PROGNOSIS OF MONTHLY UNEMPLOYMENT RATE IN THE EUROPEAN UNION THROUGH METHODS BASED ON ECONOMETRIC MODELS

Gagea Mariana

Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration, Bd-ul Carol I, no. 22, Building B, B315, Iaşi, post code 700505, mariana.gagea@uaic.ro, Tel. +40 232 201 584

Balan Christiana Brigitte

Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration, Bd-ul Carol I, no. 22, Building B, B315, Iaşi, post code 700505, itte7@yahoo.com , Tel. +40 232 201 584

In this paper we propose the prognosis of the unemployment rate in the European Union through the Box-Jenkins method and the TRAMO/SEATS method as well as the detection of the method which proves to provide the best results. The monthly unemployment rate in the European Union is affected by seasonal variations of deterministic and stochastic nature. The prognosis through the Box-Jenkins nature supposes the separate consideration of seasonal variations, according to their specific nature. The stochastic seasonal variations are modelled and prognosticated simultaneously with the other components of the time series, based on the generating stochastic process. The prognosis of the monthly unemployment rate in the European Union through the TRAMO/SEATS methods is done by aggregating the individual prognoses of the components of the time series, obtained according to the stochastic processes models that generate them.

Key words: seasonal variations, stochastic process, moving average, prognosis, performance indicators of the prognosis

1. Introduction

There exists a great variety of analysis and prognosis methods as far as seasonal time series are concerned. These methods can be grouped according to various criteria, such as: their funding principle; the way seasonal variations are treated, either individually or globally, etc. The Box-Jenkins and TRAMO/SEATS methods are both based on ARIMA stochastic processes and models. The two methods are different as far as the approach of the seasonal component is concerned. The Box-Jenkins method consists in identifying, estimating and validating the stochastic process model that generates the time series as a whole, as well as modelling and the simultaneous prognosis of the components of the series, according to this model.

As opposed to the Box-Jenkins method, which it actually incorporates, TRAMO/SEATS is a method of time series decomposition, involving the individual treatment of components. The application of this method supposes finding the stochastic process that generates the analysed series and decomposing it in stochastic processes that generate the time series components. The prognosis of the values of this phenomenon is achieved by aggregating the individual prognoses of the components of the series.

The monthly unemployment rate in the European Union, considered as an average level of the phenomenon in the27 member states, is analysed for the period between January 2000 – November 2007. The data was extracted from the Eurostat portal. Data processing is done with SPSS, Demetra and JMuTi programmes.

The phenomenon under study has an accentuated seasonal character, manifested regularly in time. The presence of seasonal variations can be verified through different methods, such as: the linear chronograph, the Fisher test, the autocorrelation function.

We wish to make a comparison between the prognosis of the monthly unemployment rate in the European Union according to the Box-Jenkins and the TRAMO/SEATS methods, in order to facilitate the choice of the method providing the best results. The comparison will take into account the advantages, the limits and the results thus obtained.

2. Method description

a) The Box-Jenkins method

The Box-Jenkins method, proposed in 1970 by the authors who gave its name, is a short-term prognosis method. The Box-Jenkins method algorithm comprises the following steps: identification of the stochastic process model that generates the analysed time series, estimation of the model, model testing and prognosis. The first step of the identification stage of the process generating the time series is the stationarity analysis. A temporal series is considered stationary if the generating stochastic process is stationary. The identification of the stochastic process which generates the time series under analysis is based on the autocorrelation and partial autocorrelation functions (see 1, pages 47-82).

The basic stochastic models used in the econometric analysis of time series describe the series behaviour through linear combinations of white noises, linear combinations of past values or combinations of past values and white noises: moving average models (MA), autoregressive models (AR) and mixed autoregressive and moving average models (ARMA).

The estimation of the parameters of the stochastic process model is made using methods such as: the smallest least squares method, which can be easily applied in order to estimate autoregressive models; the method of maximum likelihood; the Yule-Walker algorithm. The validation of he estimated model requires the verification of the parameters meaning as well as error analysis (see 2, pp. 717-728; 3, pp. 225-247; 5, pp. 348-351).

b) The TRAMO/SEATS method

The TRAMO/SEATS method was developed by Agustin Maravall and Victor Gomez (1996), its starting point being the method proposed by Burman (1980). This method consists of estimating through OLS of the 'hidden' components in the time series ("signal extraction"), on the basis of ARIMA stochastic process which generates the series. Three main components of the time series are taken into account: the trend-cycle component, the seasonal component and the irregular component.

