

# A Forecast Model of the Employment Rate in Romania

Elisabeta JABA<sup>1</sup>, Professor PhD  
Carmen PINTILESCU<sup>2</sup>, Assoc. Prof. PhD  
Elena Daniela VIORICA<sup>3</sup>,  
Christiana Brigitte BALAN<sup>4</sup>  
“Alexandru Ioan Cuza” University of Iasi, Romania

## Abstract

The paper analyzes the dynamics of the population employment rate in Romania and identifies a model of its forecast.

The evaluation of the dynamics of the employment rate and of the influence factors – demographic, economic, demo-economic – is made analyzing the indexes of the dynamics. The analysis of the dynamics employment rate correlated with the dynamics of the influence factors, using the principal component analysis, allows us to identify years clusters. Each cluster is characterized by several features expressed by the classification function. The classification function can be used for the forecast of the cluster for a given time horizon. The employment rate for a given horizon can be estimated knowing the years cluster’s features to which the time horizon belongs. The years cluster membership of time horizon is estimated using the discriminant function.

**Key words:** employment rate, forecast model, dynamics indexes, principal component analysis, discriminant analysis

## Introduction

The attention paid to the analysis of the dynamics of the population employment rate is due to the place held by this factor rate dynamics in the development strategy [Blanchard, 1997; Stiglitz, Walsh, 2004].

During the transition period, Romania followed the line of openness and closeness to the developed countries economies by restructuring the economy and by the privatization process that generated economic downfalls, unemployment and inflation. After 2003, the macroeconomic outputs of Romania improved and are reflected in the GDP growth rate generated by the high volume of investments and of private consumption, together with the decrease in inflation, unemployment and budgetary deficit.

## 1. Trends in the dynamics of the employment rate and of the influence factors

---

<sup>1</sup> Professor PhD, Head the Department of Statistics, Faculty of Economics and Business Administration, “Alexandru Ioan Cuza” University of Iasi, Romania, ejaba@uaic.ro

<sup>2</sup> Associate Professor, Department of Statistics, Faculty of Economics and Business Administration, “Alexandru Ioan Cuza” University of Iasi, Romania, carmen.pintilescu@uaic.ro

<sup>3</sup> PhD student, Department of Statistics, Faculty of Economics and Business Administration, “Alexandru Ioan Cuza” University of Iasi, Romania, danutza23@yahoo.com

<sup>4</sup> PhD student, Department of Statistics, Faculty of Economics and Business Administration, “Alexandru Ioan Cuza” University of Iasi, Romania, itte7@yahoo.com

The working hypothesis is that, on the basis of structural changes of the economy, the employment rate dynamics has a specific trend characterized by years clusters differentiated on the value of the indexes dynamics and on their sign.

This hypothesis implies the analysis of the economic phenomena in their interdependence and, implicitly, their prognosis. Therefore, the prognosis of a phenomenon depends on the prognosis of the development environment. The forecast of the employment rate for a given horizon (year) depends on the years cluster characterized by the dynamics of the economic, social and political environment specific to the analyzed time interval.

The paper aims to identify the years clusters depending on the influence factors that explain them. To express the population employment rate we use the population employment rate. Its dynamics is positively correlated with the dynamics of the economy.

In order to attain this objective, we suggest a prognosis method for the level of a phenomenon based on the prognosis of the environment. This method consists of the characterization of the environment and of the identification of the clusters using data analysis methods and the prognosis using the discriminant classification function.

For applying this method, we address several problems:

- the environment description, that is its characteristics expressed by the dynamics of the factors that define the studied phenomenon (the choice of the variables and the measurement units);
- the evaluation of the dynamics and the identification of the trend model of the phenomenon (using the dynamics indexes of the studied phenomenon and of their factors);
- the identification of the years clusters for the studied environment (principal component analysis);
- the definition of the classification function (discriminant analysis)

The analysis of the dynamics of a phenomenon, in the classical statistical analysis, consists of the time variation measurement, using the indexes, the average trend identification and the trend extrapolation [Granger, 1980; Coutrot and Droesbeke, 1990;].

