EVOLUTION OF EDUCATION OVER THE YEARS IN EUROPEAN COUNTRIES: AN APPLICATION OF DOUBLE PRINCIPAL COMPONENT ANALYSIS

Catarina Matias Albuquerque

Dissertation

Master in Data Analytics

Supervised by

Adelaide Maria de Sousa Figueiredo

Fernanda Otília de Sousa Figueiredo

2019

Biographical Note

Catarina Matias Albuquerque was born in Tarouca, district of Viseu, on the 2nd of December 1996. In 2014, finished the secondary education in Escola E. B. 2,3/S de Tarouca, she moved to Porto to attend the Bachelor's Degree in Economics at School of Economics and Management of University of Porto (FEP). Three years later, in 2017 and in the same faculty, she enrolled in the Master's Degree in Modelling, Data Analysis and Decision Support Systems.

In 2019, Catarina started her professional activity as Junior Financial & Administrative Assistant in Finance & Legal Department at HUUB.

Acknowledgments

It has been a long and challenging journey during the last two years. I would like to take this opportunity to express my heartfelt gratitude to those who always had supported me. This thesis would not have been possible without you and your encouragement.

First of all, I would like to give my warmest embrace to my parents, Almerinda and Carlos, and my sister, for inspiring me, for all the support given, endless love and for being my best friends day after day. I have no words to express what you mean to me, to thank you for always taking care of me and for giving me the opportunities that have made me who I am.

I also would like to express my gratitude to my supervisors Professors Dr. Adelaide Figueiredo and Dr. Fernanda Figueiredo for all the availability, advices, support, dedication, constructive criticism, and incredible patience during this project, and also the teachers and staff from FEP.

I owe my deepest thanks to my family and my friends from Porto and Tarouca for encouraging me all the time, for cheering me up in the difficult times and for being always by my side. Thank you guys.

Special thanks to my classmates along the five years at FEP, it was a pleasure to share this part of my life with you, for all the amazing times we spent and for all the knowledge we discovered together.

I can't help to mention "Associação Filarmónica de Tarouca" for always understanding my personal challenges and constantly making me feel like home and "Tuna Feminina de Economia do Porto" for giving me the best people that I will always keep in my heart.

I am also thankful to my colleagues from HUUB for all the patience, motivation and for always inspiring me to improve myself and my abilities.

And finally, dear God and Holy Mary, thank you for giving me the encouragement and strength, especially through these challenging times. I am really glad for all your love, guidance, blessings and grace.

Abstract

Education is an important concept that can lead to an increase of the development, employment and growth of the societies. Since this aspect is in constant growing, especially in Developed Countries, this dissertation will study the evolution of the Tertiary Education, that is the highest level of education, in some European countries between 2008 and 2016.

The main purpose of the work is to identify the differences and similarities during the period referred before, to study the evolution trends of the variables (or indicators) and the countries, and to analyse the existence of a common structure, over time, in the data tables.

To perform this analysis will be applied the Double Principal Component Analysis.

Keywords: Education, Principal Component Analysis.

Resumo

Educação é um conceito importante que leva ao aumento do desenvolvimento, emprego e crescimento das sociedades. Estando este aspeto em constante crescimento, especialmente nos Países Desenvolvidos, nesta dissertação vai ser estudada a evolução da Educação Superior, que é considerado o nível mais alto de educação, nalguns países da Europa no período entre 2008 e 2016.

Os principais objetivos da tese são identificar as semelhanças e diferenças no espaço de tempo em estudo, analisar tendências evolutivas dos países e das variáveis (ou indicadores), e analisar a presença de uma estrutura comum, ao longo do tempo, nos quadros de dados.

Para realizar esta análise vai ser aplicado o método de Dupla Análise em Componentes Principais.

Palavras-Chave: Análise em Componentes Principais, Educação.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Motivation

When we hear the word "Education", we all associate, in a simple way, the activities that involve teaching and learning at schools, but education is much more than this. UNESCO, Institute of Statistics, defined education as the "processes by which societies deliberately transmit their accumulated information, knowledge, understanding, attitudes, values, skills, competencies and behaviours across generations. It involves communication designed to bring about learning" (ISCED, 2011). We also can see education as a human right with a great power of evolution and development that can help countries to reduce poverty, improve health, reach gender equality, stability, peace and many other desired aspects.

However, some people and societies have to deal with the lack of opportunities of education. This scenario is more serious in the Least Developed Countries. According to the World Bank Group (2018), there are, approximately, 260 million of children out of school, primary and secondary, and, specifically in Africa, nearly 90% of students cannot read or write, because do not have minimum skills. This is something worrisome, because education has an important impact in the social and economic context of the countries.

In the case of Europe, the panorama is much better than in the Least Developed Countries. According to Statista (2017), the illiteracy rate, since 2015, in adults in Europe has only been 1.9%. All countries in Europe have their own education system, but these systems could be inserted in a statistical framework of UNESCO, that is ISCED (International Standard Classification of Education).

In terms of levels of education, it is common sense that a higher level of education provides more knowledge, growth and development. According to UNESCO, "Tertiary Education includes what is commonly understood as academic education but also includes advanced vocational or professional education." (ISCED, 2011). Tertiary Education, the higher level of education, implies learn with a high level of specialisation and complexity which includes diplomas, certificates, the degree of bachelor, master, doctoral and others. In general, for one year extra of education, there is an increase of 9% in earning per hour.

Actually, the systems of High Education are in constant evolution through government policies and development of schools for example. That is happening, because there is a gradual awareness in Europe that the countries must invest more and better in Higher Education.

As we can see in the Figure 1.1, the percentage of people with Tertiary Education has being increase significantly since 2000 and Europe & Central Asia are the second group with the highest percentage.

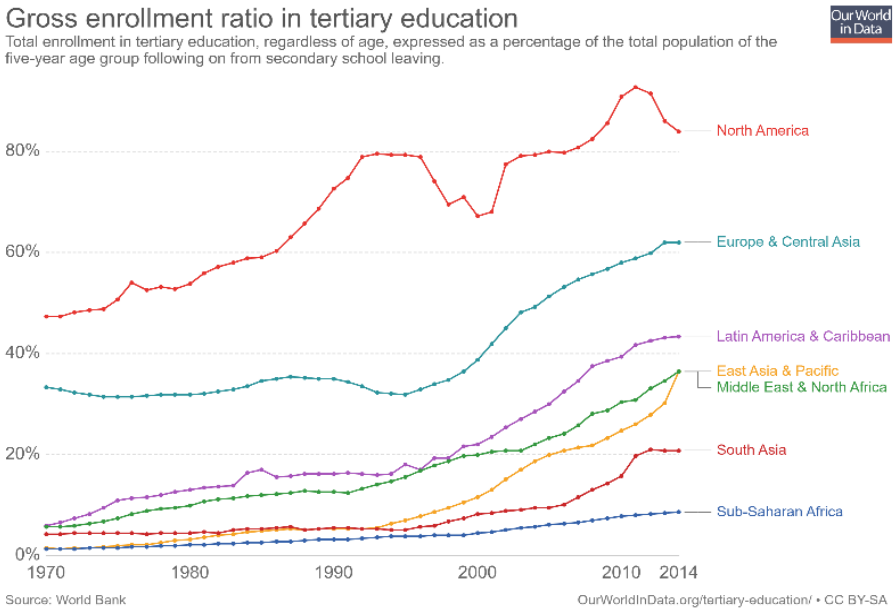

Figure 1.1 - Evolution of Tertiary Education **(Roser & Ortiz-Ospina, 2018)**

Since education, in special Tertiary Education, is such a relevant topic that can promote poverty reduction, employment, health for the individual and can drive societies to economic growth, constantly innovation, development of the human capital and so many other advantages, this work will study the evolution of Tertiary Education in European countries over the years.

