ANALELE ŞTIINŢIFICE ALE UNIVERSITĂŢII „ALEXANDRU IOAN CUZA" DIN IAŞI
Tomul LV Ştiinţe Economice 2008

EXTRACTION OF SEASONAL VARIATIONS OF UNEMPLOYMENT RATE IN ROMANIA USING SEVERAL METHODS BASED ON MOVING AVERAGE FILTER

Mariana GAGEA*, Alina Măriuca IONESCU**

Abstract

At present, both at European Union and world level, experts are preoccupied to find the best method for the deseasonalisation of a time series that should assure the comparability of statistical data. The present paper follows the line of these researches.

In the study, we undertake a comparison of the most representative methods based on moving average filter: moving average method, Census X-11 method and X-12 ARIMA method. Theoretical research shows the superiority of X-12 ARIMA method, which has incorporated the previous methods as regards the algorithm and the advantages, contributing to the improvement of the weaknesses of the former methods. The criteria for the comparison of the results obtained through the three methods applied to the time series of unemployment rate in Romania during the period 2000 - 2007 didn't indicate a unique method, as being the most adequate for deseasonalisation.

Key words: *seasonal variations, moving average filter, ARIMA models for stochastic processes*
JEL classification: C22

1. Introduction

Lately, Romania has recorded an economic growth, reflected in positive trends of the evolution of GDP, growth rhythm of industrial production, final consumption and gross fixed capital formation. As regards the economic growth, economic analyses of international organisms have placed Romania on one of the first positions among the European Union member states.

Positive aspects of the economical situation in Romania are reflected on labor market too, in the increasing number of occupied population and in the reduction of unemployment. At present, unemployment rate in Romania presents a tendency of decrease comparing with anterior period, having a level inferior to average level of European Union.

---

* Mariana GAGEA (marygag2002@yahoo.com) is assistant professor of Economy, Quantitative Analysis and Informational Systems Department at "Alexandru Ioan Cuza" University of Iasi, Faculty of Economics and Business Administration. He received his PhD in: Cybernetics and Economic Statistics. Her research interests include: Time Series Analysis with SPSS and EViews. Her teaching interests include: Econometrics, Basic Statistics.

** Alina Măriuca IONESCU (alina.ionescu@yahoo.com) is PhD of Economy, Quantitative Analysis and Informational Systems Department at "Alexandru Ioan Cuza" University of Iasi, Faculty of Economics and Business Administration. He received his PhD in: Cybernetics and Economic Statistics. Her research interests include: Multivariate Statistical Analysis. Her teaching interests include: Basic statistics, Econometrics.

Analysis of unemployment rate is of great importance since this indicator, together with other macro-economical indicators, reveals general and structural changes that occur in transition period towards an economy with a competitive market.

In the paper we aim to find the best method to analyze the seasonal variations of the time series of monthly unemployment rate in Romania, during the period January 2000 – November 2007. Data are selected from the database available on Eurostat site [Eurostat, 2008].

We use the following methods for the extraction of seasonal variations in a time series that are founded on a moving average filter: moving average method, Census X-11 method and X-12 ARIMA method. Comparison is done with respect to algorithm, advantages, limits, and also according to the obtained results.

Unlike the case of methods based on econometric models, comparison of the results provided by certain methods based on a moving average filter is very difficult. They are ad-hoc methods and they don't include instruments that may assure statistical inference basis. At the present time, experts in this field are preoccupied to find the best method for seasonal adjustment of time series and, in this context, to find unique criteria for comparison of methods. The present paper follows the line of these researches aiming to identify the "best" method for extracting seasonal variations from the time series of unemployment rate in Romania.

2. Presentation of the methods

The extraction methods of seasonal component constructed on the moving average filter have the simplest method, called *moving average method*, as their „core" (figure 1). The more complex methods, Census X-11 and X-12 ARIMA, have refined moving average method, have improved its limits and have developed the algorithm, so that moving average method may be considered as being integrated in subsequently developed methods.