The **TRAMO** method (***T****ime series* ***R****egression with ARIMA noise,* ***M****issing values and* ***O****utliers*) consists of the estimation and prognosis of the regression model with ARIMA errors, called reg-ARIMA, in the possible presence of missing values, outliers of various types and other deterministic special effects. The application of the automatic procedure of the TRAMO method consists of the following steps [Demetra 2.0 User Manual, pp. 18-58]: identification of the aggregation scheme of the time series components; identification and estimation of the reg-ARIMA model on the grounds of the method of maximum likelihood; testing and correction (wherever necessary) of the calendar effects: the working days effect, the leap year effect, the Easter effect, the legal holidays effect; testing and correction of the outlier points of the series, that can be: extreme values, bearing the AO symbols, transitory shocks, bearing the TC symbols, or level changes, bearing the LS symbols; optimum interpolation of missing values of the time series; extension of the time series for the following two years through prognosis, on the basis of the estimated ARIMA model; determination of correction factors for the total, calendar and outlier effects and calculation of the pre-adjusted or 'linearized' temporal series values. (see 4, pp. 18-58).

The **SEATS method** (***S****ignal* ***E****xtraction in ARIMA* ***T****ime* ***S****eries*) is a method of decomposition of temporal series in components, according to the ARIMA modes. The stochastic components estimated through this method are added deterministic effects (the calendar and outlier effects) eliminated through the TRAMO method, the final components being thus obtained: the corrected series of seasonal variations, the tendency, the seasonal component, the irregular component. The SEATS method offers the prognosis of these components and of the uncorrected values of the seasonal variations for the following two years.

3. Presentation of results

a) The prognosis of the monthly unemployment rate in the European Union through the Box-Jenkins method

The components of the unemployment rate in the European Union as well as their aggregation scheme are identified using the Fisher test and the band test (see 3, pp. 11-14; 21-22). The analysed time series presents significant trend and seasonal variations, consequently the series is non-stationary. The adequate aggregation scheme of the components is thus the additive one.

In order to identify the adequate stationarity techniques the Franses test is applied for testing the seasonal unit root, with the JMulTi programme (see 3, pp. 172-175). The model that includes a constant term, seasonal binary variables and trend is chosen. The estimated and tested regression model presents statistically significant coefficients of the seasonal binary variables and has no seasonal roots of unit module. Consequently, the unemployment rate in the European Union presents deterministic seasonal variations which are eliminated through the moving averages method (see 5, pp. 444-447).

The simple autocorrelation function (figure 1) and the partial autocorrelation function (figure 2) for the unemployment rate in the European Union are then analysed. Their deterministic seasonal variations have been extracted.

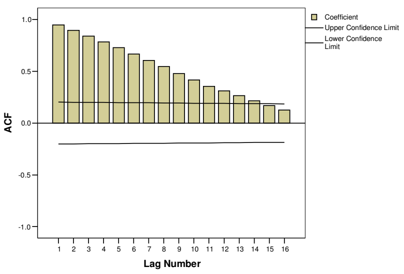

Fig. 1. The autocorrelation function for the unemployment rate in the European Union out of which the deterministic seasonal variations have been extracted

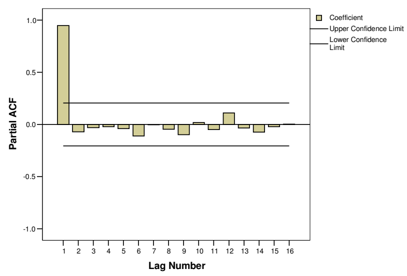

Fig. 2. The partial autocorrelation function for the unemployment rate in the European Union out of which the deterministic seasonal variations have been extracted

The autocorrelation functions of the analysed series indicate an autoregressive stochastic process, bearing the symbol AR(1), defined by the relation:

$y_t = \phi_1 y_{t-1} + \varepsilon_t$.

The order 1 autoregressive process equation with constant terms estimated through the method of maximum likelihood, with the SPSS programme. The estimated equation is: $y_t^* = \phi_1(\tilde{y}_{t-1} - \mu) + \mu = 0{,}996(\tilde{y}_{t-1} - 8{,}028) + 8{,}028$, with $\tilde{y}_t$ the value of the unemployment rate in the European Union in moment *t*, out of which the deterministic seasonal variations have been extracted.

