STATISTICS IN TRANSITION new series, March 2017
Vol. 18, No. 1, pp. 139–150

THE ACHIEVEMENTS OF STUDENTS AT THE STAGES OF EDUCATION FROM THE SECOND TO FOURTH USING FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

Mirosława Sztemberg-Lewandowska[1]

ABSTRACT

Significant demographic phenomena can be observed in Poland – the number of school age population is decreasing. It affects higher education since the immediate effect of demographic changes is the drop in the number of students. The analysis of the level of future students' knowledge also remains an important aspect of the problem.
The purpose of the article is to compare the level of knowledge presented by students at the subsequent stages of education in the period 2009-2015. The research covers the average exam results received on graduation from the second, third and fourth stage of education. Functional principal component analysis, which is based on functional data, will be applied in the study. This method allows an analysis of dynamic data.

Key words: level of knowledge, functional data, functional principal component analysis, longitudinal data.

1. Introduction

Demographic conditions exert a direct impact on higher education. Since 1990 Polish higher education has been going through a period of continuous and dynamic growth resulting from the population boom, lasting for almost 15 years. Starting from 2006, the first symptoms of this trend collapse became noticeable as the number of students was gradually decreasing. A drop in the number of 10-year-olds was recorded in each consecutive year and thus the reversal of the trend favourable for Polish higher education. This process has been continuing until today causing the ongoing decrease in the number of the traditional college-age population.

Figure 1 presents the number of students graduating from the subsequent stages of education. The students who graduated from primary school (PS) in

[1] Wrocław University of Economics. E-mail: miroslawa.sztemberg-lewandowska@ue.wroc.pl.

2003 are marked with the letter A on x-axis, in 2006 - middle school and 2009 - secondary school, i.e. theoretically those born in the same year. It can be observed that slight local extremes of the PS students correspond to larger fluctuations in the number of secondary school students. The number of PS graduates reached the minimum (M) level in 2006 and of secondary school graduates in 2012, whereas a year later – the local maximum was achieved. The number of PS graduates has been stabilizing since 2014, i.e. the students who will graduate from the middle school (MS) in 2017 and the secondary school in 2020.

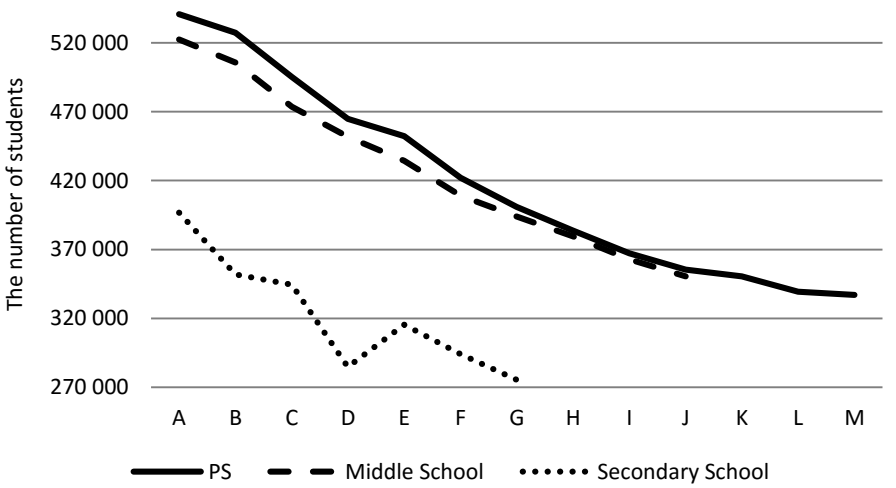

Figure 1. The number of students graduating from the subsequent stages of education

Source: *author's compilation based on CKE [Central Examination Commission] data.*

A	PS 2003	M 2006	S 2009
B	PS 2004	M 2007	S 2010
C	PS 2005	M 2008	S 2011
D	PS 2006	M 2009	S 2012
E	PS 2007	M 2010	S 2013
F	PS 2008	M 2011	S 2014
G	PS 2009	M 2012	S 2015
H	PS 2010	M 2013	
I	PS 2011	M 2014	
J	PS 2012	M 2015	
K	PS 2013		
L	PS 2014		
M	PS 2015		

The level of knowledge presented by future students is another important aspect affecting higher education. For this reason, the level of knowledge presented by students at the consecutive stages of education in the period 2009-2015 was compared. The research covers average exam results received on graduating from the second, third and fourth stages of education. Functional principal component analysis was applied in the study.