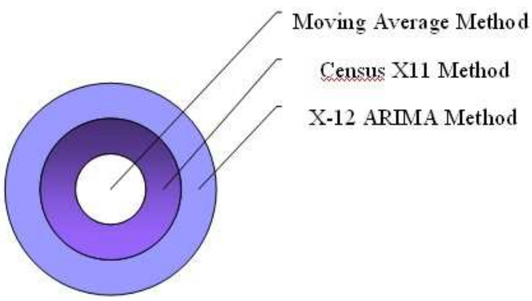

Fig. 1 Methods for the extraction of seasonal component based on moving average filter

2.1. Moving average method

Moving average method, also called the classical method, represents the most simple method for the extraction of seasonal variations [Bourbonnais, R., 2004, 23-24, 31-33; Jaba, E., 2002, 444-447; Vaté, M., 1993, 181-185]. *The algorithm* of the method contains a sequence of operations, with particularities for the model of aggregation of the temporal series components, additive or multiplicative. Application of moving average method to a time series obtained through multiplicative aggregation of its components, such is the series of unemployment rate in Romania, implies the following of the next steps: a) adjusting the time series using moving averages *(* $\bar{y}_t$ *)* – consists in estimating tendency through a moving average of order equal to the period of seasonal variations; b) computing of seasonal indices *(* i_t *)* – as ratios between observed terms and those adjusted through moving average; c) computing seasonal coefficients (S_j , $j = 1,2,\ldots,P$, with P being the period of seasonal variations) – as the average of seasonal indices corresponding to seasonal period *j* (month, trimester etc.) during the *n-1* observed years of the adjusted series; d) correction of the seasonal variations of the time series *(* y_t^* *)* – as ratios between empirical values (y_t) and corresponding seasonal coefficients ($S_{j,t}$).

The advantage of the elementary method of moving averages consists in the simple algorithm, with quick calculation, easy to effectuate with and without the help of a computer. Application of this method is *limited* by the using of moving averages in the first step of the algorithm for the estimation of the tendency, which shortens the temporal series with *P-1* terms, where *P* is the order of moving average, which is the period of seasonal variations. In addition, the determinist assumption of seasonal variations, which are repeating rigorously identical from a period *P* to another period, is very restrictive and only partly respected in real economy.

2.2. Census X-11 method

Census X-11 method was proposed by Shiskin, Youle and Musgrave, in the year 1967. The method follows principles that underlie moving average method for the extraction of seasonal variations of a time series and, in addition, it uses an *iterative principle* in estimation of the components of the series (Bourbonnais, R., 2004, 34-35; Lidary, D., 2001, 15-31; Vaté, M., 1993, 186-187).

Unlike the moving average method, Census X-11 method has instruments implemented for the identification and correction of atypical values and of calendar effects (trading day effect, Easter effect etc.). These effects are estimated using linear regression models, on the basis of residual component which incorporates them.

The algorithm of Census X-11 method incorporates the simple algorithm of the moving average method. The main stages are denoted with capital letters: A, B, C and D. Each stage is composed of two sub-stages that respect the steps of the simple algorithm of decomposition of time series.

Census X-11 method presents the *advantage* of correcting atypical values and calendar effects. The method overtakes the limits of the rigid treatment of seasonal variations, by calculating seasonal coefficients with moving averages, for each seasonal sub-stage from each year. Census X-11 method treats random seasonal variations and it is both a method of decomposition of time series and a method of forecasting.

The main *disadvantage* of Census X-11 method consists in using moving averages that are asymmetric at the ends of the temporal series. Another weak point of X-11 method is represented by the lack of explicit models. More than that, users might have different options in application of the method, so that the obtained results for the same time series may not be compared.

2.3. X-12 ARIMA method

X-12 ARIMA method was proposed in the second part of the '90s by Census Office of United States, being the most recent version of the methods from Census X11 family (Demetra 2.0X User Manual, 2002, 48-66; Fischer, R., 1995, 13-16). X-12 ARIMA method and TRAMO/SEATS represent methods of seasonal adjustment accepted and recommended, at present, by Statistical Office of the European Union (Eurostat).

X-12 ARIMA method solve the problem of the „end points" of Census X11 method by extension or prolongation of the original series at the two ends, with values estimated through ARIMA model. Thus, the estimators of the terms from the ends of the series are improved. Prolongation of the original series, adjustment of atypical values and of calendar effects are done with regARIMA models (regression models with ARIMA errors).