1.2 Contextualization in Statistics

Multivariate analysis is a broad field where even very experienced statisticians move carefully, because it is a recent area of science and much has already been revealed, but much is yet to be discovered.

In Statistics, the data tables are very hard to analyse most of time, because they are composed of a lot of information and they have correlations between them. When there

are many data tables, the procedure more intuitive to apply is treating, separately, each data table, which is obviously expensive, a waste of time and there is loss of joint information. To analyse this type of data there are different methods, with specific properties and restrictions, and the ways to perform analysis depends on the goals to be achieved.

The Three-way data tables are data structures composed by three dimensions (where one or more can be the time), n individuals, with $i = 1,2,\ldots,n$, p variables, with $j = 1,2,\ldots,p$ in t time, with $t = 1,2,\ldots,T$, and must contains two dimensions in common (columns and rows).

Tensors can be seen as a high dimensional generalization of matrices. A first-order tensor is a vector, a second-order tensor is a matrix, and tensors of order three or higher are called higher-order tensors (Kolda & Bader, 2009).

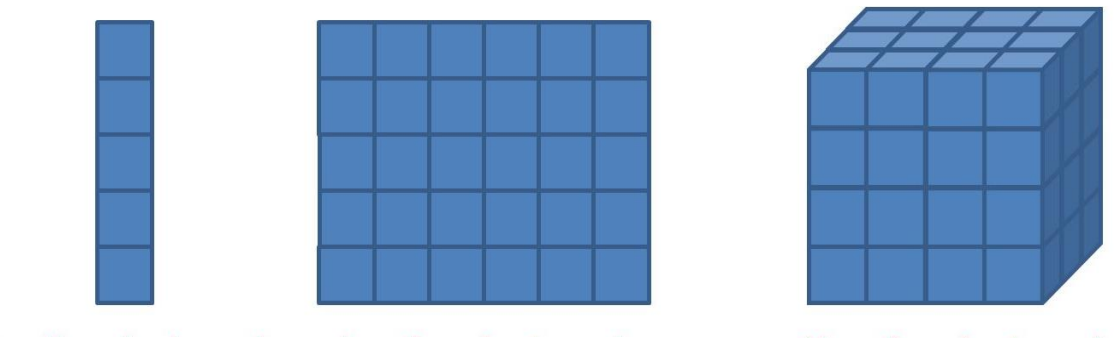

Figure 1.2 - Data Structures - Adapted Figure **(Big Data Analysis - What is a Tensor, 2017)**

"A tensor represents a high dimensional dataset like for example three or more dimensional datasets that are often encoded as multi-dimensional arrays." (Big Data Analysis - What is a Tensor, 2017).

Tensors can be quite useful with many applications such as in the case of text mining, that leads to models where the co-occurrence of 3 or more words are incorporated, in social networks, the arbitrary degrees of influence can be encoded by tensors, and so on.

1.3 Objectives

In this dissertation are presented some methods of multivariate analysis with more detail the methods based in Principal Component Analysis (PCA). PCA is a dimension-

reduction tool used to transform a set of variables into a smaller set that contains the majority of the information of the large set and other concepts related to panel data and time series. The method applied in the data set is the Double Principal Component Analysis (DPCA), that is studied deeply further in a specific chapter.

The main goals of this study are comparing globally the data tables and represent them in principal plans, finding the common structure in the data tables, analysing the evolution, the trends, making the trajectories and identifying the similarities and differences between the individuals along the time.

1.4 Thesis Structure

The thesis is composed by 5 chapters, with Bibliography and the Annexes in the end. The structure is the following:

The chapter one, that is the Introductory section, is divided in the motivation of the topic (why education and Tertiary Education are such an important topic in European countries), the contextualization in statistics (some relevant statistic information about the topic), the objectives of the dissertation (what I want with this work) and the thesis structure.

The second chapter is the State of the Art that contains some basic definitions, a study of the Factorial Analysis and the different methods to analyse data frames and joint data frames. The concepts related to cross-section data, time-series and panel data and some important aspects about multivariate data also take place in the State of the Art of this dissertation.

The following chapter, the third one, is all about Double Principal Component Analysis, because is the method applied in the database.

In the fourth chapter the database, the results and discussions obtained from the statistical analysis and the application of the Double Principal Component Analysis in the database are presented.

The last chapter of the dissertation contains the conclusions and future work.

To finish, are presented the Bibliography and Annexes that were necessary to perform this thesis.

2. State of the Art

Multivariate statistics consists in a set of methods that deals with multivariate data, that is when data tables contain more than three variables for each observation. Fundación BBVA (2013) says that the main aim of the multivariate analysis is to analyse how variables are related to one another, and how they work together to differentiate the cases in which the observations are made.

The method that should be applied in a data set depends on the type of the data and on the objective of the research. That shows why is so important studying and knowing the different methods.

The section 2.1 is about Factorial Analysis and the respective methods and in section 2.2 are presented the concepts related to cross-section data, time-series, panel data and some important aspects about multivariate data.

2.1 Factorial Analysis

Factorial Analysis is a set of techniques and its main purpose is the reduction of a huge number of variables into a smaller one with the minimal loss of information and it is used to simplify data, because of the reduction of dimensionality. There are two types of Factorial Analysis: the Exploratory Factorial Analysis, which is used to explore the data in an initial phase of the search, to find the relation between the variables and to identify patterns of correlations, and the Confirmatory Factorial Analysis, which is used to test hypotheses.

In the next sections are presented the different types of Factorial Analysis for a single data table and for several data tables.

Data	Properties		Method
One Data Table	Quantitative Variables		Principal Component Analysis (PCA)
	Qualitative Variables	2 Variables/Contingency Tables	Factorial Correspondence Analysis (FCA)
		>2 Categorical Variables	Multiple Correspondence Analysis (MCA)
Multivariate Data (Three-way Data)	Quantitative Variables		STATIS
	Quantitative or Qualitative Variables or Both		Multiple Factorial Analysis (MFA)
	All the tables with the same variables and observations		Double Principal Component Analysis (DPCA)

Table 2.1 - Methods of Factorial Analysis

2.1.1 Factorial Methods for Analysis of a Data Table

2.1.1.1 Principal Component Analysis

Principal Component Analysis, a technique of multivariate analysis, is probably the oldest and best known of the multivariate techniques and it is fundamental to the study of multivariate data. This method came up in 1901 with the contributions of Karl Pearson and was developed by Hotelling in 1993 independently (Jolliffe, 1986). Like other methods, PCA was not immediately and wisely used until the technology and the computers were developed. Actually, this and the other methods are well represented in every statistical program with adequate packages.

