

Correction of calendar effects and signal extraction in short-term time series in Eustat

The methodology used in EUSTAT to correct calendar and seasonal effects and the extraction of the trend-cycle in short-term time series follows the recommendations by Eurostat detailed in [6] and is in agreement with the methodology used by the main statistical offices around the world.

The correction of calendar effects and signal extraction of different short-term statistics is carried out by applying the methodology based on ARIMA models implemented in the software TSW developed by Gianluca Caporello and Agustín Maravall [4]. This programme is the Windows version, with some modifications, of TRAMO and SEATS [9], developed by Víctor Gómez and Agustín Maravall. The TSW programme can be downloaded free of charge from the Banco de España web page ¹.

This document is mainly based on [20].

Linearization of raw series

TRAMO ('Time Series Regression with Arima Noise, Missing Observations and Outliers') is a programme for the estimation, prediction and interpolation of regression models with unobserved values and ARIMA errors that enables the modelling of different types of anomalous values or 'outliers'.

Given the raw series $z = (z_1, \dots, z_T)$, TRAMO adjusts the regression model

$$z_t = y_t' \beta + x_t \quad \forall t = 1, \dots, T, \quad (1)$$

where $\beta' = (\beta_1, \dots, \beta_K)$ is the vector of regression coefficients, $y_t' = (y_{1t}, \dots, y_{Kt})$ includes the K regression variables and x_t follows an ARIMA general process of the type,

$$\phi(B)\delta(B)x_t = \theta(B)\epsilon_t, \quad (2)$$

where B is the backshift operator, $\phi(B)$, $\delta(B)$ and $\theta(B)$ are polynomials in B and the innovations ϵ_t are supposed to be a white noise $N(0, \sigma_\epsilon^2)$ process for $\forall t$.

The polynomial $\delta(B)$ contains the unit roots derived from the differences both from the regular part and the seasonal part, $\phi(B)$ is the polynomial with the stationary

¹<http://www.bde.es/servicio/software/tsw.htm>.

autoregressive roots and $\theta(B)$ denotes the (invertible) moving average polynomial. It is supposed that polynomials are multiplicative of the form:

$$\begin{aligned}\delta(B) &= (1 - B)^d(1 - B^s)^D, \\ \phi(B) &= (1 + \phi_1 B + \dots + \phi_p B^p)(1 + \Phi_1 B^s + \dots + \Phi_P B^{sP}), \\ \theta(B) &= (1 + \theta_1 B + \dots + \theta_q B^q)(1 + \Theta_1 B^s + \dots + \Theta_P B^{sQ}),\end{aligned}\quad (3)$$

where s denotes the number of observations per year. The model may contain the constant μ which includes the mean of the differentiated series $\delta(B)z_t$ that, if significant, will be the constant regression term (i.e., $y_{1t} = 1$ for $\forall t = 1, \dots, T$).

The regressors $y'_t = (y_{1t}, \dots, y_{Kt})$ can be classified as:

1. Regressors specified by the user to include the effects of holidays, variables that are highly correlated with z_t that improve the adjustment, etc.
2. Regressors generated automatically by *TRAMO* to include the effects of working days, the effect of Easter and intervention variables of the type:
 - Dummy variables.
 - Any type of sequence of ones and zeros.
 - $1/(1 - \delta B)$ of any sequence of ones and zeros with $0 < \delta \leq 1$.
 - $1/(1 - \delta_s B^s)$ of any sequence of ones and zeros where $0 < \delta_s \leq 1$.
 - $1/((1 - B)(1 - B^s))$ of any sequence of ones and zeros.

Once the model is specified, *TRAMO*:

1. estimates by maximum likelihood (or conditional or unconditional least squares) all the parameters of the selected model (both from the regression and the general ARIMA model);
2. detects and corrects anomalous observations or 'outliers' (additive, transitory or level shift);
3. provides predictions of the series with their corresponding mean square error;
4. optimally interpolates the non observed values and calculates the mean square errors.

TRAMO has an option that carries out the treatment of 'outliers' and identification of the ARIMA model automatically. The methodology used is described at [13], [7], [8], [9], [10] y [11].

Signal extraction from linearized series

The software *SEATS* ('Signal Extraction in Arima Time Series') decomposes time series in unobservable components or signals, following a method based on ARIMA models developed for seasonal adjustment of economic time series (see, e.g. [5], [2], [3], [14], [1] y [15]).

TRAMO provides *SEATS* with the original series, the non stochastic effects it has estimated ('outliers', effects of working days and Easter, intervention variables, etc.), the linearized series (previously interpolated, if required), i.e., x_t in (1), and the ARIMA model estimated in (2).

