in order to provide a meaningful signal of its underlying evolution. However, this task is notably more difficult than in the case of monthly or quarterly data due to the complexity of its seasonal component, formed by several sub-components (some of them linked to fractional periodicities) and its noisy nature (Ladiray *et al*., 2018).

The paper is organized as follows. The second section presents the main characteristics of the data. The third section develops the econometric methodology, which is implemented in two stages. In the first one, we apply a preliminary treatment of the deterministic effects and, in the second one, we use an univariate structural model, based on an unobserved component representation, to decompose and forecast the daily sales data. The empirical results are presented in the fourth section. The paper ends presenting the main conclusions and future developments.

2. DATA

Daily sales series of monthly VAT taxpayers comes from the Immediate Supply of Information System (SII, *Sistema Inmediato de Información*), introduced in January 2017 and officially implemented by the Spanish Tax Agency since July 2017 (Tax Agency, 2017).

This system allows the exchange of tax information between the Spanish Tax Agency and taxpayers required by the SII practically in real time, by supplying the detail of the invoicing records within four days, through the electronic platform of the Spanish Tax Agency.

In this way, both tax management and tax compliance are improved (e.g. by the taxpayers comparison of the information in their books with the information provided by their customers and suppliers).

The group compulsorily included in the SII is made up of all those taxpayers whose obligation to declare VAT is monthly:

- Large Companies (turnover greater than 6,010,214.04€ in the previous year).
- Those companies that pay taxes through the special regime for groups of companies[1].
- Registered, voluntarily, in the Monthly VAT Return Registry.

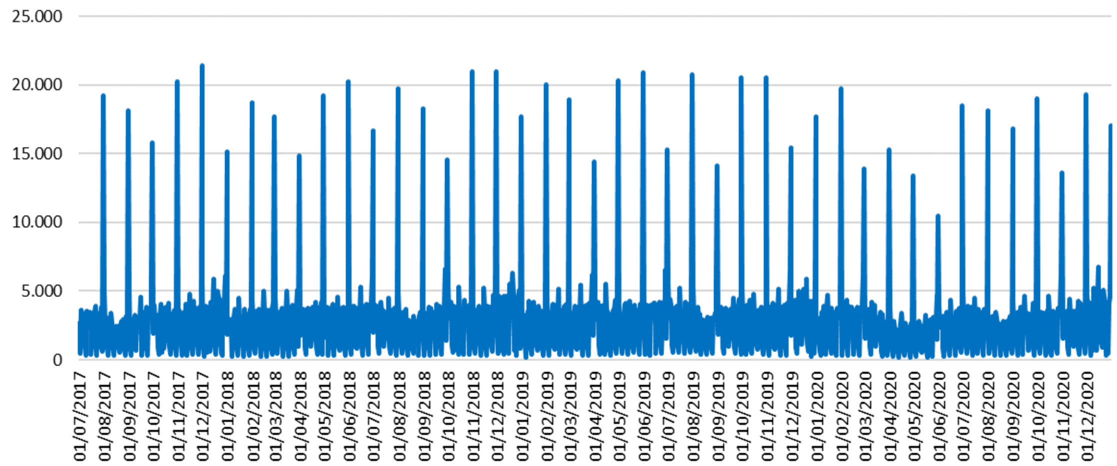
In addition, this system is applied to those taxpayers who voluntarily adopt it.

Thus, the SII comprises about 63,000 taxpayers representing around 70% of the country's total business turnover, with a great diversity of coverage by activities (see Annex 1 for more details).

Figure 1 shows the daily time series, from July 1st, 2017 to December 31th, 2020. The sample comprises from the full-fledged implementation of the SII until the end of 2020. At first glance, it is possible to see the large volatility of the series, due to the different invoicing patterns, that we will comment on later, such as the large effect of the end of the month, which makes very difficult to distinguish a relevant signal of underlying evolution.

Figure 1

Daily sales from SII (million €)



As a general rule, firms must send their invoices four days after they have been issued. However, the information received by the Tax Agency is not always the final one, as it can be completed and corrected in the following days. Experience indicates that it is necessary to wait around two weeks for the stabilization of the levels that allows us to consider the data as definitive.

3. ECONOMETRIC METHODS

In this section we present the econometric methodology used in the paper. Modeling economic daily time series poses several and difficult challenges due to the coexistence of multiple seasonal

[1] A group of companies is considered when it is formed by a parent company and its subsidiaries. They must be firmly bound to one another.

components linked to various frequencies, the complex structure of the calendar[2], the strength of its irregular component and its sensitivity to exogenous factors (e.g. outliers) that distort its usual behavior (Ladiray *et al.,* 2018).