Depending on the nature of the phenomenon, the time series analysis supposes the decomposition on components of a time series [Gourieroux and Monfort, 1990; Droesbeke et al., 1990; Bourbonnais and Terazza, 2004] and their extrapolation [Makridakis, Wheelwright, Hyndman, 1997].

When considering time series as stochastic processes, there are used ARIMA models [Lütkepohl, 1993; Hamilton, 1994; Mélard, 1990].

In most studies of time series analyses, one starts from the hypothesis of linearity. This hypothesis is opposite to the evolution of the transition and post-transition phenomena. For such situations, we recommend a forecast model that depends on the economic, social and political environment, which define the analyzed phenomenon. The model recommended supposes, firstly, the identification of the years clusters of the studied phenomenon on the basis of the influence factors dynamics, using the principal components analysis (PCA). Then, for the desired time horizon, the membership cluster is determined, using the discriminant analysis (DA). When the membership cluster is identified, and also the factors that determined such a dynamics, one can stimulate, through appropriate development policies for the influence factors, the desired level of the phenomenon.

The study starts with the evaluation, through indexes, of the dynamics of the employment rate and influence factors - demographic, economic, and demo-economic. Then, we identified the year clusters according to the dynamics of the employment rate, based on the principal component analysis, and estimated the parameters for the classification function of a year in a certain cluster, making use of the discriminant analysis. Eventually, we estimated the parameters of a forecast model of the employment rate dynamics in proportion to the phenomenon dynamics, identified on a years cluster, and not on the trend extrapolation, like in the classical forecast.

We used the official statistics as data source. The data processing was done with the SPSS program.

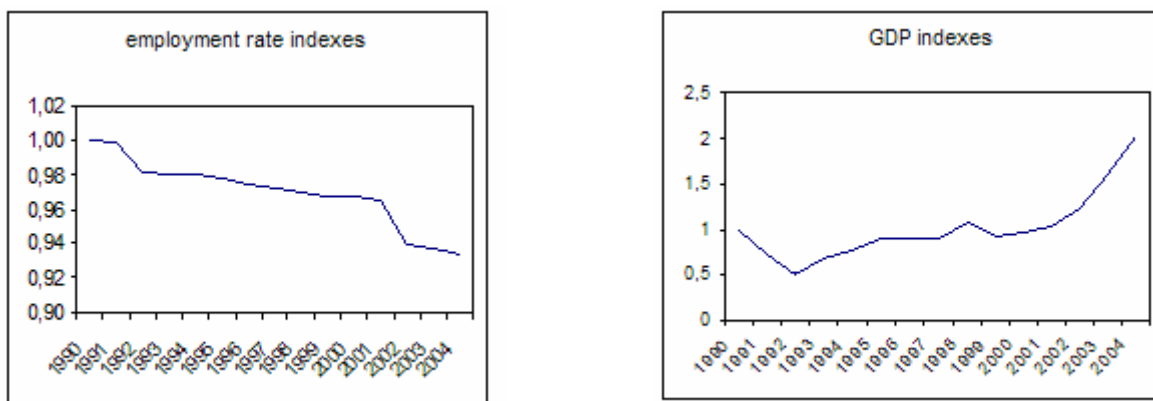
The proposed forecast model may be used to simulate forecast variances, which could be taken into consideration to found strategies of economical development.

Tendencies in the dynamics of the employment rate and influence factors  
The working hypothesis formulated based on the economical increase theory states that there are some links between the dynamics of the employment rate and the dynamics of the influence factors.

### ***1.1. The dynamics indexes for the population employment rate and the dynamics indexes for the GDP value/inhabitant***

The dynamics of the population employment rate in the period 1990-2004 shows a general decreasing trend (see *Figure 1, left*) with specific variations especially under the influence of the process of deindustrialization starting in 1990, the privatization of public enterprises and development of services specific to a market economy (financial, real-estate, banking, stock exchange services) and IT economy.

