Savić Mirko, PhD
Assistant Professor
Faculty of Economics in Subotica
Segedinski put 9
24000 Subotica, Serbia
phone: +381641183528 fax:
+38124546486
savicmirko@eccf.su.ac.yu

MULTIVARIATE ANALYSIS OF EMPLOYMENT IN EASTERN EUROPE

Abstract
In the context of transitional changes in the last decade in East European countries, the employment structure is one of the fastest changing areas. What is the best methodology to compare the employment situationin the different countries of the region? Multivariate statistical analysis methods are representing very reliable way of getting the full picture of the problem. Principal components analysis is one of the simplest multivariate methods but it could produce very useful information about the employment in the Eastern Europe in very easy and understand able way.

Key Words
Employment, Multivariate Analysis, Principal Components Analysis

ACM classification
G.3 PROBABILITY AND STATISTICS, *Multivariate statistics*

JEL classification
C4 – Econometric and Statistical Methods: Special Topics, C49 – Other

Former communist countries have lost out onat least a half century of normal economic development. The nature of their growth built serious structural distortions into their economies, which made them highly inefficient, compared to the rest of the world. This led to an unavoidable output decline after the collapse of communism. Making up for this lost time will takeat least 15 years in the case of the Czech Republic, 20 years in the case of Hungary and Poland, and 30 years in the case of Romania.

In the last decade, labour forceactivity rates increased in most Western European countries . In the Eastern European countries employment situation was different. The economic activity rates have decreased both for men and women (the decrease being often larger for men). In a majority of eastern countries women constitute a larger proportion of the total labour force between 40 and 50%. Exception is Turkey because it has considerably less women in the labour force - fewer than 30%.

Total employment as a percentage of the total populationis an indicator of acountry's capacity to support its population. Bosnia and Herzegovina, The Former Yugoslav Republic of Macedonia and Serbia have the lowest employment ratios in the whole Europe. The economic activity rates in these countries are not too much below the average, and the reasons for the low ratio can be high unemployment in Bosnia and Herzegovina and Serbia.

The general trend between in the last decade in East European countries is a continuous increase in the share of total employment in the service sector at the expense of the agricultural and industrial sectors. The re are exceptions, and the share of total employment in agriculture is still highin Albania(72%), Moldova(51%) and Romania(43%).

Women are less likely than men to be employers or selfemployed workers in Eastern Europe, especially in Turkey and Albania. Proportion of employed women among these groups is over 50%, only in Moldova and Belarus with Ukraine, Lithuania, Estonia, Latvia and Russia slightly below 50%.

In Lithuania the proportion of women among legislators, senior officials and managers is highest in Eastern Europe (47%). In the rest of the region there is clear majority of men in these occupations with lowest recorded proportion of women being in Turkey at 8%.

Part time employment remains a female domain, but varies considerably among countries . In Moldova the proportion of all employed women working part time is less than 1%. In most of the countries, the trend has been increasing since 1990. Ingeneral, employed persons in Eastern Europe have longer working hours than employed persons in Western Europe (except in Greece and Iceland). In all countries mens pend more time in paid work than women, but the difference between women and men is more pronounced in Western Europe than in Eastern Europe.

The employment is very dynamic and complex economic category that can be analyzed with different scientific methods including quantitative ones. Multivariate statistical analysis offers arange of methods and techniques, for example discriminant analysis, conjoint analysis, principal components analysis, AID and CHAID methods, factor analysis, cluster analysis, correspondence analysis etc. These methods gradually are finding their placeindifferent areas of economy. At themoment, marketing research represent the area with most intensive use of multivariatestatistical methods. If they can find their placeinmarketing, the nmorethanlogically is tospreadthe useof the semethods also intoother areas of economic science.

This paper alsonotes that employment structure dynamics is highon the priority list of countries around the worldby virtueof the implementationinmany countries of longitudinal panel surveys andother instruments measuringLabour Market Dynamics (LMD).

