QUANTITATIVE METHODS IN ECONOMICS
Vol. XIII, No 2, 2012, pp. 40 – 47

UNEMPLOYMENT RATE FOR VARIOUS COUNTRIES SINCE 2005 TO 2012: COMPARISON OF ITS LEVEL AND PACE USING FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

Stanisław Jaworski
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: stanislaw_jaworski@sggw.pl
Konrad Furmańczyk
Department of Applied Mathematics
Warsaw University of Life Sciences – SGGW
e-mail: konrad_furmanczyk@sggw.pl

Abstract: We apply the functional principal component analysis to compare the unemployment rate in euro area, Japan and USA since 2005 to 2012. For preprocessing analysis we used B-splines system with roughness penalty for smoothing the data. The analysis enables to reveal the most important type of variation in unemployment rate and its pace's in examined countries.

Keywords: B-splines basis system, functional principal component analysis, unemployment rate

INTRODUCTION

The unemployment rate is an important indicator with both social and economic dimensions. The time series analysis of unemployment are used by public institutions and the medias as an economic indicator. The banks may use this data for business cycle analysis. The general public might also be interested in changes in unemployment rate. Rising unemployment rate makes an increased pressure on the governments in order to spend on social benefits and cause a reduction in tax revenue. Rapid increase of unemployment rate may be a symptom of crisis in economy but its fixed decrease may be a signal for grown in the economy (for more information *see* . E. Burgen et al. (2012)).

In the paper we analyze seasonally adjusted monthly unemployment rate in various countries from 2005 to 2012 for euro area, USA and Japan. The source of the data is the EUROSTAT report (see *ec.europa.eu/eurostat)*. The unemployment rate is considered as a benchmark to ensure comparability of conditions of world economy. Although we should be aware of the definitional and technical pitfalls involved in the preparation of several unemployment series emanating from different sources of various countries.

Thus we expect the interpretability of the data comes not only from inspecting the level of the unemployment rate but also from the pace of the rate. Thus the great emphasis should be placed on getting sensible and stable estimation of pace. For this reason we decided to smooth the series by regular functions, possessing one or more derivatives.

In the chapter *Methods* we shortly described some mathematical tools and then in next chapter the conclusions were drown. Necessary computations were carried out by *fda* R package (see www.r-project.org).

METHODS

We assumed that unemployment rate in the *i*-th country is of the form

$$y_{ij} = x_i(t_j) + \varepsilon_{ij}, \quad j = 1, 2, \ldots, N$$

where $x_i(t) = \sum_{i=1}^{K} c_{ik}\phi_k(t)$ and $\{\phi_k\}$ is B - splines basis system (see E.W. Weisstein) and ε_{ij} is an unspecified random error.

We used the penalized sum o squared errors fitting criterion to estimate $x_i(t)$, that is we minimized

$$\sum (y_{ij} - x_i(t_j))^2 + \lambda \int \left[D^2 x_i(s)\right]^2 ds$$

with respect to coefficients $\{c_{ik}\}$. Details of this approach can be found in Ramsey an Silverman (2005).

The smoothing parameter λ measures the rate of exchange between fit to the data in the first term and variability of the function x in the second term. For small λ the curve x tends to become more and more variable since there is less and less penalty on its roughness. In practice we chose parameter λ minimizing Generalized Cross-Validation (GCV) measure with respect to λ. (see Ramsey and Silverman (2005), p.97).

After smoothing the data we carry out functional principal component analysis. This approach was taken by Besse and Ramsey (1986), Ramsey and Dalzell (1991) and Besse, Cardot and Ferraty (1997). Functional principal analysis can be defined as the search for a probe that reveals the most important type of variation in data.