At the same time, the values of the indexes for the GDP value/inhabitant increased during the same period (see *Figure 1, right*).



**Figure 1.** *The dynamics indexes for the population employment rate and the dynamics indexes for the GDP value/inhabitant in Romania, in the period 1990-2004*

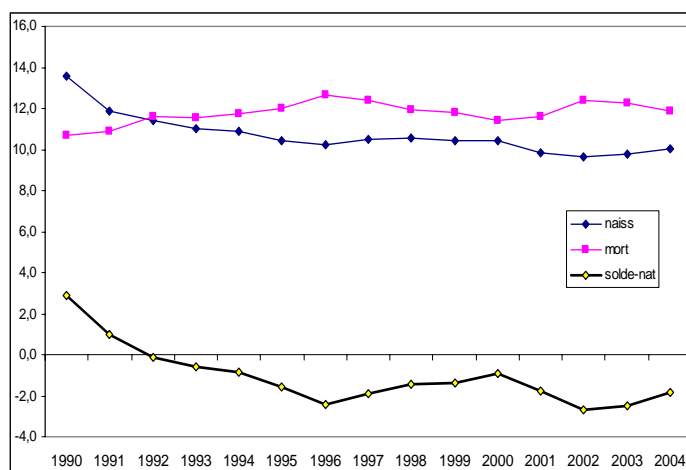
### ***1.2. The dynamics of the demographical phenomena***

The number of population in Romania beginning with 1930 and up to 1990 registered a continuous increase with positive yearly average rhythms, but with a slight decrease between 1956 (1.2%) and 1966-1970 (0.9%).

After 1990, the population number constantly decreased with negative yearly average rhythms, of -1.5%, in 1992, to -0.2%, in the previous decade.

At the last Romanian population census in 18 March 2002, the population number reached 21,680,974 inhabitants. This level was by 1,129,061 inhabitants lower than the level registered at the previous census in 1992. This decrease in the population number may be explained both by the natural decrease (*Figure 2*), as well as by the migratory decrease.

The birthrate and death rate vary from one year to another on the whole period taken into consideration. Beginning with 1992, we deal with a natural decrease (*Figure 2*).



**Figure 2.** *The evolution of the birthrate, death rate and natural decrease in Romania in the period 1990-2004*

The phenomena birthrate and death rate have a falling trend. Still, the number of births decreases even more, the natural decrease being constantly dropping on the overall period.

Another characteristic of the transition period is pointed out by the specific dynamics of the population number on residence areas: urban and rural. The population on residence areas follows a general decreasing trend.

The specificity of this evolution is given by the change of the population dynamics on the two residence areas. If until 1992 there had been an increase of the urbanization index in Romania, reaching to an urban population of 54.3%, after that period we notice a decrease of the urban population. The direction of the dynamics of population on residence areas is inverted beginning with 2001. The steps of the privatization policy of public enterprises show their effects and we notice the unemployment increase. Consequently, the remaining working force directs to the rural environment or to the external migration.

### ***1.3. The dynamics of the migratory phenomena***

Migration in Romania after the 2<sup>nd</sup> World War develops mainly under the influence of changes in the national economical structure.

This phenomenon performed as permanent migration, which had mainly political and ethnical motivations, characteristic to the period before 1989, or as temporary migration generated by economical reasons, especially after 1989.

The legal permanent migration had a reducing trend until 2001-2002. Afterwards, we notice a revival of its increase, having registered 13,082 emigrants in 2004.

The direct effect of emigration generates the appearance of the population shortfall registered at the population census in 2002 compared to 1992. Out of the 1.1 million inhabitants deficit, 362,000 inhabitants emigrated legally, 356,000 are natural decrease and the difference may be explained through illegal emigration, not well known from the statistical point of view.