Extraction of seasonal variations of a time series using X-12 ARIMA method implies the following synthetic algorithm: a) identifying the aggregation scheme of the components of the series and transformation of the series through logarithm, in the case of multiplicative scheme; b) estimating and testing reg-ARIMA model, using maximum likelihood method, including trading days effect, Easter effect, outlier effect and other special effects, and ARIMA model of regression errors, with the verification for the significance of its parameters; c) prolongation of the series using prognosis with ARIMA model, for 12 months, in the case of monthly temporal series, and for 4 trimesters, in the case of quarterly time series; d) decomposition of the „linearized" series through Census X-11 method.

Comparing with X-11 method, incorporated in it, X-12 ARIMA method presents the *advantage* of correcting calendar effects (trading day effect, Easter effect, holidays effects) and extreme values effect using regression method, in the pre-adjustment stage. This method provides a good solution for the problem of „end points" by extension or prolongation of the original series at its two ends, with values estimated using ARIMA models.

We consider that a *weakness* of the method is the possibility to apply different criteria for the approval of the time series treatment, which raises the problem of the comparability of the results obtained using different versions of the method or different time series.

2.4 Criteria for comparison of the results provided by different methods for the extraction of seasonal variations

Comparison of the methods for the extraction of seasonal variations as regards the obtained results is very difficult. In speciality literature, the american Börn Fisher suggests several criteria for the comparison of the results provided by several methods for analysis of seasonal component of a time series, such as (Fischer, R., 1995, 34-38): a) the concordance between the annual sum of the original series' terms and that of the seasonally adjusted series, measured through absolute mean difference between the annual sum of the original series' terms and that of the seasonally adjusted series; b) the orthogonality of the seasonally adjusted series and seasonal variations, using absolute mean difference between the values corrected of seasonal variations obtained from the first and from the second seasonal adjustment; c) the variability of the seasonal component, measured with relative mean square difference between the coefficients corresponding to month *j* from successive years; d) the correlation between variation of seasonal component estimated with different methods, using Spearman coefficient etc.

3. Results

The first step when analyzing a time series is the identification of its components using methods as: linear chronogram, autocorrelation function, Fisher test. The stage of the identification of the components of the series comes with the identification of their aggregation scheme too. To achieve this we could use the band test and Buys-Ballot test (Bourbonnais, R., 2004, 9-22; Gagea, M., 2004, 223-232).

The methods for the identification of the components of the series and of their aggregation scheme have unitarily indicated the presence of the following components in the time series of unemployment rate in Romania: trend component, seasonal component and random component. The adequate aggregation scheme of the components is the multiplicative one.

3.1. Extraction of seasonal variations out of the time series of unemployment rate in Romania using moving average method

Extraction of the seasonal variations out of the time series of unemployment rate in Romania using average moving method, the multiplicative version, is done with SPSS software. Seasonal component expressed with seasonal coefficients is plotted in figure 2. Graphical comparison of the original time series and that of the series corrected of seasonal variations (figure 3) indicates a significant increase of seasonal variations and of random oscillations around the trend line, so the chosen deseasonalisation model is accepted.

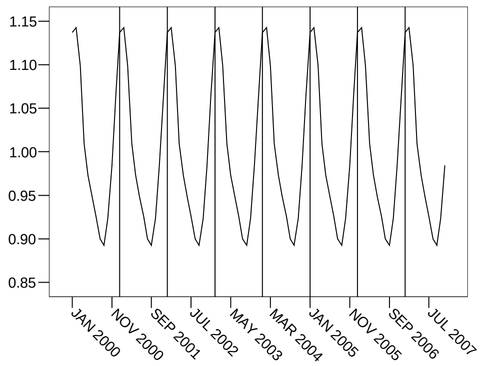

Fig. 2 Seasonal component of the series of unemployment rate in Romania, during the period 2000 –2007, extracted with moving average method

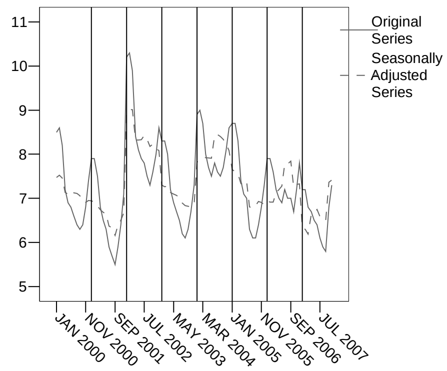

Fig. 3 Original series and deseasonalised series with moving average method, for the unemployment rate in Romania, during the period 2000 – 2007

3.2. Extraction of seasonal variations of the unemployment rate in Romania with Census X-11 method

Extraction of seasonal component of the unemployment rate in Romania with Census X-11 method, the multiplicative type, is done with EViews 5 programme.