SEATS additively breaks down ² the linearized series x_t , which follows the specified model in *TRAMO*, in the following components:

1. x_{pt} : trend-cycle component. The trend includes long term or low frequency movements, whose period is over 32 quarters (8 years). The cyclical component includes oscillations with a duration of between 2 and 8 years. Given the difficulty in discriminating between trend and cyclical oscillations, it is common to work with a mixed trend and cycle component.
2. x_{st} : seasonal component. Includes periodical or quasi-periodical movements with a duration equal to or less than one year.
3. x_{ct} : transitory component. This is a seasonal component with zero mean that captures transitory fluctuations that should not contaminate the trend-cycle and seasonality and do not follow a pattern of white noise.
4. x_{ut} : irregular component. It refers to the white noise movements that distort the linear relation between the observed series and their structural components of trend-cycle and seasonality.

The decomposition method used in *SEATS* breaks down the series x_t in orthogonal components that, except for the irregular component, satisfy the canonical property (impossibility of extracting a series of additive white noise from the components, therefore increasing their stability). Details of the methodology used may be consulted at [16], [17], [18], [19] and [12].

²The multiplicative break down is changed to additive by applying logarithms to the original series.

Referencias

- [1] Bell, W.R. y Hillmer, S.C. (1984). Issues Involved with the Seasonal Adjustment of Economic Time Series. *Journal of Business and Economic Statistics*, 2, 291-320.
- [2] Box, G.E.P. , Hillmer, S.C. y Tiao, G.C. (1978). *Analysis and Modeling of Seasonal Time Series*, en Zellner, A. (ed.), *Seasonal Analysis of Economic Time Series*, Washington, D.C.: US: Dept. of Commerce - Bureau of the Census, 309-334.
- [3] Burman, J.P. (1980). Seasonal Adjustment by Signal Extraction. *Journal of the Royal Statistical Society A*, 143, 321-337.
- [4] Caporello, G. y Maravall, A. (2004). *Program TSW. Revised Reference Manual*. Julio 2004. Banco de España.
- [5] Cleveland, W.P y Tiao, G.C. (1976). Decomposition of Seasonal Time Series: A Model for the X-11 Program. *Journal of the American Statistical Association*, 88, 284-297.
- [6] EUROSTAT (2002). *Methodology of short-term business statistics. Interpretation and guidelines. 2002 edition*. Eurostat, Luxembourg.
- [7] Gómez, V. y Maravall, A. (1992). *Time Series Regression with ARIMA Noise and Missing Observations - Program TRAM*. EUI Working Paper ECO No. 92/81, Department of Economics, European University Institute.
- [8] Gómez, V. y Maravall, A. (1994). Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter. *Journal of the American Statistical Association*, 89, 611-624.
- [9] Gómez, V. y Maravall, A. (1996). *Programs TRAMO (Time series Regression with Arima noise, Missing observations and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instructions for the User*. Documento de Trabajo 9628, Servicios de Estudios, Banco de España.
- [10] Gómez, V., Maravall, A. y Peña, D. (1999). Missing Observations in ARIMA Models: Skipping Approach Versus Additive Outlier Approach. *Journal of Econometrics*, 88, 341-364.
- [11] Gómez, V. y Maravall, A. (2001a), *Automatic Modelling Methods for Univariate Series*, Cap 7 en Peña D., Tiao G.C y Tsay, R.S. (eds.) *A Course in Time Series Analysis*, New York: J. Wiley and Sons.

- [12] Gómez, V. y Maravall, A. (2001b). *Seasonal Adjustment and Signal Extraction in Economic Time Series*, Cap 8 en Peña D., Tiao G.C. y Tsay, R.S. (eds). *A Course in Time Series Analysis*, New York: J. Wiley and Sons.
- [13] Harvey, A. (1990). *Forecasting, structural time series models, and the Kalman filter*. Cambridge University Press.
- [14] Hillmer, S.C. y Tiao, G.C. (1982). An ARIMA-Model Based Approach to Seasonal Adjustment. *Journal of the American Statistical Association*, **77**, 63-70.
- [15] Maravall, A. y Pierce, D.A. (1987). A Prototypical Seasonal Adjustment Model. *Journal of Time Series Analysis*, **8**, 177-193.
- [16] Maravall, A. (1988). *The Use of ARIMA Models in Unobserved Components Estimation*, en Barnett, W., Verndt, E. Y White, H. (eds). *Dynamic Econometric Modeling*, Cambridge: Cambridge University Press.
- [17] Maravall, A. (1993). Stochastic Linear Trends. *Journal of Econometrics*, **56**, 5-37.
- [18] Maravall, A. (1995). *Unobserved Components in Economic Time Series*, en Pesaran, H., Schmidt, P. Y Wickens, M. (eds). *The Handbook of Applied Econometrics*, vol.1, Oxford: Basil Blackwell.
- [19] Maravall, A. y Planas, C. (1999). Estimation Error and the Specification of Unobserved Component Models. *Journal of Econometrics*, **92**, 325-353.
- [20] Maravall, A. (2002). *Brief description of the programs*. Banco de España.