## **2. The identification of the model of the employment rate dynamics**

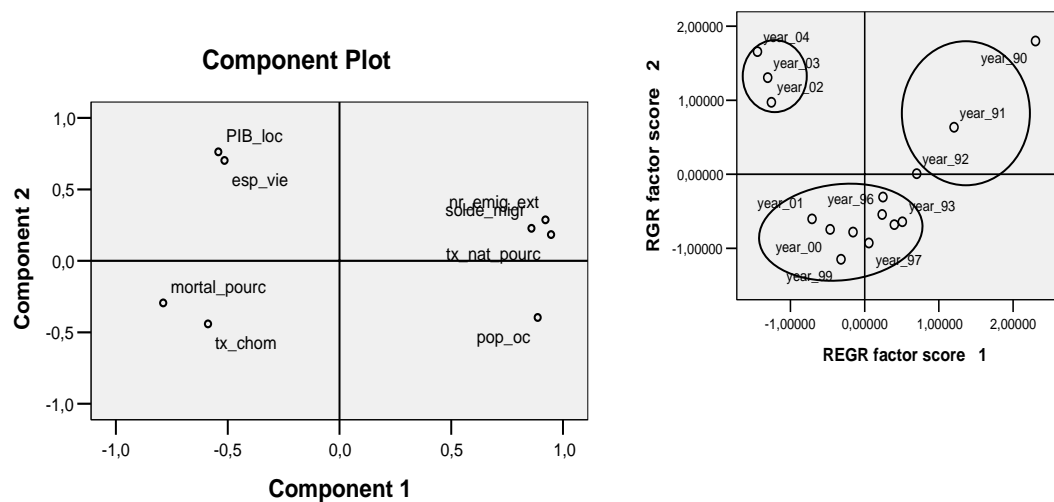
We start from the idea that the different intensity of the factors' influence on the employment rate may materialize into years clusters according to the registered dynamics.

## 2.1. Defining clusters according to the employment rate dynamics

In order to identify the years clusters according to the dynamics of influence factors upon the dynamics of the employment rate, we used the Principal Component Analysis (PCA) method. The variables considered in the PCA are the following:

- death rate,  $X_1$ , (*mortal\_pourc*);
- birthrate,  $X_2$ , (*tx\_nat\_pour*);
- life expectancy,  $X_3$ , (*esp\_vie*);
- unemployment rate,  $X_4$ , (*tx\_chom*);
- GDP value,  $X_5$ , (*PIB\_loc*);
- number of emigrants,  $X_6$ , (*emig\_ext*);
- migratory decrease/increase,  $X_7$ , (*sold\_migr*);
- population employment rate,  $X_8$ , (*pop\_oc*).

The results are shown in Figure 3.



**Figure 3.** Years clusters according to the dynamics of the population employment rate in Romania in the period 1990-2004

The years clusters show characteristics specific to the dynamics of the phenomena analyzed. Therefore:

- the 1<sup>st</sup> cluster consisting of the years 1990-1992 (*tend\_pos*) is characterized by a positive dynamics of the population employment rate, of the population natural increase, of the migratory increase and the number of external emigrants, it is also characterized by a negative dynamics of the GDP value, life expectancy, death rate and unemployment rate;
- the 2<sup>nd</sup> cluster made up of the years 1993-2001 (*tend\_ct*) is characterized by a stationary dynamics of the mentioned phenomena;
- the 3<sup>rd</sup> cluster comprising the years 2002-2004 (*tend\_neg*) is characterized through a positive dynamics of the GDP value, of the life expectancy, death rate and unemployment rate, and a negative dynamics of the population employment rate, the population natural decrease, the migratory decrease and the number of external emigrants;

## 2.2. Defining the discriminant function

In the discriminant analysis we considered as dependant variable the dynamics indexes of the employment rate, defined through the three clusters and as independent variables the dynamics indexes of the influence factors.

The coefficients of the discriminant function, calculated through the discriminant analysis, are shown in Table 1.