2. Methodology

Principal component analysis (PCA) is based on the transformation of original variables into the set of new and mutually orthogonal variables referred to as principal components [Harman 1975]. Functional principal component analysis (FPCA) is characterized by the advantages of a classical principal component analysis and, moreover, allows for the analysis of dynamic data. The type of data is the basic difference between these two methods: PCA is based on multivariate data, whereas FPCA is based on functional data. Functional data take the form of curves and trajectories, i.e. the sequence of individual observations rather than just a single observation [Hall and Hosseini-Nasab 2006, Krzyśko et al. 2012].

In the case of the functional principal component analysis (FPCA) each principal component is presented as the principal component weight function, also referred to as the time dependent eigenfunction $\xi_j(t)$ [Daniele 2006, Ramsay and Silverman 2005]. Eigenfunction maximizes the principal component function variance:

$$v(t,s) \overset{def}{=} \frac{1}{n-1}\sum_{i=1}^{n}\left[x_i(t)-\bar{x}(t)\right]\left[x_i(s)-\bar{x}(s)\right] \tag{1}$$

Similarly to the classical PCA, in the case of the functional one the problem is the function variance distribution:

$$v(t,s)=\sum_j \lambda_j \xi_j(t)\xi_j(s) \tag{2}$$

where $\lambda_j, \xi_j(t)$ satisfy eigenequation:

$$\left\langle v(u,), \xi_j \right\rangle = \lambda_j \xi_j(s). \tag{3}$$

and eigenvalues are positive and non-decreasing:

$$\lambda_j \overset{def}{=} \int_T \xi_j(t) v(t,s) \xi_j(s) dtds. \tag{4}$$

Eigenfunctions satisfy the condition:

$$\int_T \xi_j^2(t)dt = 1 \quad \text{and} \quad \int_T \xi_j(t)\xi_i(t)dt = 0 \quad (i<j). \tag{5}$$

Eigenfunctions define the principal components of variation between the sampling functions x_i [Ingrassia and Costanzo 2005, Hall et al. 2006].

3. The achievements of students

3.1. The second stage of education

Since April 2002 the exam taken by students graduating from the sixth grade of primary school has been held. The standards of examination requirements constitute the basis for carrying out the final test. Until 2014 the standards were grouped into five trans-subject categories:

- reading,
- writing,
- reasoning,
- use of information,
- practical application of knowledge.

In 2015 the test consisted of two parts:

- the first part – covers tasks in Polish and maths,
- the second part – covers tasks in a modern foreign language.

A sixth grader takes the test in one of the following foreign languages: English, French, Spanish, German, Russian and Italian. A student can choose only the foreign language learned at school as an obligatory subject.

Figure 2 illustrates the average percentage results in particular subjects received by the sixth graders from their graduation test in the period 2002-2014.

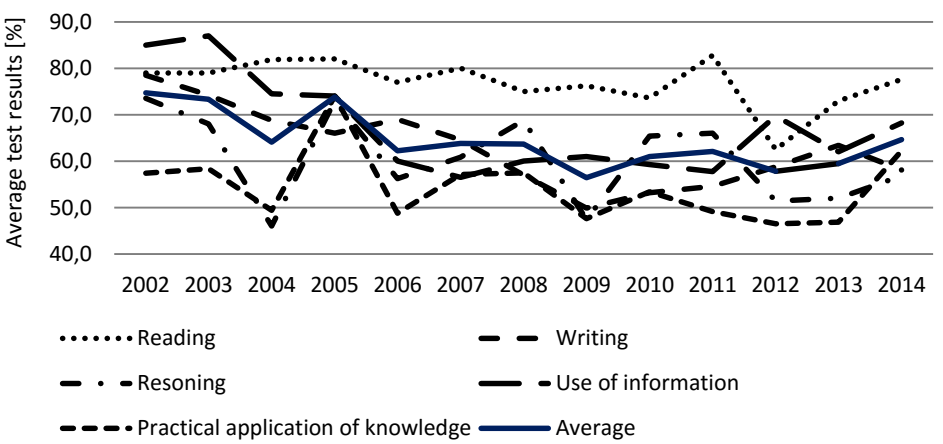

Figure 2. Average test results achieved on graduating from the sixth grade of PS in particular subjects [%]

Source: author's compilation based on CKE [Central Examination Commission] data.