We use the structural model of unobserved components proposed by De Livera *et al*. (2011) to perform the modeling and seasonal adjustment of the Spanish daily sales data. The model, called TBATS (acronym for Trigonometric seasonality, Box-Cox transformation, ARMA innovations, Trend and Seasonality), complemented with a suitable dynamic regression pre-processing, provides a flexible although parsimonious way to handle the complex nature of daily time series. The econometric methodology has two steps:

- Pre-processing (linearization). In this step we apply an 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 nature, do not fit well into the structural representation considered by TBATS (Hillmer and Bell, 1983; Hillmer *et al.,* 1983). This intervention analysis is a preprocessing step of the observed series that renders it suitable for TBATS.

- Structural decomposition using TBATS. The pre-processed time series is decomposed into trend-cycle, seasonality and irregularity. As we will expose below, the seasonal component has a complex nature due the coexistence of multiple seasonal patterns, some of them with fractional periodicities.

Let us now explain both steps with some detail.

3.1. Pre-processing (linearization)

The rank-mean analysis of the time series, as well as its wide range of variation, suggest the convenience of using a preliminary logarithmic transformation. In the same vein, the estimation of the λ parameter of the Box-Cox transformation performed by TBATS is very close to zero[3], providing additional support for the log-transformation.

The analysis of the residuals both from the structural TBATS model and from a reduced-form model (Cuevas *et al*., 2019) confirms the need to control for some effects linked to: (i) bank holidays, (ii) the inexact periodicity of the monthly seasonal component and (iii) its interaction with the weekly seasonal component.

These effects are represented by deterministic variables and their impact on the observed time series is estimated using a regression model that includes a linear trend and a multiple seasonal component affine to the one used in the trigonometric seasonal representation used by TBATS. The regression model is:

---

[2] The most notable effects are: different length and composition of the months, moving seasonality (e.g. Easter), leap years and a time-varying calendar of working days that interacts with the composition of the months.

[3] The exact value is 0.0039.

[1] $$z_t = \beta x_t + \alpha_0 + \alpha_1 t + \sum_{i=1}^{3} \sum_{j=1}^{k_i} \left[\gamma_j \sin\left(\frac{2j\pi t}{m_i}\right) + \varphi_j \cos\left(\frac{2j\pi t}{m_i}\right) \right] + e_t$$

Being:

- z_t: (log-transformed) observed variable.
- x_t: m deterministic (dummy) variables linked to the bank holidays and to the inexact periodicity of the monthly seasonality.
- m_i: periodicity of the i-th seasonal component. Based on a preliminary analysis (Cuevas et *al.*, 2019) we have considered three components (weekly, monthly and yearly) whose periodicities[4], expressed in days, are: 7, 30.4375 and 360.25.
- k_i: number of harmonics of each seasonal component.
- e_t: Gaussian error term.

It is interesting to note that the regression model [1] can be considered as a one-equation approximation to the complete structural TBATS model that will be presented below, especially due to its similar treatment of the (multiple) seasonality. This similitude enhances the complementarity of steps 1 and 2.

The number of harmonics, k_i, associated with each seasonal component (weekly, monthly and annual) can be determined by means of a preliminary estimate of the TBATS model applied to the original time series.

3.2. Structural (TBATS) decomposition

The TBATS approach is based on the representation of the unobserved components (trend, seasonality, irregularity) by means of explicit dynamic models (Harvey, 1989)[5].

Following the structural approach, the model incorporates a parsimonious but rather general representation of the trend. It also includes an explicit 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.

The TBATS model assumes that the (possibly Box-Cox transformed) observed series (z_t) results from the aggregation of three unobserved components: trend (p_t), seasonality (s_t) and a stationary inno-

[4] The periodicity of the monthly seasonal component takes into account both the different length of the months and the leap years. The fractional periodicity of the annual seasonality is only due to the presence of leap years.

[5] The alternative approach, based on reduced-form models, offers greater flexibility to include exogenous variables but does not provide an estimate of the underlying components (Espasa *et al.*, 1996; Liu, 2005). Although less complete, we have also used this approach as a cross-check (Cuevas *et al.*, 2019).

vation (u_t). This innovation plays an additional role as the stochastic input for the other two components. In this way, both the trend and the seasonality depend on a single shock that, properly scaled and filtered, generates them. In general, we assume that u_t evolves according to a stationary and invertible autoregressive and moving average (ARMA) model:

[2] $$\left(1-\phi_1 B-\cdots-\phi_p B^p\right)u_t=\left(1-\theta_1 B-\cdots-\theta_q B^q\right)e_t$$

The ultimate shock e_t is a Gaussian white noise:

[3] $$e_t \sim iid\ N(0, v)$$

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).

This complex seasonal pattern is one of the most important differences between a daily time series and its monthly/quarterly counterpart. In the latter case, the seasonal component is unique and of integer periodicity (12 months and 4 quarters, respectively). Of course, this additional complexity requires an additional layer of specific modeling.