Table 1. *The coefficients of the discriminant function*

	group_oc		
	tend_pos	tend_ct	tend_neg
mortal_pourc	108815,1	109810,4	111727,1
tx_nat_pourc	12623,330	12711,779	12914,911
esp_vie	200919,2	202767,5	206376,1
tx_chom	929,096	952,137	977,853
PIB_loc	2555,278	2591,575	2655,410
nr_emig_ext	-,015	-,016	-,016
solde_migr	-5943,700	-5947,881	-6079,347
(Constant)	-159158	-162111	-167954

Fisher's linear discriminant functions

The classification functions of a forecast horizon in a years cluster characterized by a dynamics specific to the employment rate are defined as follows:

for 1<sup>st</sup> cluster (*tend\_pos*):

$$y_{tend\_pos} = -159158 + 108815,1 \cdot X_1 + 12623,33 \cdot X_2 + 200919,2 \cdot X_3 + 929,096 \cdot X_4 + 2555,278 \cdot X_5 - 0,015 \cdot X_6 - 5943,7 \cdot X_7$$

for the 2<sup>nd</sup> cluster (*tend\_ct*):

$$y_{tend\_cts} = -162111 + 109810,4 \cdot X_1 + 12711,779 \cdot X_2 + 202767,5 \cdot X_3 + 952,137 \cdot X_4 + 2591,575 \cdot X_5 - 0,016 \cdot X_6 - 5947,881 \cdot X_7$$

for the 3<sup>rd</sup> cluster (*tend\_neg*):

$$y_{tend\_neg} = -167954 + 111727,1 \cdot X_1 + 12914,911 \cdot X_2 + 206376,1 \cdot X_3 + 977,853 \cdot X_4 + 2655,41 \cdot X_5 - 0,016 \cdot X_6 - 6079,347 \cdot X_7$$

The identification of a years cluster where a given forecast horizon will be “distributed” requires the calculation of scores for each cluster, using the classification discriminant functions shown above. The forecast horizon will be “distributed” in a years cluster for which the highest score was obtained.

### 3. The forecast of the dynamics of the employment rate

Taking into consideration the different dynamics of the phenomena on the three clusters, the achievement of a forecast of the employment rate requires the practice of a method which should consider such influences. A trend extrapolation for the entire period would bring along serious deviations from the real line of the phenomenon evolution. The forecast method we propose is one which should consider the characteristics of the years clusters, pointing out, in fact, the relationships between the dynamics of the influence factors and the dynamics of the employment rate.

This is possible though the application of the discriminant analysis with which we identified the classification functions.

The forecast of the dynamics of the employment rate characteristic to a years cluster requires the identification of the trend model corresponding to the dynamics of each influence

factor, based on which we estimate the level of factor  $X_i$ , where  $i=1,7$  in our case, for the viewed forecast horizon.

### ***3.1. The estimation of the parameters of the trend model of the influence factors and testing their significance***

#### ***a. The population death rate in the period 1990-2004***

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the square model, according to the value of the determination report, where  $R\ square=0,543$ .

The estimated values of the square model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_1 = 0,992 + 0,031 \cdot t - 0,001 \cdot t^2$$

(t=28,825) (t=3,082) (t=-2,487)  
(Sig.=0) (Sig.=0,009) (Sig.=0,029)

where  $t$  is the time variable.

#### ***b. The population natural increase in the period 1990-2004***

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the cubic model, according to the value of the determination report, where  $R\ square=0,898$ .

The estimated values of the cubic model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_2 = 4,608 - 2,096 \cdot t + 0,223 \cdot t^2 - 0,008 \cdot t^3$$

(t=6,308) (t=-5,478) (t=4,076) (t=-3,398)  
(Sig.=0) (Sig.=0) (Sig.=0,002) (Sig.=0,006)

#### ***c. The population life hope in the period 1990-2004***

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the linear model, according to the value of the determination report, where  $R\ square=0,521$ .