During the stages B and C, preceding the final extraction of the seasonal component, Census X-11 method has identified 15 outliers, for which we compute the factors of correction of the seasonal indices and the values of replacement. The most outliers are recorded during December (four values) and January (three values).

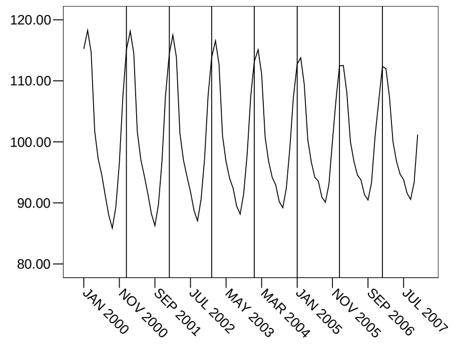

Fig. 4 Seasonal component of the series of unemployment rate in Romania, during the period 2000 –2007, extracted with Census X-11 method

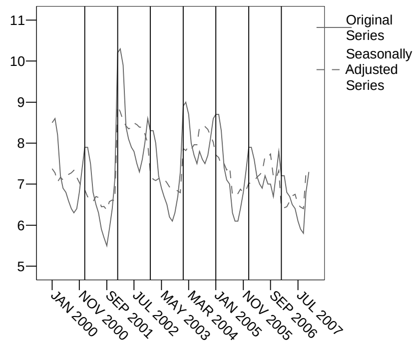

Fig. 5 Original series and deseasonalised series with Census X-11 method, for the unemployment rate in Romania, during the period 2000 – 2007

It can be noticed that the amplitude of seasonal variations extracted with Census X-11 method (figure 4) decreases in time, following the descending line of the phenomenon's tendency of evolution. The unemployment rate corrected of seasonal variations presents a significant decrease of the variation amplitude comparing with the original series (figure 5).

3.3. Extraction of seasonal variations of the unemployment rate in Romania with X-12 ARIMA method

The seasonal variations of unemployment rate in Romania are extracted using X-12 ARIMA method, the multiplicative version, with Demetra 2.1 programme, the automatic procedure.

The output of X-12 ARIMA method for the extraction of seasonal component of unemployment rate in Romania is structured on three issues: information about the used model, information about criteria for the validation of modeling and diagnosis of Demetra programme. We obtained the following information about the used model: 1) series' terms were logarithmic transformed and the log-additive model of aggregation was used; 2) trading day effect isn't statistically significant and its correction was not applied; 3) Easter effect isn't statistically significant and its correction was not applied; 4) there were identified 2 outliers: January 2002 (LS), January 2004 (LS); 5) the model of stochastic process generating regression errors is ARIMA$(2,1,2)(0,1,1)_{12}$, with all its parameters being statistically significant, according to t-Student test; 6) the estimation of the seasonal variations was done with the moving average filter MM3x3, while the estimation of the trend was accomplished with Henderson moving average filter of 9 terms.

Equation of reg-ARIMA model may be symbolically written as it follows: $y_t = 2$ outliers + x_t, where x_t represents the regression errors generated by the stochastic process ARIMA$(2,1,2)(0,1,1)_{12}$.

All the criteria for the validation of the modeling of unemployment rate in Romania through the estimated model are satisfied and Demetra programme accepts the treatment of the time series with this version of X-12 ARIMA method.