Assuming that there are *I* seasonal components of different periodicity, total seasonality is the sum of all of them:

[4] $$S_t = \sum_{i=1}^{I} S_t^{(i)}$$

Each seasonal subcomponent is linked to a basic frequency and with k of its harmonics, according to the following equation:

[5] $$w_j^i = \frac{2\pi j}{m_i}$$

Being:

- $w_{i,j}$ is the frequency of the j-th harmonic linked to the i-th seasonal subcomponent.
- m_i is the periodicity, in time units, of the seasonal subcomponent (e.g., 7 days for the weekly seasonality).

In this way, the seasonality associated with each basic frequency is obtained by adding the signals associated with that basic frequency and its k harmonics:

[6] $$S_t^{(i)} = \sum_{j=1}^{k_i} S_{j,t}^{(i)}$$

These individual terms are determined according to a bivariate vector autoregressive (VAR) process that includes S and an auxiliary factor Q[6].

[7] $$\begin{bmatrix} S_{j,t} \\ Q_{j,t} \end{bmatrix} = \begin{bmatrix} \cos(w_j) & \sin(w_j) \\ -\sin(w_j) & \cos(w_j) \end{bmatrix} \begin{bmatrix} S_{j,t-1} \\ Q_{j,t-1} \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} u_t$$

[6] Q is a shifted version of S that helps to complete the state space representation [6].

Equation [7] is a deterministic Fourier expansion centered on the frequency [5], stochastically perturbed by the common innovation of the system, u_t. This innovation is scaled using the parameters γ_1 and γ_2. This representation stands out for its parsimony, since only four parameters are involved: two scale parameters and two initial conditions, regardless of the time scale of the seasonality, which can be very large: greater than 28 and 360 periods in the monthly and annual case, respectively.

In this way, for each seasonal subcomponent (e.g. weekly), k first-order VAR representations are defined, as many as harmonics are needed to represent it.

The magnitude of the scale parameters determines the proximity to a deterministic behavior of the seasonal subcomponent. At the limit, if both are zero, the component is completely deterministic.

Finally, the trend p_t is a random walk, I(1), with a first-order autoregressive, AR(1), drift:

[8] $$p_t = p_{t-1} + \phi g_{t-1} + \alpha u_t$$

Being:

- g_t: drift.
- Φ: damping parameter that controls for the impact of the drift on the trend. In general, $0 \leq \Phi \leq 1$.
- α: scale parameter that modules the impact of the innovation on the trend. As a rule, $\alpha \geq 0$.

The next equation defines the drift:

[9] $$g_t = (1-\phi)b + \phi g_{t-1} + \beta u_t$$

Being:

- b: location parameter that represents the steady state of the drift, provided that $\Phi<1$.
- β: scale parameter that modules the impact of the common innovation on the drift. In general, $\beta \geq 0$.

Equations [8] and [9] provide a parsimonious yet flexible representation for the trend. In this way, depending on Φ we can get an I(2) or an I(1) trend. If $\Phi=1$ we obtain an IMA(2,1) trend. In addition, the scale parameters determine the closeness to a deterministic behavior. As a special case, if $0<\Phi<1$ and $\beta=0$ we get a random walk with a constant drift. If, in addition, $\alpha=0$, the trend becomes completely deterministic.

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.

4. EMPIRICAL RESULTS

We turn now to the empirical results derived from the application of the methodology presented in the previous section, following its two-step scheme.

4.1. Modeling aggregate daily sales

The first step consists of estimating equation [1]. Let us now describe the exogenous variables considered in x_t. The bank holidays 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 (Cuevas *et al*., 2019)[7]. The role of this variable is improved if we restrict it to be binary, setting 2/3 as the threshold.

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 the month, adopting the value 1 in this case and 0 in the rest. 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). Table 1 shows the results of the estimation of the deterministic effects, by means of the regression model [1].

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.

It is worth noting the strength of the holiday effect as well as the impact of the end-of-the-month effect, in itself and in its interaction with the weekly seasonality. The beginning-of-the-month effect and day-15th effect, although significant, are of a lesser magnitude.

[7] An alternative estimate, using weights from the Spanish Regional Accounts, yields similar results.

The daily time series, corrected from the deterministic effects quantified in the previous table, is decomposed by means of the TBATS model. The next table presents the estimation of its parameters.

Table 2

Estimation of TBATS parameters

Seasonality								
Weekly			Monthly			Annual		
<m, k>	γ_1	γ_2	<m, k>	γ_1	γ_2	<m, k>	γ_1	γ_2
<7, 3>	-0.0003	0.0020	<30.4375, 7>	-0.0022	0.0014	<365.25, 6>	-0.0033	0.0010
Trend			Innovation: ARMA(p,q)					
Φ	α	β	p	q	σ	Likelihood	AIC	n
0	0.012	0	0	0	0.4129	3473.39	3553.39	983

Note: <m,k> denote the periodicity and the number of harmonics for each seasonal component, respectively.