In the first step we search function ξ_1 which maximize sample variance

$$\frac{1}{n}\sum_{i=1}^{n} {f_{i1}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\int \xi_1(s)x_i(s)ds\right)^2$$

subject to

$$\int {\xi_1}^2(s)ds = 1.$$

In the second step we find ξ_2 such that $\int \xi_1(s)\xi_2(s)ds = 0$ and ξ_2 maximize

$$\frac{1}{n}\sum_{i=1}^{n} {f_{i2}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\int \xi_2(s)x_i(s)ds\right)^2$$

subject to

$$\int {\xi_2}^2(s)ds = 1.$$

Next, in m-step we find ξ_m such $\int \xi_k(s)\xi_m(s)ds = 0$ for $k < m$ and maximize

$$\frac{1}{n}\sum_{i=1}^{n} {f_{im}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\int \xi_m(s)x_i(s)ds\right)^2$$

subject to

$$\int {\xi_m}^2(s)ds = 1.$$

Very often the data are presented as a points in the graph in the first two principal components ξ_1, ξ_2. In this case the criterion of quality of functional principal components has the form

$$\frac{\sum_{j=1}^{2}\sum_{i=1}^{n} {f_{ij}}^2}{\sum_{j=1}^{K}\sum_{i=1}^{n} {f_{ij}}^2}100\%.$$

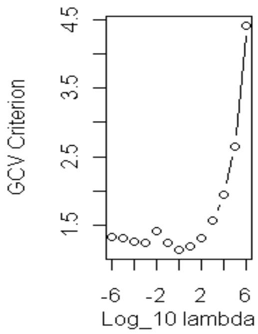
This formula compute percentage of variability of the first two principal components.

DATA ANALYSIS AND RESULTS

In this chapter we consider unemployment rate in Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, United Kingdom, Norway, Croatia, Turkey, United States and Japan.

We used B-splines system and we have taken smoothing parameter $\lambda = 10$. The choice of the parameter's value was based on generalized cross validation plot (Figure 1).

Figure 1. GCV plot.



Source: own preparation

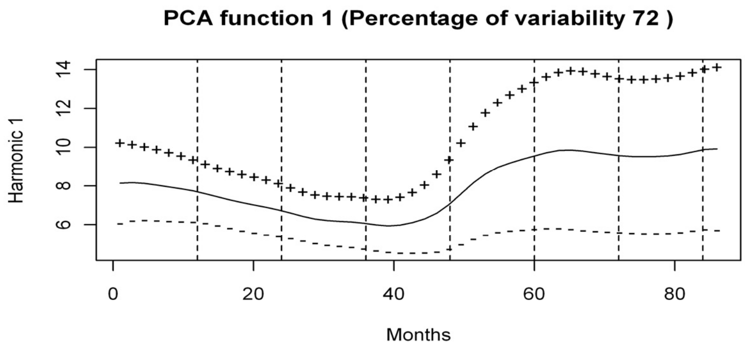
The generalized cross validation measure is minimized at $\lambda = 1$ and is slightly higher for $\lambda = 10$. We chose $\lambda = 10$ to obtain more stable unemployment pace estimate then in the case for $\lambda = 1$. We found that the residual plots were quite good. Next the smoothed data were explored by functional principal components analysis. The first two principal components account for 92% of the total variation. They are presented in Figure 2 and Figure 3 as perturbations of the mean unemployment rate, that is $\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t)$ is presented by solid line and $\hat{\mu} \pm C_m \xi_m$, where n is the counts of countries, $m = 1, 2$ are presented by pluses and minuses. Constant C_m is given by

$$C_m = \frac{1}{n}\sum_{i=1}^{n} f_{im}^2 .$$

Observe (Figure 2) that the greatest variation between unemployment rate of various countries can be found since 2009 (from 40 month). Countries with high value of the first component relate to the countries with high unemployment rate.

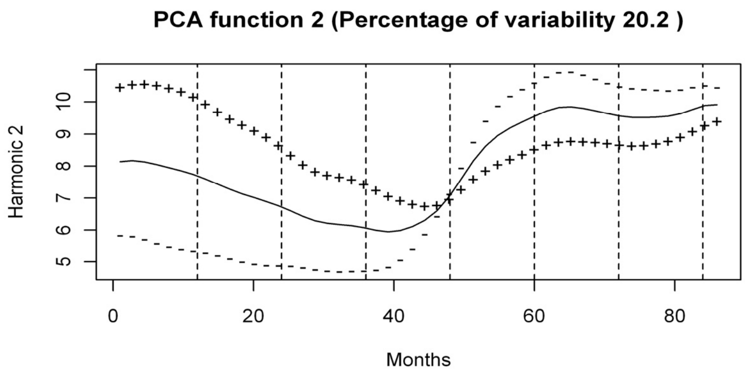
The second large unemployment rate variation is explained by the second component (Figure 3). It expresses the difference between such countries like Ireland and Poland. The unemployment rate is relatively low in Ireland and high in Poland until 2009 and after the date the difference is opposite (Figure 4).