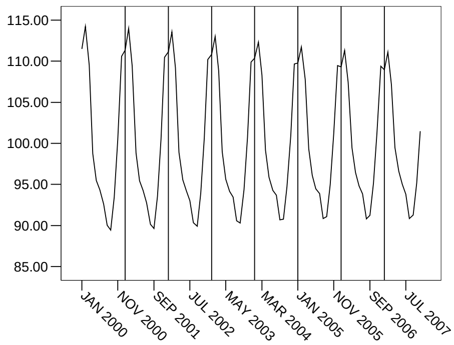

Fig. 6 Seasonal component of the series of unemployment rate in Romania, during the period 2000 –2007, extracted with X-12 ARIMA method

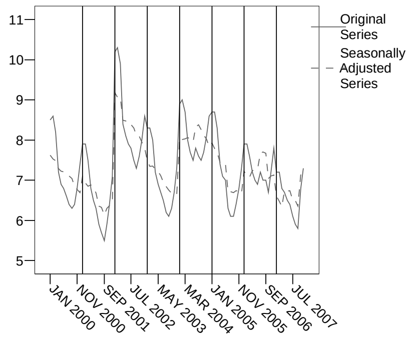

Fig. 7 Original series and deseasonalised series with X-12 ARIMA method, for the unemployment rate in Romania, during the period 2000 – 2007

As in the case of the seasonal variations extracted with Census X-11 method, the amplitude of seasonal variations extracted with X-12 ARIMA method (figure 6) decreases in time, following the descending line of the phenomenon's tendency of evolution. The time series corrected of seasonal variations presents a significant decrease of the variation amplitude comparing with the original series (figure 7).

3.4. Comparison of the results obtained with different methods for the extraction of seasonal variations, for unemployment rate in Romania

We apply the following criteria for the comparison of methods for the analysis of seasonal component of temporal series: the concordance between the annual total of the original series and that of the seasonally adjusted series, the orthogonality of seasonally adjusted series and seasonal variations, the variability of seasonal component, the correlation between components extracted with different methods.

a) *Concordance between the annual total of the original series and that of the seasonally adjusted series*

The absolute mean difference between the annual sum of original series' terms and the annual sum of the terms of each of the four seasonally adjusted series (table 1) is calculated with the relation:

$$Ma = \frac{\sum_{i=1}^{n} \left| S_i - S_i^* \right|}{n},$$

where:

- Ma represents absolute mean difference between the annual sum of the original series' terms and the annual sum of the adjusted series' terms;
- S_i - sum of the original series' terms, corresponding to year *i*, $i = \overline{1,n}$, with $S_i = \sum_{j=1}^{P} y_{ij}$, *P* being the period of seasonal variations;

- S_i^* - sum of the seasonally adjusted series' terms, corresponding to year *i*, $i=\overline{1,n}$, with $S_i = \sum_{j=1}^{P} y_{ij}^*$, y_{ij}^* being the adjusted value corresponding to year *i* and seasonal period *j*.

We consider to be a better method the one for which absolute mean difference is minimal. In the considered example, the criterion indicates Census X-11 method as being the most efficient one (table 1, column 2).

Table no. 1 – Elements of calculation

Method	***Absolute mean difference between the annual sum of the terms***	***Absolute mean difference between the terms adjusted once and twice***	***Relative mean square difference between successive coefficients***
1	*2*	*3*	*4*
Moving average method	*0.263*	***0.0038***	*-*
Census X-11 method	***0.216***	*0.0419*	*0.0046*
X-12 ARIMA method	*0.226*	*0.1625*	***0.0022***

b) Orthogonality of seasonally adjusted series and seasonal variations

Consider that the series corrected of seasonal variations shouldn't contain traces of seasonal variations. Thus, if we apply a new seasonal filtering to the series already corrected of seasonal variations, the resulting series should coincide with the first one.

In order to verify this assumption we do as it follows: we calculate absolute differences between the values corrected of seasonal variations obtained through the first and the second seasonal adjustment; then, we determine absolute mean difference with the relation:

$$M = \frac{\sum_{t=1}^{T} \left| y_t^* - y_t^{**} \right|}{T},$$

where: y_t^* represents the adjusted value through the first filtering, at moment *t*, and y_t^{**} is the adjusted value through the second filtering, at moment *t.*

It is considered to be a better method the one for which we get the smallest value of the absolute mean difference between the values obtained through the first and the second seasonal adjustment. In the considered example, the method with the best results is the moving average method (table 1, column 3).

c) Variability of the seasonal component

The variability of the seasonal component is measured using the relative mean square difference between the seasonal coefficients corresponding to month *j*, from successive years:

$$M_{Sj} = \frac{\sum_{k=2}^{n} \left(\frac{S_{kj} - S_{(k-1)j}}{S_{(k-1)j}} \right)^2}{n-1}.$$

It is desirable that the variability of the coefficients corresponding to the same period to be as small as possible.