Although this technique is one of the oldest multivariate techniques, it continues to be target of many researches, ranging from new approaches to ideas of algorithms from neural networks. PCA is tremendously versatile and has applications in numerous disciplines (Jolliffe, 2002).

Principal Component Analysis is applied in data tables that contain n individuals and p numerical variables. The principal idea of PCA is to reduce the dimensionality of data (dimension-reduction) with the minimal loss of information, preserving the maximum variation present in the data. This transformation aims to reduce a large volume of correlated variables to a smaller number of variables non-correlated, retaining the greatest part of the variability of the original data.

The new variables created are named principal components, they are ordered and uncorrelated, and are, also, a linear combination of the original variables. The principal components are ordered, which means the first principal component explains more proportion of data variability than the other ones, so it is more important statistically. The number of the principal components should be lower or equal to the original variables number.

The principal components should have associated high eigenvalues. The eigenvalue of one factor measures the amount of variance in all initial variables that is described by the factor. Therefore, a higher eigenvalue indicates that a large amount of variance in the data is explained by the principal component. Basically, the objective of the analysis is that only the highest eigenvalues should be retained. The eigenvalues and their eigenvectors can be calculated from its correlation or covariance matrix. The eigenvectors are standardized, and

the principal components are linear combinations of the variables. The coefficients of the linear combinations are the eigenvectors of the correlation or covariance matrix.

Some authors, Adbi & Williams (2010), said that "Principal component analysis is a multivariate technique that analyses a data table in which observations are described by several inter-correlated quantitative dependent variables. Its goal is to extract the important information from the table, to represent it as a set of new orthogonal variables called principal components, and to display the pattern of similarity of the observations and of the variables as points in maps.".

This method allows the reduction of dimensionality, because, although the statistical information present in the p original variables is the same that in the p principal components, PCA can obtain more than 80% of this information in only 2 or 3 principal components.

The Principal Component Analysis is the technique that supports the method that is presented in the next chapter and that was applied in the data base.

2.1.1.2 Factorial Correspondence Analysis

Factorial Correspondence Analysis (FCA) is an exploratory analysis method mainly projected by the French Jean-Paul Benzèrci in 1969, posteriorly developed by English researchers.

This method is a statistical technique with two main particularities. FCA can only be applied when the variables are categorical, and it is conventionally applied in contingency tables (2 variables).

The main objective of Correspondence Analysis is generating a bi-dimensional representation of the data to check the associations between the categorical variables (the columns and the rows). "It allows the construction of an orthogonal system of axes (called factors) where observations (rows of the table) and variables (columns of the table) can be simultaneously displayed, making easy to discover the salient information included in a given contingency table" (Soares, 1999).

In FCA, when variables and observations are proximal this is understood as a strong relationship and a strong similarity. The interpretation in Correspondence Analysis

of an axis is defined by opposition in the extreme points. These points are often the ones that have the larger contribution to the inertia of the axis.

To apply this method, all the data should be on the same scale and it provides a summarization of the data in a two-dimensional graphical. One of the results of FCA is the simultaneous presentation of the columns and the rows on a lower dimensional vector space, that is typically a two-dimensional representation.

This methodology is principally applied in the area of text mining (Balbi, Stawinoga, & Triunfo, 2012).

Since Factorial Component Analysis is only applied when the variables are categorical and in contingency tables, that means, data with only 2 variables, this method cannot be applied in the analysis of the data base in this dissertation.

2.1.1.3 Multiple Correspondence Analysis

Multiple Correspondence Analysis (MCA), a statistical technique, came up also with a work of the French Jean-Paul Benzèrci in 1973 that was improved by other statistics.

This method can be considered as an extension of the previous method, the Factorial Correspondence Analysis, when there are more than 2 categorical variables, and can be seen as the equivalent of Principal Component Analysis, in this case, to categorical variables instead of quantitative. Clarifying, MCA is a technique used to explore, summarize, describe and visualize information from a data table with n individuals and q categorical variables, where q is higher than 2 and should be lower than n.

The main goals of the method are studying and identifying individuals with similarities in a multidimensional perspective, analysing the associations between the variable categories and connecting the study of the variables and the individuals to define the individuals through the variables.

With MCA it is possible to analyse the pattern of the relations of many categorical variables. To apply this technique there are differences from the previous method. To perform MCA is previously necessary to code each categorical variable as a binary variable, using the number "1" when the level of each variable is present, and number "0" when it is not. Finishing this transformation, the data is represented by a table with binary columns,

which means the possibility of existence of only one column with the number "1" for each variable.

The Multiple Correspondence Analysis is frequently used to investigate questionnaire data.

2.1.2 Factorial Methods for Joint Analysis of Data Table

2.1.2.1 STATIS

The STATIS method, "Structuration des Tableaux À Trois Indices de la Statistique", was introduced by Y. Escoufier in 1973 and H. l' Hermier des Plantes in 1976. It was later developed by Lavit in 1988 and can also be seen as an extension of Principal Component Analysis as a way to perform three-way exploratory analysis that allows the exploration of many data tables simultaneously.

This method can be applied in quantitative data in one of these two situations:

- Several data tables collected on different occasions, on the same individuals, and the variables can be different along the data tables;
- Several data tables collected on different occasions, with the same variables, and the individuals can differ along the data tables.

Each one of these two situations require a different strategy. In the first case, more importance should be attributed to the proximity between the individuals and study if the distances among the individuals are stables over the time (STATIS method); in the second case, it should be privileged the relations between the variables and study the stability of the correlations among the variables (Dual STATIS method). When the tables have all the same individuals and the same variables the two strategies can be applied (STATIS and Dual STATIS). So, to perform the methodology STATIS, it is not required that the number of columns should be equal for all the data tables.

This methodology has as objectives: to find a common structure for the data tables, to analyse the differences and similarities between the tables and to represent the trajectories of the variables and the individuals.

The goals defined can be represented in 3 phases. There are some ways to classify the phases of the method STATIS. The following phases are inspired by the formulations presented by Escoufier (1978) and Lavit (1988) (Ambapour, 2001):

- **Interstructure:** corresponds to the global analysis of the data tables to compare them and determine the similarities or the differences between the data tables, based on the distance between two objects. If the distances were not too high, a compromise can be made;
- **Compromise:** the objective is summarizing the data tables in only one, the compromise table, according to certain criteria. The compromise should have the same nature of the objects to better represent the individuals. Through the compromise, it is possible to represent the positions of the individuals and the interpretation of the compromise axes to explain the similarities or differences of the interstructure;
- **Intrastructure:** this phase allows the representation of the positions compromise of the individuals that result in the compromise Euclidean image. In the same way as PCA, the axes of the principal plan of the compromise are also interpreted through the correlations of the variables with the axes of the compromise. There is also the representation of the trajectories of individuals, that describe the evolution of each individual or variable through the image of compromise.