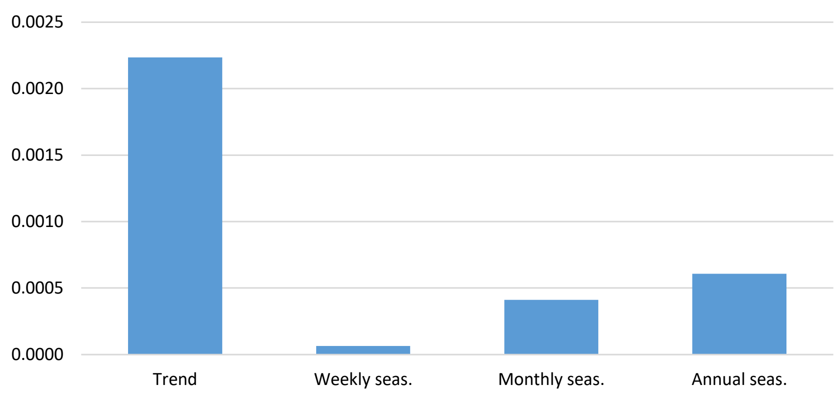
The estimated parameters indicate that all the seasonal components are quite stable, especially the one related to the weekly seasonality. This stability does not prevent the estimated components from being relatively complex, given the high number of harmonics required to represent them.

The estimated trend is a random walk whose specific innovation is much less volatile than the corresponding to the common innovation u_t.

From the previous table, we can ascertain that the common innovations u_t that affect the series are mainly reflected in its trend and its annual seasonality. Monthly seasonality and, above all, weekly seasonality are relatively immune to the shocks. This ranking is an alternative way to quantify the proximity to a deterministic scheme and is shown in the figure 2.

Figure 2

Transmission of an innovation to the components

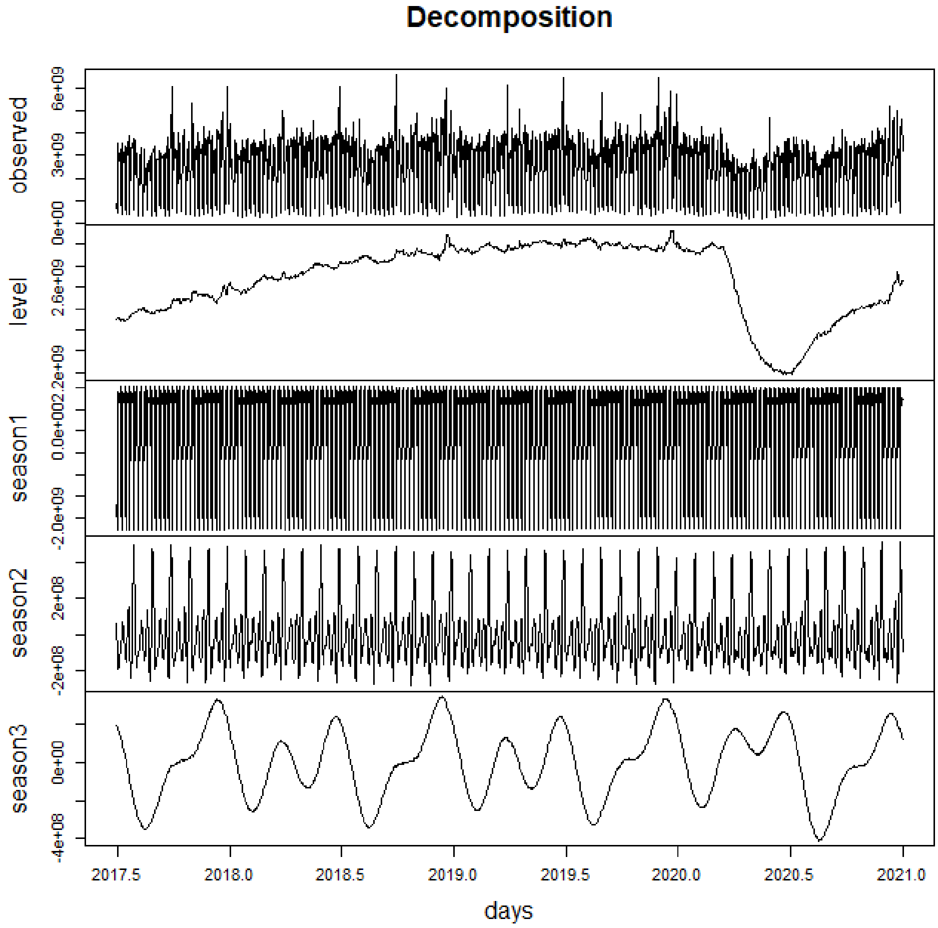


Note: The units are fractions of the standard deviation of the innovation.

The TBATS model, whose parameters are shown in the table 2, allows the estimation of the unobserved components that underlie the observed time series by means of the Kalman filter. These components are presented in the figure 3.