The average result in all subjects is denoted by a solid line and presents the decreasing tendency in the period 2002-2014.

The functional principal component analysis allowed for distinguishing two component functions. The practical explanation of the functional principal components is supported by the graphs showing each component deviation from the average in all subjects (Fig. 3).

The first weight function (72%) The second weight function (18%)

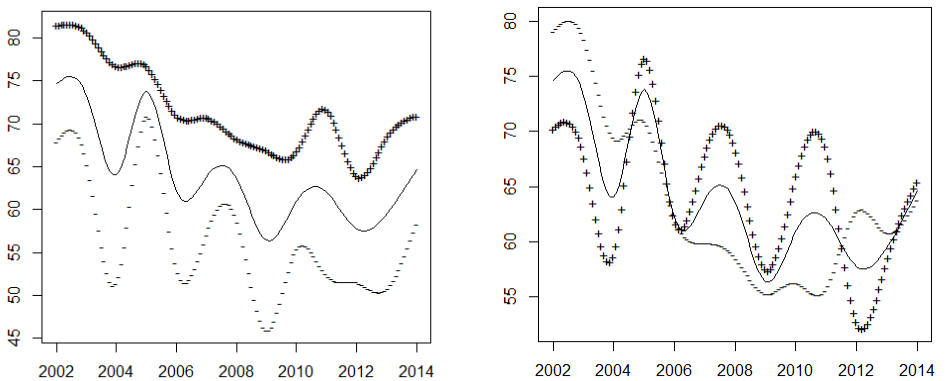

Figure 3. Weight functions

Source: author's compilation using R program.

The first functional principal component explains 72% of the joint variation, whereas the second one 18%. The first component reflects the overall tendency. The plus sign on this component means that the curve describing the result in a particular subject remains above the average. The second component shows the tendency in both outlier and mid years against the average. The plus sign on the second component means that the test result in a given subject before 2005 and after 2011 was below average, whereas in the years 2005-2011 the result was above average.

Based on the results of the functional factor analysis, data visualization and a comparison of analysed objects can be performed. Figure 4 presents data projection on the plane defined by two functional principal components.

Students received the best results in reading – above average. Writing and using information remained on an average level, the situation was better before 2005 and after 2011. Both practical application of knowledge and reasoning were below average.

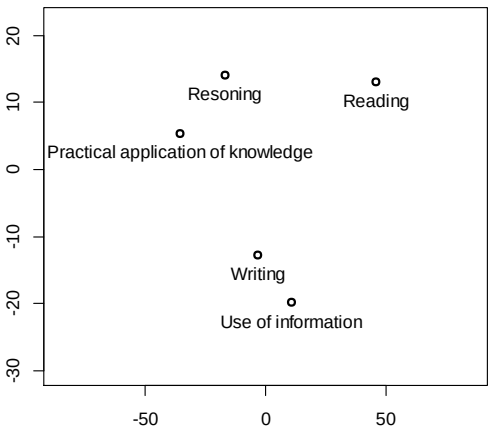

Figure 4. Objects in the components' space
Source: author's compilation using R program.

3.2. The third stage of education

The exam for the third graders of middle school covers the acquired knowledge and skills specified in the core curriculum of general education for the selected subjects taught at the third and earlier stages of education.

Before 2009 middle school students took the exam in humanities as well as maths and natural science only. In the period 2009-2011 middle school graduation exam consisted of three parts:

- humanities,
- maths and natural science,
- modern foreign language.

In the years 2012-2015 the following scopes were identified within the framework of each exam part:

- humanities covering history and social studies and also Polish,
- maths and natural science covering subjects teaching natural sciences and maths,
- modern foreign language at either basic or extended level.

A middle school student takes an exam in one of the following foreign languages: English, French, Spanish, German, Russian, Ukrainian and Italian. A student can choose only the foreign language learned at school as an obligatory subject.