The method now presented and the Dual STATIS could be methods applied in the database of this dissertation, because the variables are quantitative, and the number of variables and individuals are the same for all data tables.

2.1.2.2 Multiple Factorial Analysis

Multiple Factorial Analysis (MFA), introduced by Escofier and Pagès in 1985, can be seen as an extension of Principal Component Analysis, when variables are quantitative and of Multiple Correspondence Analysis, when variables are qualitative, in this case in the analysis of several data tables simultaneously. So, MFA can be applied when the data tables contain qualitative or quantitative variables or even both and allows studying the relations between variables, observations and tables.

This methodology is focused on studying tables, when a set of variables that are structured in groups describe a certain group of individuals. Each group can contain a different number of variables, but within the same group, the type of the variables (qualitative or quantitative), should be the same (Kassambara, 2017). The main goal is to integrate diverse sets of variables describing the same observations.

MFA consists in two steps. The first one is to perform a PCA, MCA or CA in each group of variables, according to the type of the variables in the group. The focus on this phase is the achievement of the first eigenvalue of each analysis. The distribution of the eigenvalues provides important information about the dimensionality. In the next step, a unique matrix is formed by a fusion of the normalized data groups and a PCA is applied on this matrix. In the application of the PCA, each table is weighted by the inverse of the first eigenvalue, obtained in the first step, to balance the group's influence. With these procedures, it is possible to study and analyse the evolution of the different individuals and variables.

Economy, ecology, chemistry, sensory evaluation are some of the domains in which MFA is used, but it can also be very useful in analysing surveys when it is possible to identify groups of variables or, for example, when are asked the same questions in different periods of time.

2.1.2.3 Double Principal Component Analysis

Double Principal Component Analysis (DPCA) was introduced by Bouroche in 1975 and it is also an extension of Principal Component Analysis that is applied in three-way data with quantitative variables. However, there are adaptations of the DPCA, that were "developed for the analysis of three-dimensional quantitative data, to categorical data by means of the chi-squared distance between the rows and columns profile of a contingency table" (Lera, Pérez, & Boquet, 2006).

Moreover, a DPCA application was made for categorical data in a study of internal migration in Cuba (Lera, Pérez, & Boquet, 2006). It is well known that many social, economic and medical analyses, etc., analyse qualitative longitudinal variables, which shows the importance of adapting methods that can be used for this type of data.

This method has a particularity, because it can only be applied when all the data tables, collected on different occasions, have the same individuals and the same variables. The variables should have the same units of measure. In the cases where the third dimension is not the time, the analysis is still possible, but the interpretation is extremely harder.

DPCA is more restrict when comparing to the STATIS and the Multiple Factorial Analysis. Still, this will be the method applied to the data base in this dissertation.

2.2 Multivariate Time Series and Panel Data

In this section, it is important to differentiate cross-sectional data, time-series data and panel data.

Cross-sectional data is a type of data that is collected at a single point of the time (no time dimension) for many individuals. In this case, the order of data is not important. The analysis of this data consists, mostly, on the differences between the individuals/subjects. Cross-sectional data contrasts with time-series data since the data is collected at multiple time periods and the order of data is important. The observations are repeated for the same variables and, usually, they are not independent over time. Typically, time series are sequences collected at successive equally spaced moments in time.

Multivariate time series, in a simple way, are multiple variables that change over time. It happens when there is more than one variable that is time-dependent. Each one of the variables depends on its past values and also has a dependency on the other variables. The existence of this dependency is what is used for forecasting the future values.

Kotu & Deshpande (2019) said that "Multivariate time series may use many time-concurred value series for prediction. If one has a series of points spaced over time, conventional forecasting uses smoothing and averaging to predict where the next few points will likely be." One of the most used methods for multivariate time series forecasting is Vector Auto Regression (VAR) that is used to find linear interdependencies between multiple time series. In this model, each one of the variables is a linear function of its own past values and the past values of the other model variables. One requirement of VAR is that the variables should, hypothesized, affect by each other intertemporally.

"A longitudinal, or panel, data set is one that follows a given sample of individuals over time, and thus provides multiple observations on each individual in the sample." (Hsiao, 2003). In other words, this type of data studies, simultaneously, the variations of the variables over time and between different individuals, and allows a quantitative analysis of economics relations by combining time-series and cross-section in the same model (pooling). There are balanced panels, all observations are observed in every periods of time, and unbalanced panels, the observations are not observed in every periods.

A panel has the following form:

$$X_{it}, \text{ with } i = 1,2,\dots,N \text{ and } t = 1,2,\dots,T$$

where i is the dimension of individuals and t the dimension of time. In general, a panel data regression model is

$$y_{it} = \alpha + \beta X_{it} + u_{it},$$

where u_{it} is an individual-specific that does not vary with time, in other words, is fixed over time. Two significant models are the random effects model and the fixed effects model.

Contrarily to panel data, where data only has two dimensions (cross and time sections), multidimensional panel data is when data has three or more dimensions. A multidimensional panel with four dimensions, for example, usually has this form:

$$X_{isth}, \text{ with } i = 1,2,\dots,N, s = 1,2,\dots,S, t = 1,2,\dots,T, h = 1,2,\dots,H,$$

where i is the dimension of individual, s the dimension of series, t the dimension of time and h is the dimension of horizon.

In general, the multidimensional panel data regression model, an econometric model designed for panel data, could be written as:

$$y_{isth} = \alpha + X_{sith}\beta + u_{sith}.$$

In terms of economic research, panel data have some more advantages than cross-sectional and time-series data sets, because allow the researcher to analyse a lot of important economic questions that the other ones cannot. "The oft-touted power of panel data derives from their theoretical ability to isolate effects of specific actions, treatments, or more general policies. This theoretical ability is based on the assumption that economic data are generated from controlled experiments in which the outcomes are random variables with a probability distribution that is a smooth function of the various variables describing the conditions of the experiment." (Hsiao, 2003). This is one of the reasons why this type data are in constant evolution and have always more and more importance.

Panel data has several advantages, one of them is the relief of individual heterogeneity. Thus, panel data suggest the existence of differentiating characteristics of

individuals, understood as "basic statistical unit". Temporal or sectional studies that do not consider such heterogeneity will, most of times, obtain strongly skewed results. This type of data, with the inclusion of the sectional dimension, in aggregated time-series, confers larger variability to the data. This increase in data variability contributes to the reduction of eventual collinearity between variables, particularly in models with distributed mismatches. In addition, studies with longitudinal samples provide more efficient analysis of the adjustment dynamics than the sectional studies that do not consider the possibility of a dynamic reality, giving a fake idea of stability. So, the use of panel data allows the combination of diversity of individual behaviours and the existence of adjustment dynamics to typify the answers of different individuals to certain events at different times. On the other hand, the efficiency of the estimation increases due to the fact that there are more amount of information available and, it is possible to identify and measure effects that would not be possible to detect with exclusively sectional or temporal studies, as well as to construct and test complex behavioural models by the use of distributed mismatches with few restrictions. Summing up, panel data provide more amount of information, larger variability of data, lower collinearity between variables, higher number of freedom degrees and higher efficiency in estimation of the econometric models.