The parameters ϕ_1 and μ of the model are statistically significant, the condition of rejection of the nil hypothesis being fulfilled: $(Sig = 0{,}00) < (\alpha = 0{,}05)$.

According to the autocorrelation function and the Jarque-Bera test, the modelling errors define a stochastic process of the 'white noise' type. The autoregressive model AR(1) is validated. Based on this model the unemployment rate in the European Union is estimated for the period December 2007 – November 2008, according to the relation: $y_{T+1}^* = 0{,}996(y_T - 8{,}028) + 8{,}028$. Finally, to the values obtained one adds the deterministic seasonal variations, presented as seasonal coefficients: $y_{T+1}^{(P)} = y_{T+1}^* + S_j$.

The measurement of the performance of the unemployment rate in the European Union through the Box-Jenkins method is made by means of the absolute prognosis errors average. Prognosis errors are calculated as the difference between the observed value and the estimated value, according to the relation: $e_t = y_t - y_t^{(P)}$. The average of the absolute prognosis errors is: $M(|e|) = \frac{\sum_{t=1}^{T} |e_t|}{T} = 0.079$.

b) The prognosis of the unemployment rate in the European Union through the TRAMO/SEATS method.

The TRAMO/SEATS method is applied with the Demetra programme, automatic procedure. The output of the decomposition of the unemployment rate in the European Union through the T/S method is structured on three points and presents the results below.

Information regarding the model used: 1) The terms of the series have been logarithmically transformed and the log-additive aggregation model was used; 2) Three outliers with significant statistic effect were identified, being presented in the inversed order of their importance : May 2006 (AO), January 2007 (LS), November 2005 (LS); 3) The model of the stochastic process generating the time series is ARIMA$(1,1,0)(0,1,0)_{12}$, with all the ϕ_1 parameters corresponding to the statistically significant non-seasonal autoregressive part, according to the t-Student test. The estimation of the model parameters was made by means of the maximum likelihood method; 4) The decomposition in components was made based on the exact ARIMA model, estimated in the pre-adjustment stage with the TRAMO method.

The equation of the reg-ARIMA model can be symbolically written in the following manner: $y_t = 3$ outliers + x_t. The equation of the ARIMA$(1,1,0)(0,1,0)_{12}$ mode of the regression errors x_t is:

$$(1-\phi_1 L)\Delta\Delta_{12} x_t = \varepsilon_t \Leftrightarrow (1+0{,}127L)\Delta\Delta_{12} x_t = \varepsilon_t,$$

where: - Δ represents the difference operation, resulted from the relation $\Delta y_t = y_t - y_{t-1}$; $\Delta_{12} y_t = y_t - y_{t-12}$

- L - the delay operator, resulted from the relation $Ly_t = y_{t-1}$.

Information regarding the validation criteria of the modelling (table 1): 1) The estimated value of the Ljung-Box statistic of order k=24 is Q_{24}=15.55. The trust interval for Q_{24}, guaranteed with a probability of 95%, is [0; 35.2]. The empiric value is covered by the trust interval, so the estimated value of the Q_{24} statistics is not statistically significant, the errors are not auto-correlated to the k=24 disparity; 2) The estimated value of the Jarque-Berra statistics is $JB = 0.65$. The trust interval for *JB*, guaranteed with a probability of 95%, is [0; 5.99]. The calculated value of the JB statistics is included in the accepted interval of the nil hypothesis, the errors are normally distributed; 3) The identified outliers represent 3.61% of the total number of observations of the monthly unemployment rate in the European Union. The percentage is smaller than the maximum allowed limit of 5% (see 4, pp. 27-28).

Information on Diagnostics	Model 1 (Tramo-Seats)
SA quality index (stand. to 10)	1.555 [0, 10] ad-hoc
STATISTICS ON RESIDUALS	
Ljung-Box on residuals	15.55 [0, 35.20] 5%
Box-Pierce on residuals	0.63 [0, 5.99] 5%
Ljung-Box on squared residuals	25.22 [0, 35.20] 5%
Box-Pierce on squared residuals	0.36 [0, 5.99] 5%
DESCRIPTION OF RESIDUALS	
Normality	0.65 [0, 5.99] 5%
Skewness	0.22 [-0.55, 0.55] 5%
Kurtosis	2.87 [1.91, 4.09] 5%
OUTLIERS	
Percentage of outliers	3.16% [0%, 5.0%] ad-hoc

Table 1. Synthesis of statistics used for the validation of decomposition by means of the TRAMO method

Programme diagnosis. All tests being passed, the modelling of the analysed series through this variant of the TRAMO/SEATS method is accepted and there follows the next step: prognosis for the components of the series for the following two years.