Every middle school student is obliged to take an exam in a modern foreign language at the basic level. An extended level exam is obligatory only for those students who choose to take an exam in the language they used to learn in a primary school. The other middle school students can also take it if they wish to check the level of their language skills.

The exam has a written form. Taking it is the condition to graduate from the middle school, however, the minimum result to be achieved by a student is not defined and, therefore, it is not possible to fail the exam.

Figure 5 presents the average percentage test results in particular subjects taken after the third grade of middle school in the years 2006-2015. The average result in all subjects is denoted by a solid line and shows the decreasing tendency in the period 2006-2015.

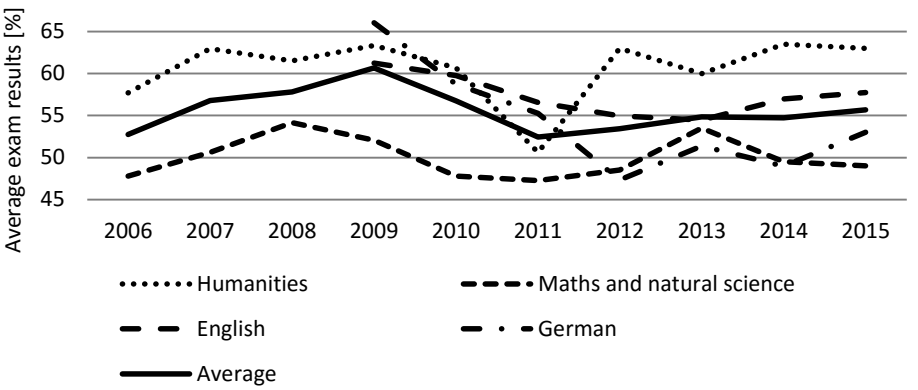

Figure 5. Average exam results achieved on graduating from the third class of middle school in particular subjects [%]

Source: author's compilation based on CKE [Central Examination Commission] data.

Two component functions were distinguished by means of the functional principal component analysis. Fig. 6 presents the graphs of each component deviation from the average in all subjects.

The first weight function (74%) The second weight function (24%)

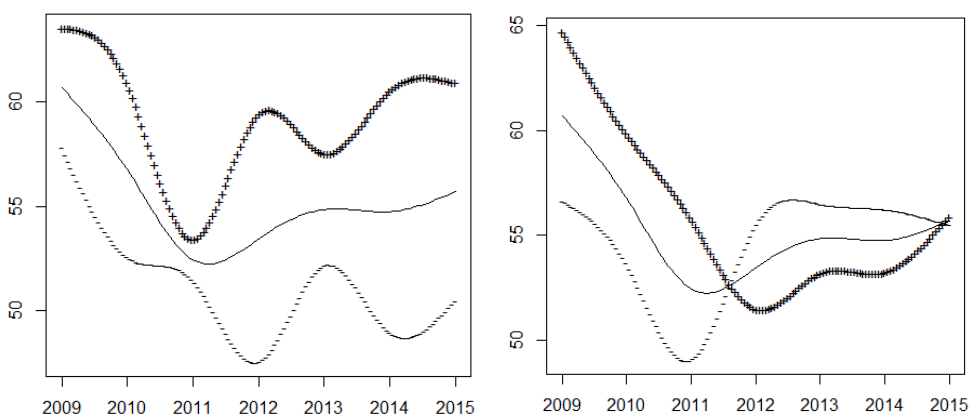

Figure 6. FPCA weight functions

Source: author's compilation using R program.

The first functional principal component explains 74% of the joint variation, whereas the second one 24%. The first component is responsible for the overall tendency. The plus sign of this component means that the curve describing the result in a particular subject remains above the average. The second component shows the tendency in the initial and final years against the average ("the beginning vs. the end") and compares the period until 2011 and after 2012 against the average result. The plus sign on the second component means that the test result in a given subject at the beginning of the analysed period was higher than the average, whereas in the years 2012-2014 the result was below the average.