The estimated values of the linear model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_3 = 0,991 + 0,002 \cdot t$$

(t=212,523) (t=3,758)  
(Sig.=0) (Sig.=0,002)

#### ***d. The unemployment rate in the period 1990-2004***

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the cubic model, according to the value of the determination report, where  $R\ square=0,542$ .

The estimated values of the cubic model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_4 = 1,359 \cdot t - 0,18 \cdot t^2 + 0,007 \cdot t^3$$

(t=3,475) (t=-3,23) (t=3,022)  
(Sig.=0,005) (Sig.=0,008) (Sig.=0,012)

#### ***e. The GDP value in the period 1990-2004***

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the square model, according to the value of the determination report, where  $R\ square=0,823$ .

The estimated values of the square model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_5 = 0,961 - 0,099 \cdot t + 0,01 \cdot t^2$$

(t=6,311) (t=-2,262) (t=3,861)  
 (Sig.=0) (Sig.=0,043) (Sig.=0,002)

*f. The number of emigrants in the period 1990-2004*

As to the variable « number of emigrants », after estimating parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the cubic model, where *R square*=0,844.

The estimated values of the cubic model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_6 = 109846,5 - 32794,2 \cdot t + 3564,881 \cdot t^2 - 122,996 \cdot t^3$$

(t=8,188) (t=-4,667) (t=3,552) (t=-2,976)  
 (Sig.=0) (Sig.=0,001) (Sig.=0,005) (Sig.=0,013)

*e. The migratory increase in the period 1990-2004*

After estimating the parameters of the models identified graphically (linear, square and cubic) and testing their significance, we chose the square model, according to the value of the determination report, where *R square*=0,823.

The estimated values of the square model parameters, the values of the calculated test statistics, and their associated probabilities are as follows:

$$X_7 = 0,751 - 0,132 \cdot t + 0,006 \cdot t^2$$

(t=5,34) (t=-3,271) (t=2,417)  
 (Sig.=0) (Sig.=0,007) (Sig.=0,032)

**3.2. The estimation of the employment rate for a forecast horizon in a years cluster**

The prognosis of the employment rate dynamics suppose the prognosis of the value of each factor variable for the forecast horizon based on the estimated trend model.

For instance, in the case of the variable *death rate*, the estimated value of the level of this phenomenon in 2007 is:

$$X_1 = 0,992 + 0,031 \cdot 16 - 0,001 \cdot 16^2 = 1,232$$

where *t=16* corresponds to the value set for the time variable for the given horizon.

Similarly, for all the factor variables we obtain estimations of their level for a given forecast horizon.

Based on these estimated values of the factor variables we may classify the horizon, through the discriminant function, for instance 2007, in a years cluster. According to the cluster where the viewed forecast horizon will be “classified”, we estimate the population employment rate.

## Conclusions

We achieved the analysis of the phenomenon dynamics which, consistent with the economic theory, has an impact on the population employment rate. This analysis highlighted a different dynamics of the phenomenon for the period 1990-2004. During the transition period of Romania, the dynamics of the phenomena was different, producing years clusters.

The clusters were identified using the principal component analysis. These clusters are characterized through different dynamics of the phenomena shaping the dynamics of the population employment rate.

Traditionally, the statistical forecast is done by trend extrapolation. Such a forecast which takes into account the trend during all the mentioned years would rely on the theory of a similar evolution for the whole period. In this paper, we pointed out the fact that the dynamics of the phenomena in a transition period is different, being able to identify years clusters on the overall considered period.

Consequently, the forecast through a trend extrapolation on the whole would lead to serious digressions from the real evolution line of the phenomenon.

In order to eliminate this shortcoming, we propose a forecast method of a phenomenon level based on its influential environment. In other words, we analyze the dynamics of the phenomenon determining the employment rate and identify years clusters for the employment rate, using the principal component analysis. Later, based on the discriminant analysis, we estimate the coefficients of the discriminant function for classification for each of the clusters identified.