We wish to make a comparison between the prognosis of the monthly unemployment rate in the European Union according to the Box-Jenkins and the TRAMO/SEATS methods. Due to the fact that the two methods are based on different principles and ways of treating the seasonal variations, a comparison between the results obtained with these two methods is quite difficult to achieve. The simplest solution seems to be a graphic comparison (figure 3), yet this method only provides intuitive information. The simple graphical comparison between the two prognoses does not say exactly which of the two methods has led to better results.

Prognosis methods based on re-seasonalization do not allow the determination of the theoretical values of the series, necessary for the comparison with the empirical values. These situations prevent the determination of prognosis errors, on which the performance indicators of the prognosis are actually built. The prognosis absolute error average, which equals 0.079, was calculated for the prognosis of the unemployment rate in the European Union through the Box-Jenkins method. The performance measurement of the prognosis obtained through the TRAMO/SEATS method, by means of this specific indicator, is not allowed.

One possible solution in order to compare the results of the two prognosis methods is the following: the analysed time series is divided into two sub-series, the work series and the control series. The work series can be considered the unemployment rate in the European Union in the period January 2000 – November 2006, while the control series is represented by the unemployment rate in the period December 2006 – November 2007. Modelling and prognosis are based on the data comprised in the work series, while the control series is used for the calculation of prognosis errors and the performance indicators, according to which the methods are compared and the best method is chosen. Following this procedure, the following results have been obtained: the average of the absolute prognosis errors for the Box-Jenkins method is 0.669, while for the TRAMO/SEATS method the average is 0.295. The method that provided the best results in this particular situation is thus the TRAMO/SEATS method.

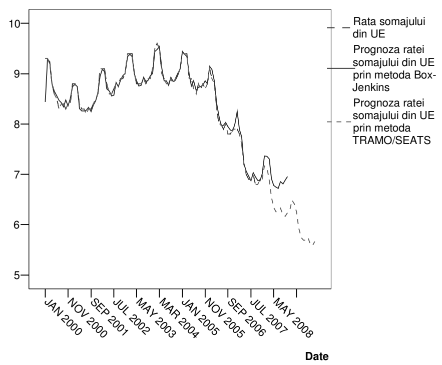

Fig. 3. Unemployment rate in the European Union initiated and prognosticated through the Box-Jenkins and TRAMO/SEATS methods

4. Conclusions

The Box-Jenkins and TRAMO/SEATS methods have the advantage of using analysis instruments that ensure the basis for the statistical interface. The Box-Jenkins method is limited by the experience and subjectivity of the user in the identification stage of the stochastic process model. Such an example could be provided the modelling of the unemployment rate in the European Union based on the initial series in comparison with the work series. In the latter case the prognosis errors that were obtained were substantially bigger then in the case of the initial series. As far as limiting user subjectivity is concerned, the TRAMO/SEATS method provides the advantage of the automatic procedure of the Demetra programme and the automatic guidance of the user when it comes to finding a solution to an identified problem.

The direct numerical comparison of the results of prognosis methods based on different approaches of seasonal variations is possible neither individually nor globally. We are reserved in accepting the solution of prognosis methods' comparison by splitting the initial series into working series and control series, even if this is one of the few solutions that exist. Therefore, the application of this method raises two important issues: data representativity for the two sub-series, and subjectivity in applying the methods on different time series: the initial and the working series.

Bibliography

1. Alonso, J. H., *Econometria de series temporales*, Editorial Universitas, S.A., Madrid, 2000
2. Andrei, T., *Statistică şi Econometrie,* Ed. Economică, Bucureşti, 2003
3. Bourbonnais, R., Terraza, M., *Analyse des series temporelles. Aplication à l'economie et à la gestion*, DUNOD, Paris, 2004
4. Demetra 2.0x User Manual, http://circa.europa.eu
5. Jaba, E., *Statistica,* 3rd edition, Ed. Economică, Bucureşti, 2002