Figure 3

Stochastic decomposition of the daily SII sales

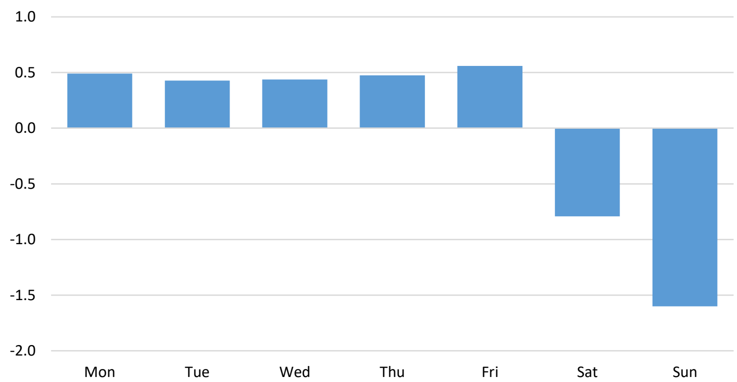


Note: level refers to trend, season1 refers to weekly seasonality, season2 refers to monthly seasonality and season3 refers to annual seasonality.

The information provided by this decomposition makes it possible to infer, in the first place, the weekly seasonal profile, the most stable component of all. As can be seen in the following graph, the weekly pattern of daily sales shows a strong contrast between the weekend, especially Sunday, and the rest of the week.

Figure 4

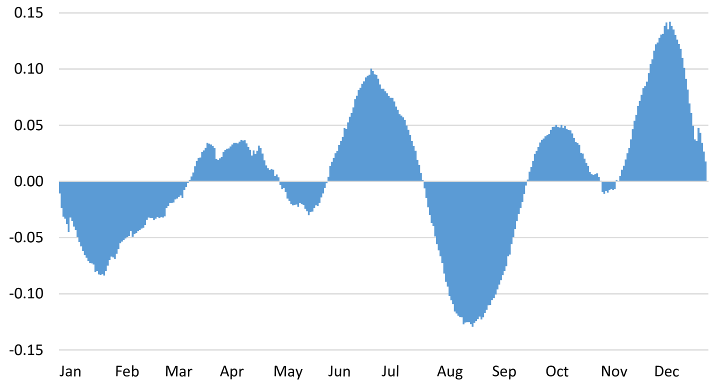
Mean profile of the weekly seasonality



In the same vein, we can check the mean profile of the annual seasonality. As can be seen in the next figure, its pattern is relatively complex, comprising several peaks and troughs. Among the former, the days that belong to December, June and July stand out. On the other hand, the trough days are registered in August, January and February.

Figure 5

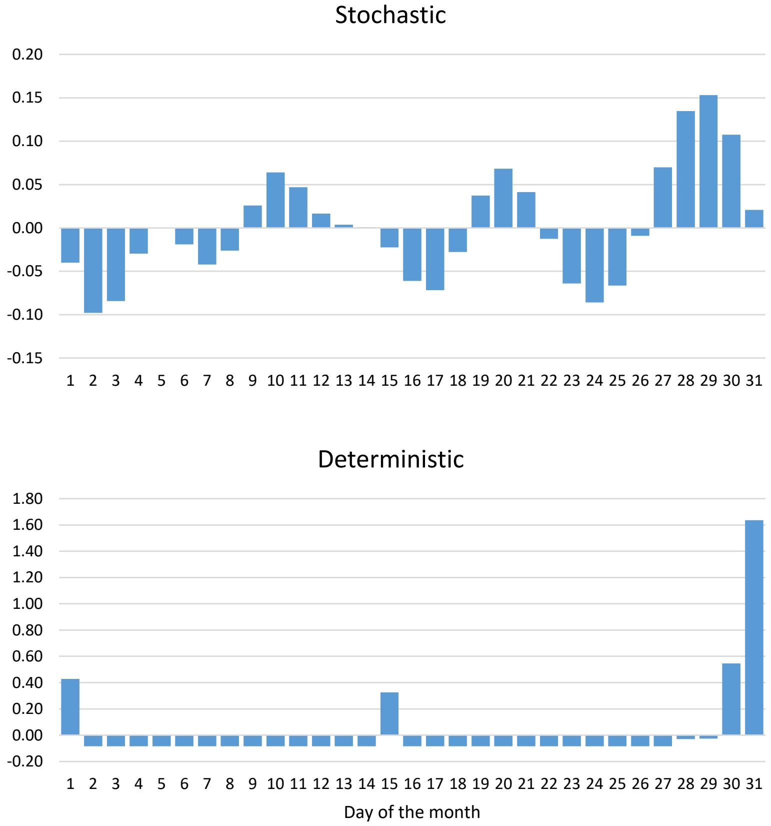
Mean profile of the annual seasonality



To fully analyze the monthly seasonal profile, it is necessary to combine the stochastic component estimated by TBATS, as shown in figure 3, with the estimation of the deterministic effects of monthly periodicity, (beginning-of-the-month, day 15th and end-of-the-month effects), presented in table 1. These effects, properly centered to average zero in the long term, are shown in the following graph.