Although, econometric analysis with panel data is not immune to problems, because it increases the risk of having incomplete samples or data with serious collection problems, as well as the importance of measurement errors; occurs the heterogeneity bias (bias resulting from a poor specification due to the not consideration of a possible differentiation of the coefficients along the sectional units and/or over time); problems related to selectivity bias, i.e. errors resulting from the collection of data, which means that they do not constitute a random samples, and other disadvantages.

However, the method selected to apply in this dissertation was Double Principal Component Analysis and the approach of this statistical technique is detailed in the next chapter.

3. Double Principal Component Analysis

3.1 Introduction

The main objectives of Double Principal Component Analysis are globally comparing the different data tables, studying the evolution of the relations between the different variables and the evolution of the individuals as in STATIS and Multiple Factorial Analysis methodologies.

3.2 Method and Notations

This method can be divided in three different phases:

- **1st Step: Interstructure** – Analysis of the global evolution over time. It is based in a PCA of the centres of gravity in the clouds;
- **2nd Step**: **Analysis of the clouds of individuals –** Study of the deformation of the clouds around its centre of gravity. The analysis of the clouds, centred relative to its centre of gravity is made to eliminate the phenomenon of global evolution;
- **3rd Step**: **Intrastructure** – Representation, in the same space, of the evolution of the different individuals and variables over time, to synthetize the similarities and the differences between the clouds. The main problem is finding a space in which the trajectories of individuals and variables can be represented.

To better explain the method, there is a need of explaining first, the notation that will be used.

Initially, there are T data tables of order $n \times p$, X_t with $t = 1, 2, \ldots, T$, composed of n individuals measured by p variables in T different instants. The entries in each one of the tables are denoted by $(x_i^j)^{(t)}$, which represent the value of the individual i for the variable j, with $i = 1,2, \ldots, n$ and $j = 1,2, \ldots, p$ in time t, with $t = 1,2, \ldots, T$, and X_t is the matrix in the instant t:

$$X_t = \begin{bmatrix} (x_1^1)^{(t)} & (x_1^2)^{(t)} & \cdots & (x_1^p)^{(t)} \\ (x_2^1)^{(t)} & (x_2^2)^{(t)} & \cdots & (x_2^p)^{(t)} \\ \vdots & \vdots & \ddots & \vdots \\ (x_n^1)^{(t)} & (x_n^2)^{(t)} & \ldots & (x_n^p)^{(t)} \end{bmatrix}$$

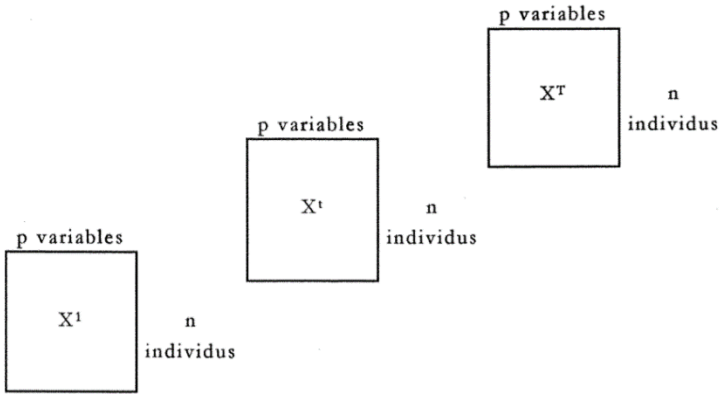

Figure 3.1 - T Data Tables over time (Guirao, 1998)

At time t, the $j\ -th$ variable is denoted by the vector of $\mathbb{R}^n$ and given by:

$$\left(\boldsymbol{x}^j\right)^{(t)} = \begin{bmatrix} \left(x_1^j\right)^{(t)} \\ \left(x_2^j\right)^{(t)} \\ \vdots \\ \left(x_n^j\right)^{(t)} \end{bmatrix}, \text{with } j = 1,2,\dots,p$$

and the $i\ -th$ individual is denoted by the vector of $\mathbb{R}^p$ and given by:

$$(\boldsymbol{e}_i)^{(t)} = \begin{bmatrix} (x_i^1)^{(t)} & (x_i^2)^{(t)} & \dots & \left(x_i^p\right)^{(t)} \end{bmatrix}, \text{with } i = 1,2,\dots,n.$$

Individuals are weighted and this matrix with the weights of the individuals is defined by:

$$D = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_n \end{bmatrix}, \text{with } 0 < p_i < 1 \text{ and } \sum_{i=1}^{n} p_i = 1.$$

Let Q be the diagonal matrix of the inverse of the variables' variances defined by:

$$Q_{1/s^2} = \begin{bmatrix} 1/s_1^2 & 0 & \cdots & 0 \\ 0 & 1/s_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/s_p^2 \end{bmatrix},$$

where $s_1^2, s_2^2, \dots, s_p^2$ are the variances of the p variables.

At time t, the centre of gravity of table X_t is the vector of weighted means of the p variables defined by:

$$\boldsymbol{g}^{(t)} = \begin{bmatrix} (\bar{x}^1)^{(t)} \\ (\bar{x}^2)^{(t)} \\ \vdots \\ (\bar{x}^p)^{(t)} \end{bmatrix}, \text{ where } (\bar{x}^j)^{(t)} = \sum_{i=1}^{n} p_i (x_i^j)^{(t)}, j = 1,2,\dots,p \text{ and } t = 1,2,\dots,T$$

For each instant t, there is a cloud of individuals defined by table X_t, which is denoted by:

$$N_I{}^{(t)} = \left\{ \boldsymbol{e}_i{}^{(t)}, i = 1,2,\dots,n \right\}, \text{ with } t = 1,2,\dots,T.$$

In the sections below are explained with more detail the three phases of the Double Principal Component Analysis.

3.3 Interstructure (1st Step)

The Interstructure is represented by the analysis of the global evolution over time. The main purpose of the first phase is analysing and describing the global evolution of the individuals as well as studying the global evolution of the data tables through its centres of gravity of each cloud $N_I{}^{(t)}$.