1. METHODS AND DATA

Principal components analysis represents one of the simplest multivariate methods. The objective is to takeporiginal variables (X_1, X_2, ..., X_p) and findcombinations of the setoproduce indices or new variables (Z_1, Z_2, ..., Z_p) that areuncorrelatedinorder of their importance, and that describe the variationin the data. The lack of correlation means that the indices aremeasuringdifferent "dimensions" of the data, and the ordering is such that Var(Z1) ≥ Var(Z2) ≥ ... ≥ Var(Zp). The new variables arethe n the principal components. Inprincipal components analysis, there is always the hope that the variances of most of the new variables will besolow as to benegligible. Inthat case, most of the variationin the full dataset can be a dequately described by the few Z variables with variances that are not negligible, and some degreeof economy is the n achieved.

Principal components analysis will not always work, in the sensethat alargenumber of original variables arereducedto asmall number of transformedvariables. Indeed, if the original variables arenot correlated, the n the analysis achieves nothing. The best results areobtainedwhenthe original variables arevery highly correlated, positively or negatively. If that is the case, the n it is quiteconceivable that 20 or moreoriginal variables can beadequately representedby two or threeprincipal components. If this desirablestate of affairs does occur, the n the important principal components will beof someinterest as measures of the underlying dimensions in the data. It will alsobeof valuetoknow that the reis agooddeal of redundancy in the original variables, with most of them measuringsimilar things.

The principal component analysis will helpus makethe plot of East European countries against their values for twoprincipal components. The pictureis rather meaningful interms of what is knownabout employment in the region. The countries with similar employment situationwill be grouped togethe r and it is possibletosee the positionof eachcountry in the comparison with other countries.

The procedurefor principal component analysis starts with dataon pvariables for nindividuals, as indicatedinTable 1.

Table 1: **The form of datafor a Principal Components Analysis**

Case	X_1	X_2	...	X_p
1	x_{11}	x_{12}	...	x_{1p}
2	x_{21}	x_{22}	...	x_{2p}
...	...	...	...	...
n	x_{n1}	x_{n2}	...	x_{np}

The first principal component is the n the linear combinationof the original variables ($X_1, X_2,...,X_p$):

$$Z_1 = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \quad (1)$$

that varies as much as possiblefor the individuals, subject tocondition that

$$a_{11}^2 + a_{12}^2 + K + a_{1p}^2 = 1 \qquad (2)$$

Thus the variance of Z1, Var(Z1), is as large as possible given this constraint on the constants a1j. The constraint is introduced because if this is not done, the n Var(Z1) can be increased by simply increasing any one of the a1j values. The second principal component is

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \ldots + a_{2p}X_p \qquad (3)$$

It is chosen so that Var(Z2) is as largeas possiblesubject tothe constraint that

$$a_{21}^2 + a_{22}^2 + K + a_{2p}^2 = 1 \qquad (4)$$

and also to the condition that Z1 and Z2 have zero correlation for the data. Further principal components are defined by continuing in the same way. If there are p variables, the n there will be up to p principal components.

The steps ina principal components analysis arethe following:

1. Start with coding the variables to have zero means and unit variances. Sometimes it is omitted when it is thought that the importance of variables is reflected in their variances. In the case of employment all the variables will have the same importance.
2. Calculation of the covariance matrix. This is a correlation matrix if step 1has been done.
3. Finding the eigenvalues and the corresponding eigenvectors. The elements of eigen vectors matrix are the coefficients of the principal components while the eigenvalues are their variance.
4. Discarding any components that account for only a small proportion of the variation in the data. For example, for the 20 variables there will be 20 principal components but only the first four component account over 90% of the total variance. On this basis, the other 16 component may reasonably beignored.

The principal components analysis is giving us very useful information about differences in employment structure but it could be just the start for more serious multidimensional analysis. For example, factor analysis, on the basis of the principal components analysis, helps to reduce avast number of variables (for example, all the questions tapping several variables of interest in a questionnaire) to a meaningful, interpretable, and manageable set of factors.

The data were collected from UNECE (United Nations Economic Commission for Europe) statistical database maintained by the Statistical Division of the UNECE Secretariat. It provides detailed statistical informationon countries in Europe, North America and Central Asia. These are probably not the latest data on employment in the East European region, but it is fresh enough to get the insights.

Only several variables were taken for analysis but it is just a start for further and moredetailed scientific research.

2. RESULTS

It is appropriate tobegin with step1 of the four parts of the analysis that have just been described. Standardization of the measurements ensures that they all have equal weight in the analysis.