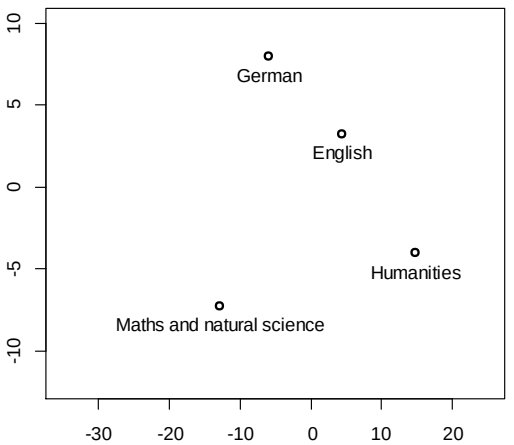

Figure 7. Objects in the components' space
Source: author's compilation using R program.

The visualisation of objects in the space of the component function (fig. 7) allows for drawing the following conclusions:

- humanities: the test result below the average, the situation was worse at the beginning of the analyzed period than in the years 2012-2014,
- English and German: average test result, higher than average at the beginning of the studied period, after 2012 – lower than average,
- maths and natural science: test result lower than average, at the beginning of the analysed period the situation was worse than after 2012.

3.3.The fourth stage of education

A graduate taking "the old type of graduation exam" (before 2015) is obligated to take: two exams in the oral part and three exams in the written part. The obligatory exams in the oral part are as follows:

- an exam in Polish (without defining the level),
- an exam in a modern foreign language (without defining the level).

The obligatory exams in the written part are as follows:

- an exam in Polish (basic level),
- an exam in maths (basic level),
- an exam in a modern foreign language (basic level).

A graduate taking "the new type of graduation exam" (from 2015) in the written part is obliged to take an exam in the chosen additional subject (extended level).

In order to receive the graduation diploma a student has to get at least 30% points at the exam in each obligatory subject in the oral part and receive at least 30% points at the exam in each obligatory subject in the written part.

Figure 8 presents the number of students who passed / failed the secondary school graduation exam in the period 2009-2015. It is noticeable that the number of students taking the graduation exam is decreasing each consecutive year and, moreover, the number of those who failed this exam remains disturbingly high in the recent two years.

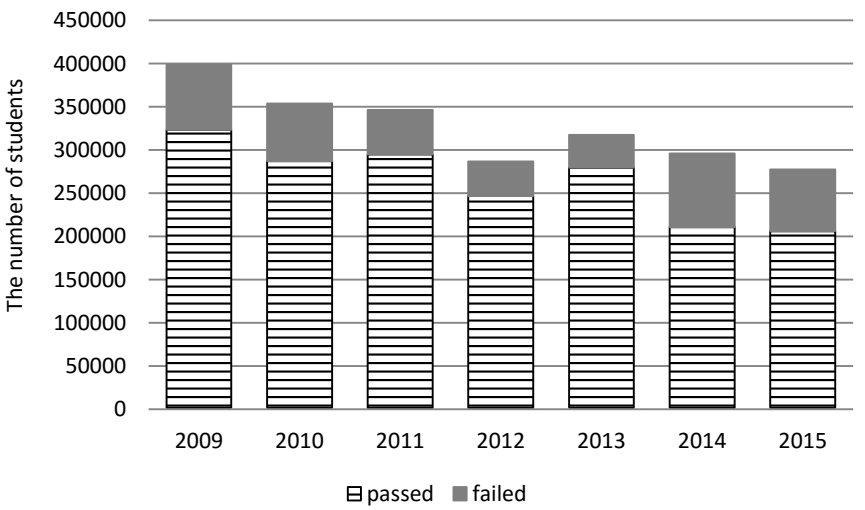

Figure 8. The number of students taking graduation exam

Source: author's compilation based on CKE [Central Examination Commission] data.

Figure 9 presents the average percentage results obtained at the secondary school graduation exam in particular subjects in the period 2008-2015. The bold line shows an average result in all subjects – a decreasing tendency is noticeable.

The functional principal component analysis was applied to distinguish two component functions. The graphs of each component deviation from the average in all subjects are presented in Figure 10.