Figure 6

Mean profile of the monthly seasonality



The main features of the monthly seasonality are, first and above all, the preeminence of its deterministic component, the important role of the end-of-the-month effect and, finally, the lesser relevance of both the beginning-of-the-month and the 15th day effects.

One additional feature of the TBATS model is that can be used to forecast the daily sales in order to have permanently filled the current month. Of course, the forecasts can be extended as much as required if, for whatever reasons, more distant forecasts are required.

4.2. Results by activity breakdown

As mentioned before, the firms included in the Immediate Information Supply system (SII) are classified according to the activity declared by their own companies. Specifically, the activities are classified according to the breakdown of the Economic Activity Tax (IAE, *Impuesto de Actividad Económica*), which has a correspondence with the standard NACE-2009 classification (see Annex 1 for details on its coverage).

In this way, the methodological approach previously described can be applied to all the branches of activity considered. For simplicity we are going to show the most significant results for total sales and the main branches of activity. Additionally, we are going to focus in the recent recessive episode due to the COVID-19 health crisis, because the extraordinary and unusual nature of this shock deserves special attention, showing how it has affected the different economic activities.

Figures 7 and 8 show, respectively, the final series corrected from deterministic effects and seasonal and calendar components, and their corresponding year-on-year rate growth. We have used a 28 days moving average filter to render the levels comparable with those of a monthly time series.

Figure 7

Final corrected series (MA-28 levels)

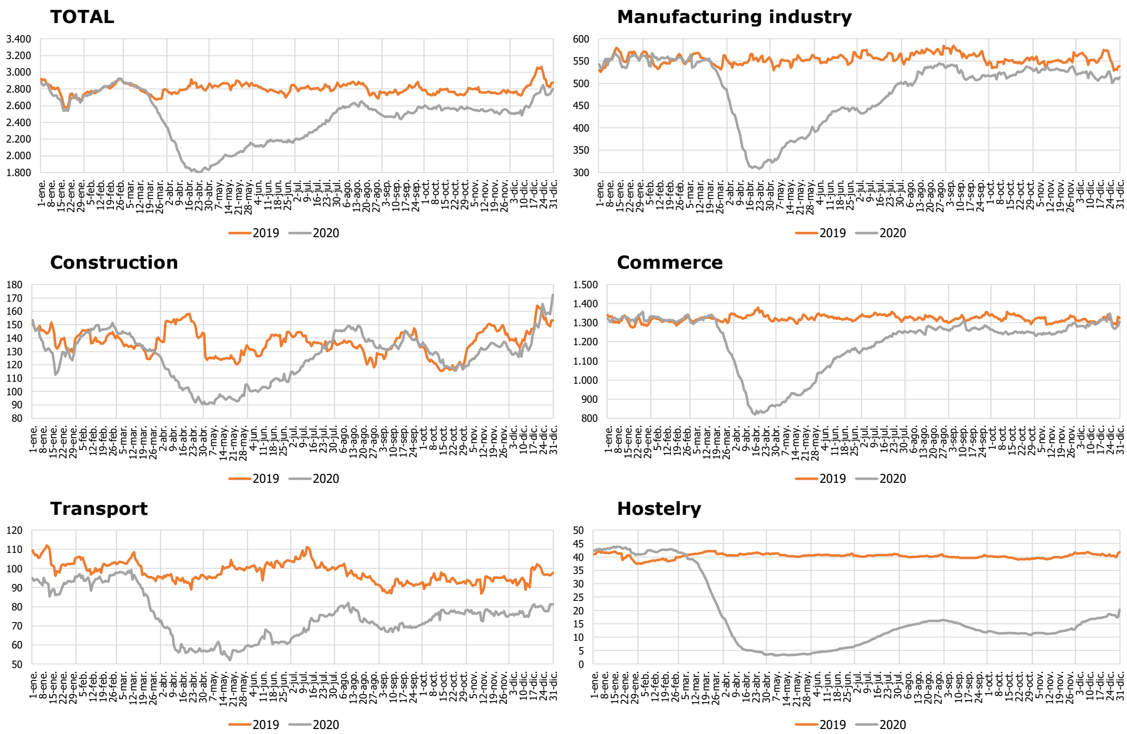
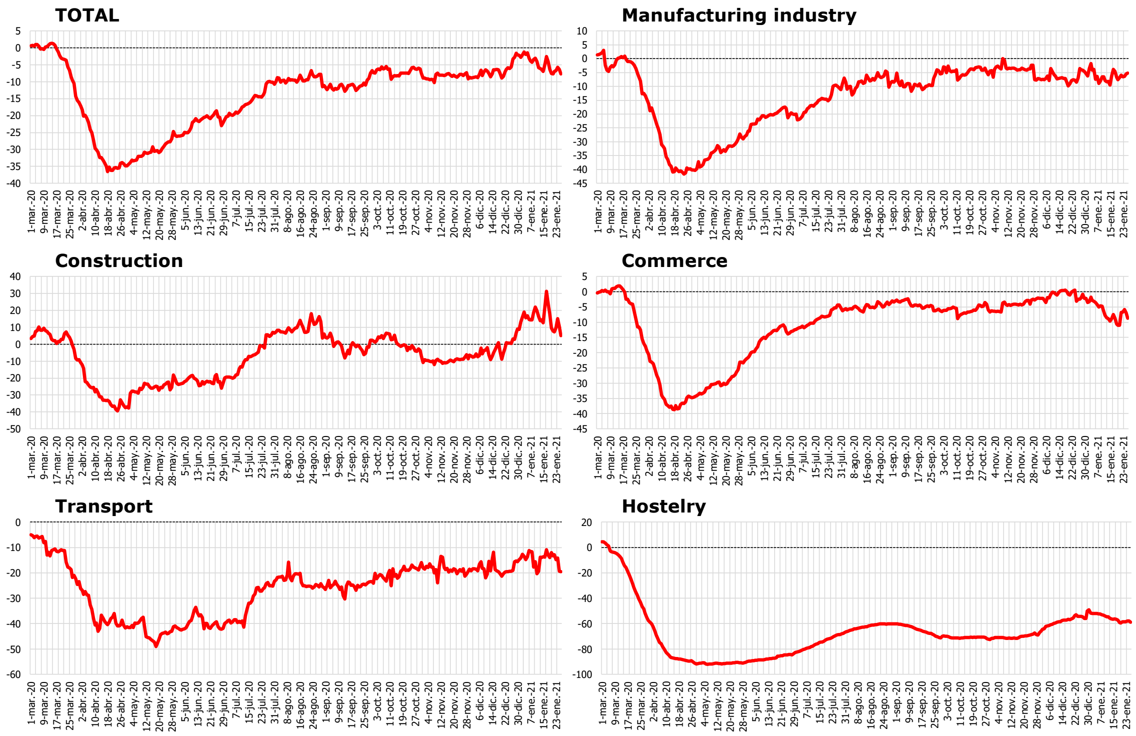