The variability of the seasonal component extracted using moving average method is always zero, since the method is based on the determinist principle, so that seasonal variations are identical from a seasonal period *j* to another, $j = \overline{1,P}$. We study this criterion only for Census X-11 and X-12 ARIMA methods. Monthly relative mean square difference between successive seasonal coefficients is minimal for X-12 ARIMA method (table 1, column 4).

d) Correlation between components obtained with different methods

The correlations between seasonal components extracted with different methods are calculated using *Spearman's rank correlation coefficient*. All the estimations of Spearman's coefficient are statistically significant and greater than 0.95, meaning that seasonal components estimated using different methods are very closely correlated (table 2).

Table no. 2- Correlation between seasonal components extracted using different methods

			Seasonal Coefficients MA	Seasonal Coefficients X11	Seasonal Coefficients X-12 ARIMA
Spearman's rho	Seasonal Coefficients MA	Correlation Coefficient	1.000	.979**	.961**
		Sig. (2-tailed)	.	.000	.000
		N	95	95	95
	Seasonal Coefficients X11	Correlation Coefficient	.979**	1.000	.967**
		Sig. (2-tailed)	.000	.	.000
		N	95	95	95
	Seasonal Coefficients X-12 ARIMA	Correlation Coefficient	.961**	.967**	1.000
		Sig. (2-tailed)	.000	.000	.
		N	95	95	95

**. Correlation is significant at the 0.01 level (2-tailed).

4. Conclusions

The methods for the extraction of seasonal variations based on moving average filter are ad-hoc methods, making difficult the comparison of the results obtained with different methods. *Criteria for the comparison*, suggested by Börn Fisher in a study done for Eurostat, *didn't indicate a unique method as being the most adequate for the extraction of seasonal variations from the time series of unemployment rate in Romania*.

As regards the criterion of the concordance between the annual total of the original series and that of the seasonally adjusted series, the method with the best results is Census X-11. The criterion of the orthogonality between the seasonally adjusted series and the seasonal variations has indicated the moving average method to be the best, closely followed by Census X-11 method. The third criterion, the variability of the seasonal component, has placed X-12 ARIMA method on the first place, having a small advantage comparing with Census X-11 method.

The three compared methods led to much closed results. This aspect is emphasized both with the comparative graphical analysis of seasonal variations extracted using different methods and with the measurement of their correlation, with Spearman's coefficient.

We consider that the best method for the extraction of seasonal variations should lead to the best prognosis, too. In this context, the evaluation of the results of the different methods for the extraction of seasonal variations may be done, indirectly, using prognosis performance indicators. They are measured during the last stage of the analysis of a time series, which is the prognosis stage.

References

Bourbonnais, R., Terraza, M., Analyse des series temporelles. Aplication à l'economie et à la gestion, DUNOD, Paris, 2004

Demetra 2.0x User Manual, 2002, http://circa.europa.eu, accesta la 10 decembrie 2007

Fischer, B., *Descomposition of Time Series. Comparing Different Methods in Theory and Practice*, martie/aprilie 1995, http://ec.europa.eu/comm/eurostat/, accesat la 20 februarie 2008

Gagea, M., *Analiza statistică a variațiilor sezoniere ale ratei şomajului din Romania* în „Probleme Actuale de Statistică. Dinamica fenomenelor şi proceselor economico-sociale", coordonatori Elisabeta Jaba, Alecsandru Puiu Tacu, Ed. Junimea, Iaşi, 2004

Jaba, E. Statistica, Ediția a treia, Ed. Economică, Bucureşti, 2002

Lidary, D., Quenneville, B., *D*ésaisonnaliser avec la méthode X-11, 2001, www.census.gov/ts/paper/X11_french.pdf, accesat la 15 iunie 2003

Vaté, M., *Statistique chronologique et prevision*, Ed. Economica, Paris, 1993

*** *Eurostat*, http://epp.eurostat.ec.europa.eu/, accesat la 15 ianuarie 2008