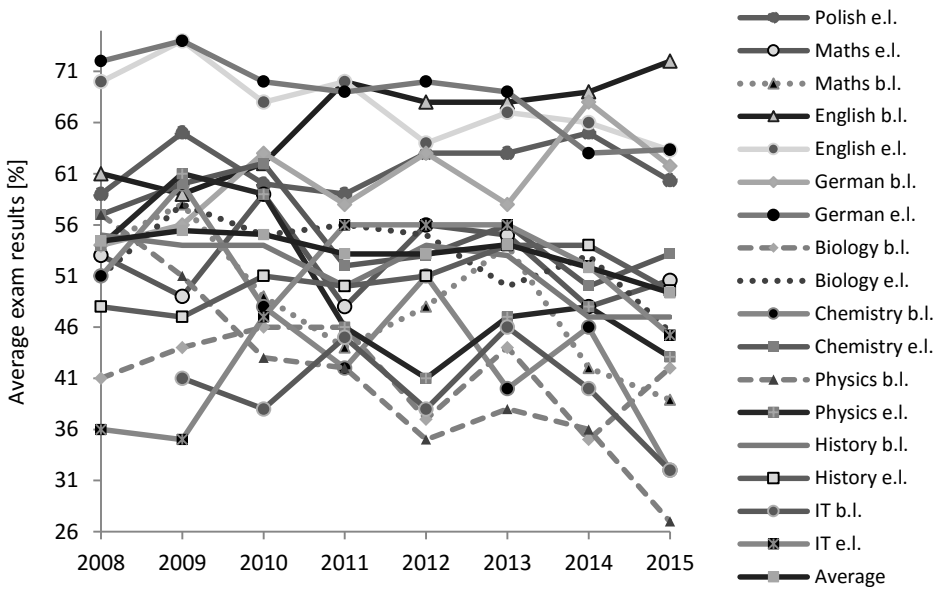

Figure 9. Average results of the secondary school graduation exam

Source: author's compilation based on CKE [Central Examination Commission] data.

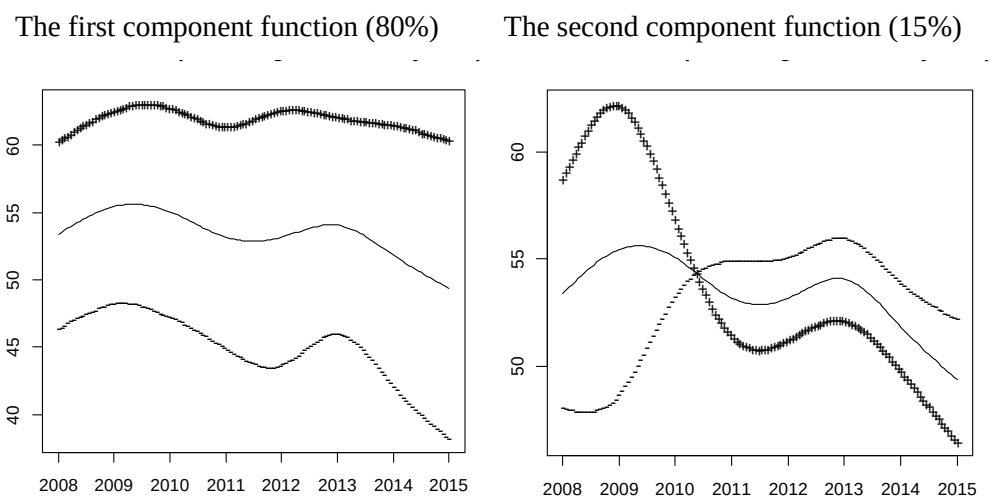

Figure 10. FPCA components

Source: author's compilation using R program.

The first functional principal component explains 80% of the joint variation, whereas the second one 15%. The first component is responsible for the overall tendency. The plus sign of this component means that the curve describing the result in a particular subject remains above the average. The second component

shows the tendency in the initial and final years against the average ("the beginning vs. the end") and compares the period until 2010 and after 2010 against the average result. The plus sign on the second component means that the secondary school graduation exam result in a given subject was higher than the average at the beginning of the analysed period, whereas at the end the result was worse than the average.

Next the data were projected on the plane determined by two functional principal components (Fig. 11).