Figure 8

Final corrected series (MA28 levels): Year-on-year (YoY) growth rate



The adjustment carried out allows us to extract a clear signal, useful for the monitoring of sales in each sector on a real-time basis. Likewise, the decline in activity caused by the impact of the COVID-19 shock has been clearly uneven by activity branch. The greatest impact caused by the lockdown and related policies have had the largest impact on the hostelry and transport sectors, with falls that remain at 60% in the first case.

5. CONCLUSIONS

The use of structural, unobserved components time series model has been successful, allowing us to extract non-seasonal information useful for the short-term monitoring of the Spanish economy.

From a purely statistical view, tax-based data stands out as an inexpensive and reliable way to use the information routinely provided by the economic agents when complying with their tax duties, replacing the use of costly sample surveys or qualitative information derived from data sources whose representativeness cannot be properly assessed. Of course, part of this reliability comes from the positive reinforcement due to penalties and fines if compliance is not perfect. Few data sources are backed by this reinforcement schedule.

For future work, we plan to compare TBATS with STLM in the realm of seasonal adjustment, in order to ascertain the performance of model-based methods (TBATS) vs non-parametric filtering methods (STLM), see Ollech (2018).

ANNEX 1: BREAKDOWN OF SALES

The branches of activity that can be considered, with their corresponding coverage in 2019, comprise the following sections of the NACE:

	Number of		Domestic sales (mill. €)		Percentage of domestic sales of SII companies with respect to the total
	SII	Total	SII	Total	
TOTAL DOMESTIC SALES	57.855	3.796.021	1.213.602	1.730.463	70,1
C. Manufacturing industry	11.499	175.039	245.057	307.734	79,6
C10. Food industry	2.605	25.819	64.929	75.588	85,9
C11 + C12. Manufacture of beverages and tobacco	703	5.775	12.563	14.754	85,2
C13 + C14 + C15. Textile industry, garment manufacturing and leather and footwear industry	935	22.910	6.545	11.254	58,2
C17 + C18. Paper industry; graphic arts	559	15.844	10.537	15.256	69,1
C20. Chemical industry	810	3.866	21.475	24.215	88,7
C21. Manufacture of pharmaceutical products	233	473	12.225	12.867	95,0
C22 + C23. Manufacture of rubber and plastic products and other non-metallic mineral products	1.256	11.866	19.336	25.477	75,9
C24 + C25. Metallurgy; manufacture of iron, steel, ferro-alloy and metal products, except machinery and equipment	1.686	33.920	23.679	36.751	64,4
C26 + C27. Manufacture of computer, electronic and optical products; manufacture of electrical material and equipment	457	4.158	7.655	9.147	83,7
C29. Manufacture of motor vehicles, trailers and semi-trailers	370	1.635	30.628	33.859	90,5
C16 + C31. Wood and cork industry; furniture manufacturing	604	22.114	5.086	10.533	48,3
C28 + C30 + C33. Manufacture of machinery and equipment; manufacture of other transport equipment; repair and installation of machinery and equipment	1.063	19.290	10.806	17.202	62,8
C19 + C32. Coke and refined petroleum products: other manufacturing industries	218	7.369	19.595	20.831	94,1
D. Supply of electricity, gas, steam and air conditioning	760	29.695	36.647	49.318	74,3
F. Construction	6.959	409.901	61.406	123.140	49,9
F41. Building construction	4.114	235.355	28.841	68.620	42,0
F42. Civil Engineering	679	12.989	9.603	14.396	66,7
F43. Specialized construction activities	2.166	161.557	22.963	40.124	57,2
G. Wholesale and retail trade; motor vehicle and motorcycle repair	19.964	496.720	547.723	684.794	80,0
G45. Sale and repair of motor vehicles and motorcycles	2.133	76.504	70.328	85.865	81,9
G46. Wholesale trade and trade intermediaries, except of motor vehicles and motorcycles	15.140	212.881	341.390	405.586	84,2
G47. Retail trade, except of motor vehicles and motorcycles	2.691	207.335	136.005	193.343	70,3
G471. Retail trade in non-specialized stores	378	24.833	80.038	91.693	87,3
G473. Retail sale of automotive fuel in specialized stores	486	4.871	11.772	18.618	63,2
G472, G474 to G479. Rest of retail trade	1.827	177.631	44.195	83.032	53,2
H. Transportation and storage	2.975	181.033	51.556	80.024	64,4
I. Hostelry	1.456	269.340	18.344	65.287	28,1
I55. Accommodation Services	907	33.369	11.049	22.640	48,8
I56. Food and beverage services	549	235.971	7.295	42.647	17,1
J. Information and communications	1.642	69.112	57.953	70.739	81,9
M + N. Professional and administrative activities	5.600	647.369	112.607	175.550	64,1
Z. Rest of activities	7.000	1.517.812	82.308	173.877	47,3

Bibliography

CLEVELAND, R. B.; CLEVELAND, W. S.; MCRAE, J. E., and TERPENNING, I. (1990): "STL: A seasonal-trend decomposition procedure based on Loess". *Journal of Official Statistics*, 6(1): 3-73.

CUEVAS, A.; LEDO, R., and QUILIS, E.M. (2019): "Incorporando información fiscal de frecuencia diaria en la previsión macroeconómica a corto plazo", Instituto de Estudios Fiscales, *Papeles de Trabajo*, n. 11/2019.

DE LIVERA, A.; HYNDMAN, R. J., and SNYDER, R.D. (2011): "Forecasting time series with complex seasonal patterns using exponential smoothing", *Journal of the American Statistical Association*, 106(496): 1513-1527.

ESPASA, A.; REVUELTA, J. M., and CANCELO, J. R. (1996): "Automatic modelling of daily series of economic activity". In: Prat, A. (ed.), *Proceedings in Computational Statistics*. Physica-Verlag, Berlin, Germany: 51-63.

HARVEY, A. C. (1989): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge, UK.

HARVEY, A. C.; KOOPMAN, S., and RIANI, M. (1997): "The modelling and seasonal adjustment of weekly observations". *Journal of Business and Economic Statistics*, 15(3): 354-68.

HILLMER, S. C., and BELL, W. R. (1983): "Modeling time series with calendar variation". *Journal of the American Statistical Association*, 78(383): 526-534.

HILLMER, S. C.; BELL, W. R., and TIAO, G.C. (1983): "Modeling considerations in the seasonal adjustment of economic time series". In: Zellner, A. (Ed.) *Applied Time Series Analysis of Economic Data*, US Bureau of the Census, Washington, USA: 74-100.

HYNDMAN, R. J., and KHANDAKAR, Y. (2008): "Automatic time series forecasting: the forecast package for R", *Journal of Statistical Software*, 27(3): 1-22.

HYNDMAN, R. J., and ATHANASOPOULOS, G. (2013): *Forecasting: principles and practice*, OTexts: Melbourne, Australia.

LADIRAY, D., PALATE, J., MAZZI, G. L., and PROIETTI, T. (2018): "Seasonal adjustment of daily and weekly data". In: Ladiray, D., and Mazzi, G.L. (Eds.) *Handbook on Seasonal Adjustment*, Eurostat, Luxembourg.

LIU, L.M. (2005): *Time series analysis and forecasting*, Scientific Computing Associates, River Forest, Illinois, USA.

OLLECH, D. (2018): "Seasonal adjustment of daily time series", Deutsche Bundesbank, Discussion Paper n. 41/2018.

TAX AGENCY (2017): "Suministro Inmediato de Información (SII) del Impuesto sobre el Valor Añadido (IVA)", Agencia Estatal de Administración Tributaria (AEAT).