The covariance matrix for the standardized variables is the correlation matrix. The correlation matrix for all variables shows that correlation variables arenot particularly high, which indicates that several principal components will berequired to account for the variation in the data. The eigenvalues of this matrix are shown in the table2.

Table 2: **Eigenvalues of correlation matrix, and relatedstatistics (Employment EEC + Developed Countries in Employment in Eastern Europe) Activevariables only**

	Eigenvalue	% Total	Cumulative	Cumulative
1	2.633371	43.88952	2.633371	43.8895
2	2.065042	34.41736	4.698413	78.3069
3	0.821423	13.69038	5.519836	91.9973
4	0.371016	6.18360	5.890851	98.1809
5	0.109108	1.81846	5.999959	99.9993
6	0.000041	0.00069	6.000000	100.0000

Table 3: **Eigenvectors of correlation matrix (Employment EEC + Developed Countries in Employment in Eastern Europe) Active variables only**

	Principal Components					
	1	2	3	4	5	6
Agriculture	0.278519	0.608002	0.189799	0.085429	0.015234	0.713587
Industry	0.534732	0.142312	0.100197	0.728522	0.015719	0.390867
Services	0.015761	0.648560	0.303039	0.385817	0.022126	0.581330
% of Women	0.137201	0.365365	0.902360	0.093115	0.157320	0.003238
% Unemployment	0.559808	0.183099	0.004075	0.361583	0.722708	0.004858
% Youth Unemployment	0.551389	0.149781	0.218709	0.416761	0.672291	0.006121

The corresponding eigenvectors are shown in table 3, standardized so that the sum of the squares of the coefficients is one for each of them. The seeigen vectors are providing the coefficients of the principal components.

The Scree Method is one of the approaches to deciding how many principal components to select as themost important. The analyst looks for natural break points in the percentage of total variance dependent from each principal component. Figure 1 shows a scree plot for Eastern European countries.

It is a matter of judgment as to how many components are important. To some extent, the choice of the number of components that are important will depend on the use that is going to be made of them. For the present case, it will be assumed that a small number of indices are required in order to present themain aspects of difference between the countries, and for simplicity only the first two components will be examined further. Between them, they account for about 78% of the variation in the original data.

The first component is

$$Z_1 = 0.278519\,X_1 - 0.534732\,X_2 + \\ + 0.015761\,X_3 + 0.137201\,X_4 - \\ - 0.559808\,X_5 - 0.551389\,X_6$$

As the analysis has been done on the correlation matrix, the variables in this equation are the original values after they have each been standardized to have a mean of zero and standard deviation of one.

The second principal component is

$$Z_2 = - 0.608002\,X_1 + 0.142312\,X_2 + \\ + 0.64856\,X_3 + 0.365365\,X_4 - \\ - 0.183099\,X_5 - 0.149781\,X_6$$

Figure 1: **Principal Components Scree Plot**

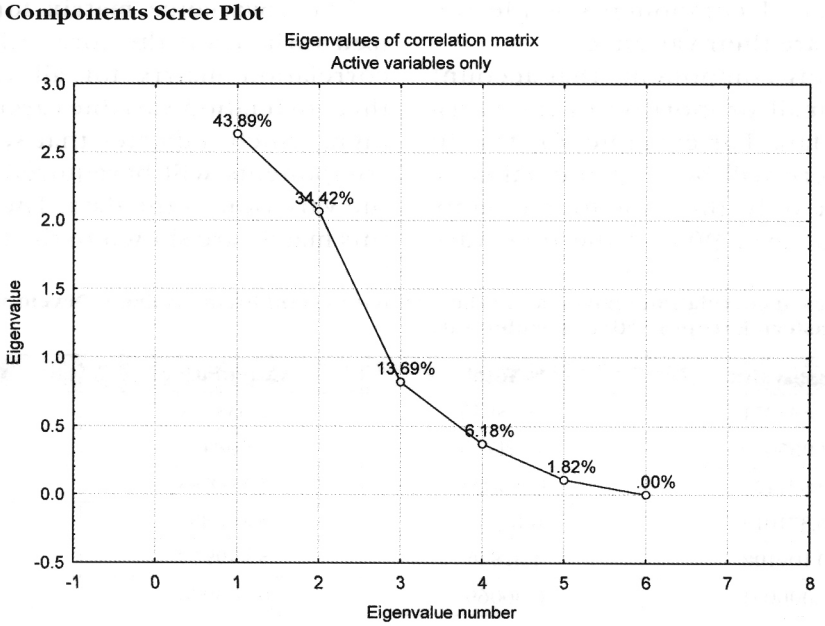


Figure 2 shows a plot of 21East European countries with three West European countries (Norway, Switzerland and Germany) for comparative estimate. The plot shows the positions of the countries against their values for the first two principal components. The picture is certainly rather meaningful in terms of what is known about the employment in the countries .