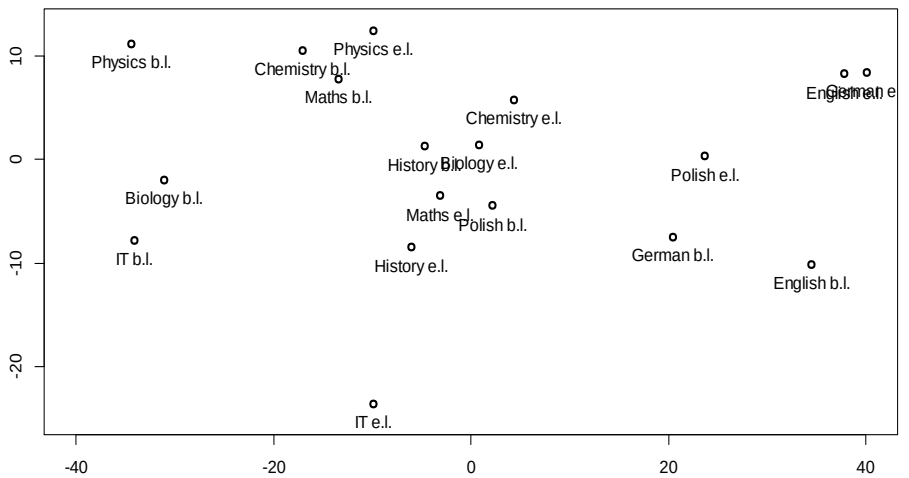

Figure 11. Objects in the component's space

Source: author's compilation using R program.

The best exam results were achieved by the secondary school graduates in German and English at the basic level (plus sign on the first and minus sign on the second component). The worst results were achieved in basic physics, basic chemistry and extended maths (minus sign on the first and plus sign on the second component).

4. Final remarks

The number of sixth graders in primary schools has stabilized since 2014, the number of final year students in middle schools and secondary schools has been continuously decreasing.

The decreasing tendency in average exam results is noticeable at all levels of education. Starting from primary school, the students receive better results in humanities, whereas science causes more problems. Primary school leavers achieved an above average result in reading, but their practical application of knowledge was below average in all subjects. An above average result was obtained in humanities part at the middle school graduation exam, whereas maths and natural science part showed below average results in all subjects. Secondary

school graduates have biggest problems in science, especially in physics (basic and extended level), in basic chemistry and extended maths. The best results were recorded in German and English at the basic level.

Due to the demographic situation and the decrease in the population representing basic educational age groups the number of students has been decreasing since 2006. This situation will persist until 2020, when the students who graduated from primary schools in 2014 will start their university education.

Similarly to the classic analysis of principal components, FPCA makes this visualisation possible. It makes the analysis easier and shows nontrivial correlations, which are difficult to find in a different way. A vital property of both methods is data reduction with maximum information being kept. Thanks to the functional analysis of principal components, which enriches possibilities of the classic analysis of principal components, it is possible to analyse dynamic data, showing both the tendency as well as the pace of changes in time.

REFERENCES

DANIELE, M., (2006). Functional principal components analysis to study environmental data, the article available on the website

http://www.old.sis-statistica.org/files/pdf/atti/Spontanee%202006_677-680.pdf

HALL, P., HOSSEINI-NASAB, M., (2006). On properties of functional principal components analysis, Journal of the Royal Statistical Society, Series B (Statistical Methodology), Vol. 68, No.1, pp. 109–126.

HALL, P., MÜLLER, H. G., WANG, J. L., (2006). Properties of Principal Component Methods for Functional and Longitudinal Data Analysis, The Annals of Statistics, Vol. 34, No. 3., pp. 1493–1517.

HARMAN, H., (1975). Modern Factor Analysis, The University of Chicago Press.

INGRASSIA, S., COSTANZO, G. D., (2005). Functional principal component analysis of financial time series, in: Vichi M., Monari P., Mignani S., Montanari A. (ed.), New Developments in Classification and Data Analysis, Springer-Verlag, Berlin, pp. 351–358.

KRZYŚKO, M., GÓRECKI, T., DERĘGOWSKI, K., (2012). Jądrowa i Funkcjonalna Analiza Składowych Głównych [Nuclear and Functional Analysis of Principal Components], the meeting of Polish Statistical Association Division in Poznań, presentation available on http://stat.gov.pl/cps/rde/xbcr/pts/Krzysko_wyklad_7_11_12.pdf (access date 1st March 2015).

RAMSAY, J. O., SILVERMAN, B. W., (2005). Functional Data Analysis, Springer.

RAMSAY, J. O., HOOKER, G., GRAVES, S., (2009). Functional Data Analysis with R and MATLAB, Springer.