The next step is to estimate the parameters of the trend models for the dynamics of each influence factor. Based on the coefficients of the trend function we find out the level of the phenomenon for a certain forecast horizon. The estimated values of each factor in a certain forecast horizon are replaced in the classification discriminant function, finding out variances of possible evolutions for the dynamics of the population employment rate.

## References:

1. Aydalot, P., *Economie régionale et urbaine*, Economica, Paris, 1985.
2. Bertoneche, M., Teulie, J., *Théorie macroéconomique*, PUF, Paris, 1977.
3. Blanchard, O., *The Economics of Post-Communist Transition*, Oxford: Clarendon Press, 1997.
4. Boeri, T., Burda, M. C., and Kollo, J., "Mediating the Transition: Labor Markets in Central and Eastern Europe," *CEPR Economic Policy Initiative No.4*, 1998.
5. Bourbonnais, R., Terazza, M., *Analyse des séries temporelles*, Dunod, Paris, 2004.
6. Coutrot, B., Droesbeke, J.-J., *Les méthodes de prévision, Que Sais-je? n°2157*, Presses Universitaires de France, Paris (2e edition), 1990.
7. Droesbeke, J.-J., et al. (editors), *Modélisation ARCH - Théorie statistique et applications dans le domaine de la finance*, Collection Association pour la Statistique et ses Utilisations, Editions de l'Université de Bruxelles, Bruxelles et Editions Ellipses, Paris, 1994.
8. Gourierou, C., Montfort, A., *Séries temporelles et modèles dynamiques*, Economica, Paris, 1990.
9. Granger, C. W. J., *Forecasting in Business and Economics*, Academic Press, New York, 1980.
10. Hamilton, J., *Time Series Analysis*, Princeton University Press, Princeton, 1994.
11. Jaba, E., Iașu, E., *Distribution régionale des migrations en Roumanie dans les années 1989-1997*, in volume „In Memoriam, Profesor Universitar doctor Alexandru Bărbat, Junimea Printing House, Iași, 2000.

12. Jaba, E., *Statistics*, Third edition, Economica Printing House, Bucuresti, 2002.
13. Jaba, E., *Labour Force – Iasi Area. Social Statistics Study*, (Force de travail – Région de Iasi. Etude de statistique sociale), “Alexandru Ioan Cuza” University, Iasi, 1979.
14. Jaba, E., Iatu, E., Pintilescu, C., *Tendances démographiques, phénomènes migratoires et marché du travail en Roumanie postcommuniste*, volume du Colloque de l’ASRDLF, Tusinia, 2006.
15. Jaba, E., Iatu C., Pintilescu C., *Dynamique de la structure économique en Roumanie et l’impact sur le marché de travail*, XLIIe Colloque de l’ASRDLF, Villes et territoires face aux défis de la mondialisation, Dijon, sept. 2005.
16. Lütkepohl, H., *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin, (2<sup>nd</sup> edition), 1993.
17. Makridakis, S., Wheelwright, S. S., Hyndman, R. J., *Forecasting: Methods and Applications*, Wiley, New York (3<sup>rd</sup> edition.), 1997.
18. Melard, G., *Méthodes de prévision à court terme*, Editions de l’Université de Bruxelles, Bruxelles, and Editions Ellipses, Paris, 1990.
19. Richardson, H.M., *Input-Output and Economic Base Multipliers. Looking backward and forward*, in *Journal of Regional Science*, 25, pp. 607-661, 1985.
20. Stiglitz, J.E., Walsh, C.E., *Principes d’économie moderne*, 2<sup>e</sup> édition, Editions de Boeck Université, 2004.
21. Voineagu, V., Furtuna, F., Voineagu, M., Stefanescu, C., *Factorial Analysis of the Social-Economical Phenomena in Regional Profile*, (Analyse factorielle des phénomènes socio-économiques en profil régional), Aramis Print Printing House, Bucharest, 2002.