The centres of gravity of the clouds $N_I{}^{(t)}$ are defined by a matrix G with dimension $T \times p$ denoted by:

$$G = \begin{bmatrix} \boldsymbol{g}^{(1)\prime} \\ \boldsymbol{g}^{(2)\prime} \\ \vdots \\ \boldsymbol{g}^{(T)\prime} \end{bmatrix} = \begin{bmatrix} (\bar{x}^1)^{(1)} & (\bar{x}^2)^{(1)} & \cdots & (\bar{x}^p)^{(1)} \\ (\bar{x}^1)^{(2)} & (\bar{x}^2)^{(2)} & \cdots & (\bar{x}^p)^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ (\bar{x}^1)^{(T)} & (\bar{x}^2)^{(T)} & \dots & (\bar{x}^p)^{(T)} \end{bmatrix}, \text{ with } t = 1,2,\dots,T.$$

The interstructure phase is based on performing a Principal Component Analysis on the cloud defined by the centres of gravity of the matrix G.

Through this procedure, it is possible to achieve an Euclidean image of the tables in a desired dimensional space. The first axis of this Euclidean image is explained by the global evolution of the time, and, in general, the centres of gravity ($\boldsymbol{g}^{(t)}$) vary continuously along this axis and over time.

3.4 Analysis of the Clouds of Individuals (2nd Step)

As presented in the section 3.2, this phase consists in studying the deformation of the clouds around its centre of gravity. A Principal Component Analysis is performed on each of the T data tables centred on its centres of gravity. These PCAs that were applied allow the interpretation of the results of each one of them with the help of graphical representations and interpretations of the individuals and variables. The higher is the number of matrices T, more difficult and time consuming is the analysis.

This procedure has as main objective the elimination of the phenomenon of global evolution in time that was already observed and analysed. The centred table is defined by:

$$Y_t = X_t - \mathbf{1}_n \boldsymbol{g}^{(t)\prime} = (I_n - \mathbf{1}_n \mathbf{1}_n' D) X_t,$$

where I_n is the identity matrix of order $n \times n$, $\mathbf{1}_n$ is the vector of $\mathbb{R}^n$ with all the components equal to 1, of dimension $n \times n$, and the general term is

$$(y_i^j)^{(t)} = (x_i^j)^{(t)} - (\bar{x}_i^j)^{(t)}, \text{ with } t = 1,2,\dots,T.$$

The T PCAs performed provide, each one of them, two orthogonal axes systems and allow the determination of T systems of principal axes in each cloud ${N_I}^{(t)}$, with $t = 1,2,\dots,T$. Being V_t the matrix of variance and covariance of X_t, W_t the matrix of scalar products between individuals and q the number of axes retained in each one of the PCAs and $q < \min(p,n)$,

- T systems of principal factors, $(\boldsymbol{u}_l)_{l=1,2,\dots,q}^{(t)}$, vectors with size p, that are the eigenvectors of QV_t, associated to the q largest eigenvalues $\lambda_1^{(t)}, \lambda_2^{(t)}, \dots, \lambda_q^{(t)}$, with $t = 1,2,\dots,T$;
- T systems of principal components, $({\boldsymbol{F}_l}^{(t)})_{l=1,2,\dots,q}$, vectors with size n, that are the eigenvectors of $W_t D$, associated to the q largest eigenvalues $\lambda_1^{(t)}, \lambda_2^{(t)}, \dots, \lambda_q^{(t)}$ with $t = 1,2,\dots,T$.

These axes systems will serve to search for a common space to the representation of individuals.

3.5 Intrastructure (3rd Step)

The last phase is the Intrastructure that answers to the principal objective of the method, that is, the representation of individuals in a common space over time.

The notation to the two axes systems desired is $(\boldsymbol{v}_l)_{l=1,2,\dots,q}$, to the principal factors and $(\varsigma_l)_{l=1,2,\dots,q}$, to the principal components.

The trajectories of each individual over time are obtained by the projection of each one of them in the new axes system.

With $QV_t\boldsymbol{u}_l^{(t)} = \lambda_l^{(t)}\boldsymbol{u}_l^{(t)}$, at instant t, the inertia explained by the principal factor l is defined by:

$$\langle \boldsymbol{u}_l^{(t)}, \boldsymbol{u}_l^{(t)} \rangle_{V_t} = (\boldsymbol{u}_l^{(t)})' V_t \boldsymbol{u}_l^{(t)} = \lambda_l^{(t)}, t = 1,2, \dots, T.$$

The index $\Phi(t,\tau)$, that represents the percentage of inertia loss of the cloud $N_I{}^{(t)}$of individuals defined by table t when these individuals are projected on the first q principal factors of cloud $N_I{}^{(\tau)}$ instead of cloud $N_I{}^{(t)}$, is defined by:

$$\Phi(t,\tau) = \left(\sum_{l=1}^{q} \lambda_l^t - \sum_{l=1}^{q} (\boldsymbol{u}_l^{(\tau)})' V_t \boldsymbol{u}_l^{(\tau)} \right) \Big/ \sum_{l=1}^{q} \lambda_l^t$$

In other words, when the cloud $N_I{}^{(t)}$ is projected in the subspace formed by the vectors $(\boldsymbol{u}_l^{(\tau)})_{l=1,2,\dots,q}$, the percentage of loss of inertia explained by the cloud $N_I{}^{(t)}$ decrease to $\Phi(t,\tau)$.

It is important to be aware that $\Phi(t,\tau) \neq \Phi(\tau,t)$ and $\Phi(t,t) = 0$.

By performing the projection of the T clouds $N_I{}^{(1)}, N_I{}^{(2)}, \dots, N_I{}^{(T)}$on the subspace formed by $(\boldsymbol{u}_l^{(\tau)})_{l=1,2,\dots,q}$, the average loss of inertia is

$$\Phi(.,\tau) = \frac{1}{T} \sum_{t=1}^{T} \Phi(t,\tau)$$

3.5.1 Criteria for Choosing the Best Axes System

According to Bouroche (1975), there are 4 criteria to the selection of the better axes system. These criteria are presented in the next subsections.

3.5.1.1 Minimization of the Sum of the Relative Losses of Inertia

The first axis selection criterion consists in choosing among T axes systems found, where the sum of the inertia losses, when all the clouds are projected on the axis system is minimum.

The criterion is basically about selecting the axes system $(\boldsymbol{u}_l^{(\tau)})_{l=1,2,\dots,q}$ where:

$$\Phi(.,\tau) = \min_{t=1,2,\dots,T} \Phi(.,t),$$

that means, in other words, that the percentage of inertia losses should be minimal where the individuals of each one of the clouds $N^{(t)}$, $t = 1,2,\dots,T$ are projected in the subspace formed by this axes system.

According to the first criterion, the systems in which the trajectories of individuals are represented are $\boldsymbol{v}_l = \boldsymbol{u}_l^{(\tau)}$, with $l = 1,2,\dots,q$ to the principal factors and $\varsigma_l = \boldsymbol{F}_l^{(\tau)}$, with $l = 1,2,\dots,q$ to the principal components.

Summarizing, this criterion is based in the minimum sum of the relative losses of inertia.