Figure 2. The first principal component as perturbations of the mean unemployment rate



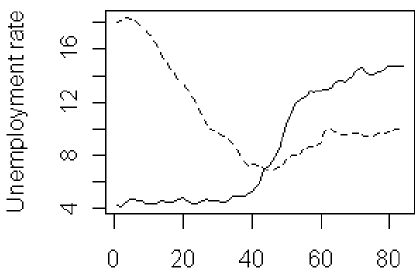
Source: own preparation.

Figure 3. The second principal component as perturbation of the mean unemployment rate



Source: own preparation

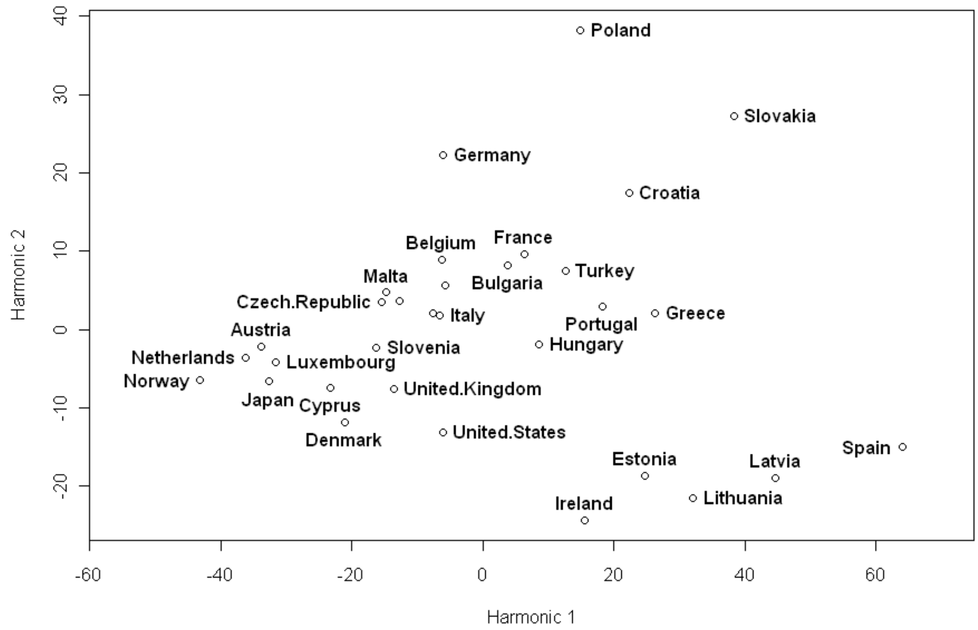
Figure 4. Unemployment rate of Poland (dashed line) and Ireland (solid line)



Source: own preparation

A good insight into the differences between countries can be made by plotting principal scores (Figure 5). Norway, Netherlands, Austria and Japan are placed to the left side of the plot. It means that the countries have low unemployment rate. Opposite side of the plot relates to the countries with large unemployment rate. Spain is the special example of them.

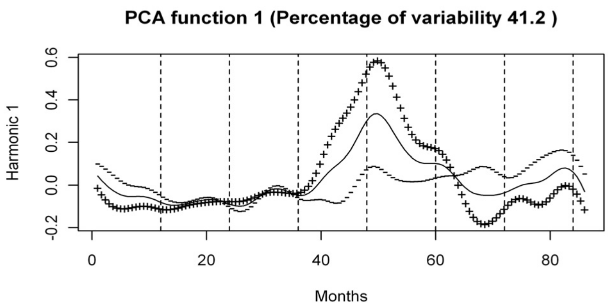
Figure 5. Principal components of unemployment rate