The plot on the figure 2 is the result of principal components analysis. The useof this result is obvious. With one look at the plot one can get the full picture about the employment in the region andthe position of each country.

The West European countries are grouped in the upper left corner with very similar employment situation in all three major economic sectors (agriculture, industry and services) and with similar values for almost every other variable.

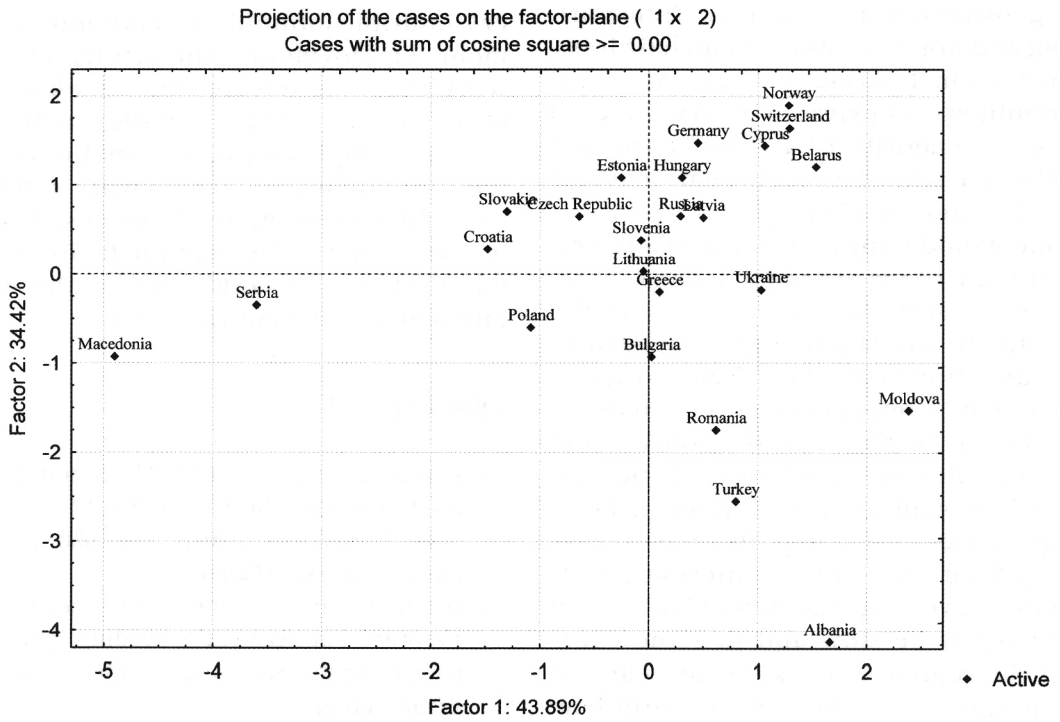
It is interesting to notice that most similar employment structure to West European countries have Cyprus, Belarus and Hungary. All countries from former USSR are very close to one another, with exception of Moldova. The reason for this is that Moldova has larger percentage of employees in agriculture, the percentage of employed women is above 50% and percentage of total unemployment is very low.

It is very easy to notice that themost dramatic difference in employment structure is in Serbia, Macedonia and Albania. Serbia and Macedonia have the same problems: very low percentage of employment in the agricultural sector and largest percentage of total unemployment in the whole Europe. Albania has the problem on the opposite side: the greatest percentage of employment in the agricultural sector and beside that, the very low percentage of employed women.

In Eastern Europe unemployment rate is more varied than in the rest of the continent, but themajority of countries had higher unemployment rates in recent years. In most countries there are small differences between women and men in unemployment rates. However, in Greece, Turkey and Albania considerably more women are unemployed than men. In Lithuania the situationis the other way round - the unemployment rate for men is higher than for women.