3.5.1.2 Maximization of Inertia Explained

The system desired is $(\boldsymbol{v}_l)$, $l = 1,2,\dots,q$. The inertia of the cloud $N_I{}^{(T)}$ explained by the system $(\boldsymbol{v}_l)_{l=1,2,\dots,q}$ is $\sum_{l=1}^{q}(\boldsymbol{v}_l)'V_t\boldsymbol{v}_l$.

The second criterion aims to maximize the inertia of all projected clouds.

$$\max_{v_1,v_2,\dots,v_q} \sum_{t=1}^{T}\sum_{l=1}^{q}(\boldsymbol{v}_l)'V_t\boldsymbol{v}_l = \max_{v_1,v_2,\dots,v_q} \sum_{l=1}^{q}(\boldsymbol{v}_l)'V_t\boldsymbol{v}_l$$

The vectors $(\boldsymbol{v}_l)_{l=1,2,\dots,q}$ are the q eigenvectors of the matrix $QV = \sum_{t=1}^{T} QV_t$ associated to the q highest eigenvalues, through the PCA.

This procedure, therefore, leads to a PCA in the cloud of individuals nT with respect to its centre of gravity and defined by the p variables, where $\sum_t^T V_t$ is the inertia matrix obtained by overlap of tables $Y_1, Y_2, \dots, Y_T$.

3.5.1.3 Sequential Search of a New Axis System

The third criterion, the sequential search of a new axis system, is performed in the following sequential way: firstly, should be selected $\boldsymbol{v}_1$, where $\sum_{t=1}^{T} \cos^2(\boldsymbol{u}_1^{(t)}, \boldsymbol{v}_1)$ is maximum and, on average, the angle $(\boldsymbol{u}_1^{(t)}, \boldsymbol{v}_1)$ is minimum. Then, select $\boldsymbol{v}_l$ orthogonal to the subspace formed by $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{l-1}$ and $\sum_{t=1}^{T} \cos^2(\boldsymbol{u}_l^{(t)}, \boldsymbol{v}_l)$ should be maximum. These steps should be applied successively until $\boldsymbol{v}_q$ is found.

$\boldsymbol{v}_1$ is the eigenvector associated with the highest eigenvalue of the matrix $U_1 U_1{}'$

$\boldsymbol{v}_l$ is the eigenvector associated with the highest eigenvalue of the matrix

$$\left[\prod_{h=1}^{l-1}(I_p - \boldsymbol{v}_h \boldsymbol{v}_h{}')\right] U_l U_l{}' \left[\prod_{h=1}^{l-1}(I_p - \boldsymbol{v}_h \boldsymbol{v}_h{}')\right]$$

where I_p is the identity matrix of size p and U_l is the matrix of size $p \times T$ defined by:

$$U_l = \left[\boldsymbol{u}_l^{(1)}\ \boldsymbol{u}_l^{(2)}\ \dots\ \boldsymbol{u}_l^{(T)}\right].$$

This method favours the first axes compared to the last ones, which means that gives decreasing importance to the axes. This way, the first axes are more important, because they explain most of the inertia.

Through the calculation of the indices, it is possible to compare the results obtained by the three previous criteria. The indices are the following:

$$\Phi(t, \boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_l) = \frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{l=1}^{q} \lambda_l^t - \sum_{l=1}^{q} (\boldsymbol{v}_l)' V_t \boldsymbol{v}_l}{\sum_{l=1}^{q} \lambda_l^t}$$

and

$$f_2(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_l) = \sum_{t=1}^{T} \sum_{l=1}^{q} \cos^2(\boldsymbol{v}_l, \boldsymbol{u}_l^{(t)})$$

3.5.1.4 Global Search of a New Axis System

The fourth criterion consists in selecting, directly, vectors $(\boldsymbol{v}_l)_{l=1,2,\dots,q}$ orthonormed, where $f_2(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_l)$ is maximum.

Unlike the sequential method, this one assigns identical importance to the axes (first and last axes).

The resolution of this problem is not possible analytically, but there are numerical methods to find a solution.

3.6 Compromise and Interpretation of the Trajectories of Individuals

The trajectories in a specific axis allow the evaluation of the proximities and oppositions between the individuals relative to the variables correlated with this specific axis, as well as the stability over the several years.

To obtain the trajectories of individuals and the compromise object there are different ways. In one of them, the compromise is the object V_τ. The trajectories of individuals are obtained from the projection of the individuals of each table on the selected axis system, where the other tables are taken as additional elements in the PCA on the table X_τ. In this criterion, the compromise object of the individuals are their positions in the year τ and to obtain the compromise object of the variables it is performed a PCA on the table X_τ.

Following other criterion, the compromise is the object $\sum_{t=1}^{T} V_t$ and through the PCA on this matrix, the compromise object of the variables is obtained. In this case, the compromise object of the individuals is obtained from an average of the coordinates of the trajectories of each individual on each axis. The trajectories are determined by the individuals' positions in the PCA of the overlap of tables $Y_1, Y_2, \dots, Y_T$ mentioned before.

4. Data Base, Preliminary Analysis and Application of the Method

4.1 Database

The database applied on this dissertation was collected from PORDATA (PORDATA, 2018) and contains 5 different tables, that refers the 5 different years in study (2008, 2010, 2012, 2014 and 2016). For each data table there are 30 observations, that in this case are some European countries[1] that are represented in Table 4.1:

Austria	Italy
Belgium	Latvia
Bulgaria	Lithuania
Croatia	Malta
Cyprus	Netherlands
Czech Republic	Norway
Denmark	Poland
Estonia	Portugal
Finland	Romania
France	Slovakia
Germany	Slovenia
Greece	Spain
Hungary	Sweden
Iceland	Switzerland
Ireland	United Kingdom

Table 4.1 - European Countries in study

Therefore, the data tables contain the same individuals described by the same variables in different years. The 14 variables/indicators and the respective abbreviations, in parentheses, are also presented in Table 4.2 and are grouped by categories.

When we deal with large databases, it is common the existence of missing values. Missing values happen when there are not any data value to an observation in a variable. It can appear due to diverse reasons and have a significant impact on the conclusions. In order to achieve the best conclusions possible, there was a need of solving these problems by looking at the years immediately before or after that there are not in study, or in other cases through the mean of the closest years.

[1]Only a few European countries were analysed, mostly because of the missing values. The other ones did not have all the information needed to apply the method.

Enrolled Students as % of Resident Population by Level of Education
▪ Pre-Primary Education **(Pre-Primary/PP)** ▪ Primary Education **(Primary/P)** ▪ Lower Secondary Education **(Lower Secondary/LS)** ▪ Upper Secondary Education **(Upper Secondary/US)** ▪ Tertiary Education **(Tertiary/T)**
Students Enrolled in Tertiary Education in Public System (%) (Public/Pb)
Population having Tertiary Education as a % of the Population by Age Group
▪ Age 25 to 34 **(25-34)** ▪ Age 35 to 44 **(35-44)** ▪ Age 45 to 54 **(45-54)** ▪ Age 55 to 64 **(55-64)**
Females in Total of Enrolled Students (%) having Tertiary Education (Females/F)
Unemployment Rate by Level of Education Completed
▪ Less than Primary, Primary and Lower Secondary Education **(Unemployment less than Secondary/ULS)** ▪ From Upper Secondary Education to Post-Secondary Non-Tertiary Education **(Unemployment with Secondary to Non-Tertiary/US-NT)** ▪ Tertiary Education **(Unemployment with Tertiary/UT)**

Table 4.2 - The 14 Variables

The data tables described previously can be consulted in Annexes (Table A.1 to Table A.5).