Source: own preparation

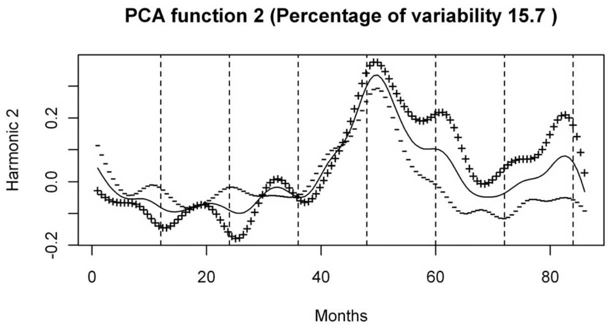
Countries close to the top of the plot are able to cope with the unemployment since the crisis in 2008-2009 than the countries close to the bottom of the plot. The pace of unemployment rate in various countries was investigated in the paper by taking the first derivative of the smoothed unemployment rate series. Then functional principal component analysis was provided for the derived functions. The outcomes are visualized in Figures 6,7 and 8. It is seen that Estonia, Lithuania and Latvia were strongly influenced by the crisis and their unemployment rates extremely increased. These states quite good are managing with the problem of unemployment. It is encouraging that the growth in unemployment slowed there and became a decreasing. In contrast, in Greece, Bulgaria and Croatia the unemployment rate was higher and is increasing faster than in the rest of states.

Figure 6. First principal component as perturbations of the mean unemployment pace



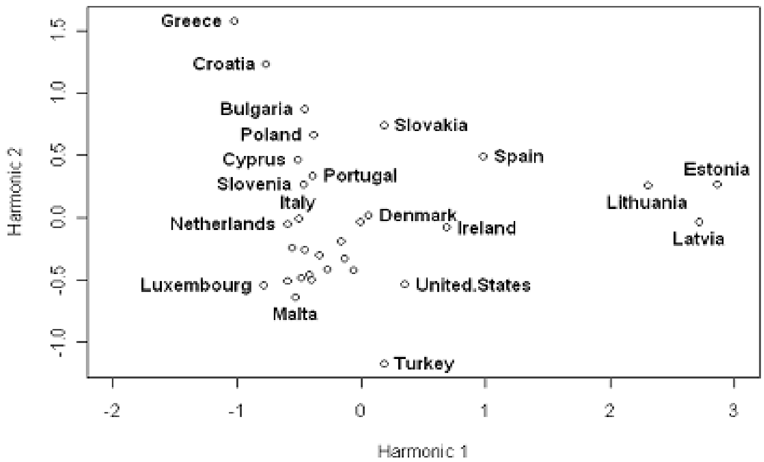
Source: own preparation

Figure 7. Second principal component as perturbations of the mean unemployment pace



Source: own preparation

Figure 8. Principal components of unemployment pace



Source: own preparation

SUMMARY

Functional principal component analysis revealed the most important type of variation from the unemployment rates. The analysis gave us possibility to find out which states were the most influenced by the crisis and in which way. The 2008-2009 crisis highly separated states with respect to differences in unemployment rates but influenced them in different way. After the crisis a list of unemployment rates for various countries changed its order. Some states with high unemployment rate now have it at moderate level. Some states have dynamically growing unemployment rate. In the beginning of 2012 the difference in unemployment rates was high but what is promising it is a flattening of the dynamic of the unemployment rates.

REFERENCES

Besse P.C., Cardot H. and Ferraty F. (1997) Simultaneous nonparametric regressions of unbalanced longitudinal data. Computational Statistics and Data Analysis. Vol. 24, 255-270.

Besse P. Ramsay J.O. (1989) Principal components analysis of sampled functions, Psychometrika, 51, 285-311.

Burgen E., Meyer B. and Tasci M. (2012) An Elusive Relation between Unemployment and GDP Growth: Okun's Law. Cleveland Federal Reserve Economic Trends.

Ramsay J.O. and Dalzell C.J. (1991) Some tools for functional data analysis (with discussion). Journal of the Royal Statistical Society Series B, 53:539-572.

Ramsey J.O. and Silverman B. W. (2005) Functional Data Analysis. Second Edition, Springer, NY.

Weisstein E.W. B-Spline. From MathWorld. A Wolfram Web Resource. http://mathworld.wolfram.com/B-Spline.html