In a majority of region countries there is a higher youth unemployment rate for women than for men, but in many countries the differences are not large and may be subject to annual variations based on which sectors of the economy are hardest hit by unemployment. The youth unemployment rate is generally higher than the overall

Figure 2: **The plot diagram on the basis of the scores for the first two principal components**



unemployment rate through out the East European region.

In some countries more than half of all the unemployed have been in that position far morethan 12 months. The recorded longterm unemployment rate is highest in Albania, Slovenia (65%) and Bulgaria (62%).

The unemployment rates were high for all East European counties except for Belarus, Cyprus, Moldova and Ukraine. Since by present standards any rate below 3% can be regarded as exceptionally good, Moldova would stand out in Eastern Europe. Ten countries in the region (Albania, Bulgaria, Croatia, Estonia, Latvia, Lithuania, Macedonia, Poland, Serbia and Slovakia) have double digit rates and hence can be classified as countries with extremely worrying unemployment rates, the same can be said about the other four nations since they have rates in excess of 7% (CzechRepublic, Greece, Russia and Turkey). With this arbitrary classification, only Moldovacan claim to have performed extremely well. However, the reader should note that absolute levels of unemployment for these countries are not strictly comparable owing to differences in measurement techniques.

When it comes to employment by sectors, of particular interest is the relative size of the service sector. This is mainly the tertiary sector (it comprises such divergent items as banking, distribution, insurance, transport, catering and hotels, laundries and hairdressers, professional services of amore varied kind, publicly and privately provided, etc.) and was, in majority of Eastern European countries, the largest inrecent years, exceeding 50% in majority of them.

As one would expect (since it is anatural characteristic of development), most of the countries considered show adecline in the percentage of the labour force engaged in agriculture. However, many East European countries have double digit percentages.

Conclusion Employment is a political and socioeconomic issue which needs to be tackled in all its manifestations. Eastern European countries are going through transitional period and the achievement of acceptable levels of manpower utilization will inevitably be slow and in some countries of the region may take many years. In addition, the reis the longer term problem: the effect of evolving structures of the labour force, attitudes to work and changing social objectives which may affect employment in a fundamental sense.

EU Social Policy Agenda and European Employment Strategy emphasise the importance of ensuring a positive, mutually reinforcing interaction between economic, employment and social policies among the countries . Good employment and social policies are needed tounder pin productivity and to facilitate the adaptation to change. They alsowill play an essential role towards the full transition to the knowledge based economy.

It is necessary tomonitor not only the employment in each country but to compare the employment charactestistics among the countries . The more countries from Eastern Europe are getting closer to EU membership the more complicated is the analysis of current situation. Multivariate analysis is very useful tool for monitoring the employment, in this case in the region of Eastern Europe.

The principal components analysis represented in this paper is conducted on the basis of just a few employment variables and many key indicators of the labour market and labour market dynamics aremissing (labour market transitions, occupational and geographical mobility, discouraged workers, multiple jobholders, temporary employment, employment contractors, shiftwork, work time and leisure time, work at home, time related under employment, in a dequate employment, international movement of labour, etc.). This work represents only the beginning of themoreserious and deeper analysis, for example factor analysis, which should be conducted on the basis of much more relevant variables.

REFERENCES

1. European Commission (2002), National Reports on the Demographic Situation in 12 Central European Countries, Cyprus and Malta in 2000 Eurostat Working Papers
2. Group of authors (2002), European Union and Eastward Enlargement - Reading, The International Summer School in European Studies, Belgrade, Serbia

3. Hardarson O. S. (2005) The EU Labour Force Survey and Indicators of Quality inWork, Eurostat, Statistical Commission and Statistical Office of the UN Economic Commission for European Communities (UNECE) // http://www.unece.org/ /stats/documents (referred on 15/09/2006)

4. UNECE, United Nations Economic Commission for Europe (2005), Trends in Europe and North America // http://www.unece.org/stats/trends2005 (referred on 5/10/2006)

Biography:

Savić Mirko is assistant professor at the Faculty of Economics in Subotica, Serbia. Heis teaching statistics, business statistics and multi variate statistical analysis at the Deparment of Business Informatics and Quantitative Methods. Hehas published more than 20 articles and four books in the areas of statistics, multivariatestatistical methods andstatistical software. Member of MENSA.