To perform the statistical analysis of the databases was used the software R (www.r-project.org), SPSS (***S****tatistical* ***P****ackage for the* ***S****ocial* ***S****ciences)*, SPAD (***S****ystème* ***P****our l'****A****nalyse des* ***D****onnées*) and Excel (Microsoft Office Excel).

4.2 Preliminary Analysis of the Data Set

Firstly, a preliminary analysis was performed to know more about each variable and the relations between them.

4.2.1 Univariate Analysis

Univariate Analysis is the analysis of only one variable/attribute. The main goal in this analysis is to describe the data and analyse the values of each one of the variables. To perform this analysis, were calculated some descriptive statistics, such as Mean, Median, Standard Deviation, Coefficient of Variation, Skewness Coefficient, Kurtosis Coefficient, Minimum and Maximum to describe and summarize the information about the variables. The tables with this information and the boxplots relative to each variable in study are presented below.

	Mean	Median	Standard Deviation	Coefficient of Variation (%)	Skewness Coefficient	Kurtosis Coefficient	Mini mum	Maxi mum
Pre-Primary	2.85	2.80	0.80	28.06	0.14	-0.31	1.30	4.60
Primary	6.06	6.10	1.80	29.73	0.67	0.23	3.50	10.80
Lower Secondary	4.42	4.25	1.13	25.54	1.94	5.57	3.00	8.60
Upper Secondary	4.71	4.65	1.29	27.45	1.02	1.93	2.40	8.30
Tertiary	4.23	3.95	1.06	25.16	0.41	-0.72	2.30	6.40
Public	75.49	85.95	28.02	37.12	-1.72	1.98	0.00	100.00
25-34	31.88	32.45	9.49	29.76	-0.05	-1.40	17.70	48.00
35-44	26.81	27.95	9.27	34.57	-0.15	-1.27	11.30	43.80
45-54	22.81	23.50	8.08	35.41	-0.23	-1.19	8.30	36.60
55-64	18.95	19.25	6.94	36.59	-0.07	-1.10	7.20	31.60
Females	56.23	55.90	4.09	7.27	0.09	-0.31	49.00	64.40
Unemployment less than Secondary	11.29	10.35	6.66	58.99	2.68	10.34	4.60	39.30
Unemployment with Secondary to Non-Tertiary	5.75	6.05	2.25	39.06	-0.04	-0.62	1.80	10.50
Unemployment with Tertiary	3.29	3.15	1.45	44.18	0.88	0.51	1.30	6.80

Table 4.3 - Summary Statistics for 2008

	Mean	Median	Standard Deviation	Coefficient of Variation (%)	Skewness Coefficient	Kurtosis Coefficient	Mini mum	Maxi mum
Pre-Primary	2.94	2.90	0.82	27.90	0.12	-0.32	1.40	4.70
Primary	5.94	5.95	1.81	30.53	0.84	0.89	3.50	11.10
Lower Secondary	4.21	4.05	1.09	25.92	1.51	3.08	2.70	7.80
Upper Secondary	4.67	4.55	1.30	27.77	1.15	2.06	2.50	8.30
Tertiary	4.39	4.20	0.95	21.57	0.40	-0.53	2.60	6.50
Public	74.77	83.60	27.50	36.78	-1.72	2.02	0.00	100.00
25-34	34.18	36.65	8.96	26.22	-0.08	-1.36	20.70	48.80
35-44	28.92	30.25	9.68	33.48	-0.10	-1.35	13.70	45.80
45-54	23.99	24.75	8.44	35.19	-0.10	-1.08	10.50	38.90
55-64	19.89	20.60	6.84	34.40	-0.20	-1.14	8.00	30.50
Females	55.94	56.35	4.02	7.18	-0.35	0.08	46.10	63.40
Unemployment less than Secondary	17.10	14.15	9.84	57.53	1.25	1.04	5.90	44.20
Unemployment with Secondary to Non-Tertiary	10.07	8.15	5.13	50.97	0.99	0.12	2.90	21.80
Unemployment with Tertiary	5.40	4.70	2.59	47.88	0.79	-0.16	1.70	11.20

Table 4.4 - Summary Statistics for 2010

	Mean	Median	Standard Deviation	Coefficient of Variation (%)	Skewness Coefficient	Kurtosis Coefficient	Mini mum	Maxi mum
Pre-Primary	3.06	3.05	0.76	24.96	-0.24	-0.49	1.50	4.40
Primary	5.92	5.75	1.87	31.60	0.85	0.91	3.50	11.30
Lower Secondary	3.93	3.85	0.97	24.79	1.49	3.32	2.60	7.20
Upper Secondary	4.58	4.40	1.25	27.23	1.39	2.16	2.70	8.20
Tertiary	4.38	4.20	0.85	19.40	0.44	-0.61	2.90	6.00
Public	75.18	83.10	26.52	35.28	-1.86	2.58	0.00	100.00
25-34	36.34	39.00	8.70	23.94	-0.07	-0.99	22.50	53.60
35-44	31.47	32.60	9.97	31.66	-0.13	-1.41	15.60	47.50
45-54	25.47	27.50	8.40	32.98	-0.11	-1.16	11.90	41.30
55-64	21.24	21.75	7.37	34.70	-0.05	-0.96	8.60	36.00
Females	55.69	55.80	3.47	6.24	-0.20	-0.61	49.10	62.50
Unemployment less than Secondary	18.40	15.60	9.32	50.68	0.97	0.65	6.20	44.60
Unemployment with Secondary to Non-Tertiary	10.72	9.25	6.09	56.85	1.12	0.96	2.70	27.70
Unemployment with Tertiary	6.02	5.25	3.80	63.26	1.75	3.28	1.80	18.40

Table 4.5 - Summary Statistics for 2012

	Mean	Median	Standard Deviation	Coefficient of Variation (%)	Skewness Coefficient	Kurtosis Coefficient	Mini mum	Maxi mum
Pre-Primary	3.13	3.05	0.83	26.44	0.27	0.35	1.50	5.20
Primary	5.96	5.85	1.84	30.92	0.93	1.43	3.50	11.50
Lower Secondary	3.78	3.70	0.90	23.78	1.27	1.88	2.60	6.50
Upper Secondary	4.40	4.00	1.17	26.55	1.23	0.98	2.90	7.50
Tertiary	4.28	4.35	0.87	20.30	0.45	-0.25	2.90	6.20
Public	75.75	83.70	26.50	34.98	-1.91	2.76	0.00	100.00
25-34	39.12	40.45	8.10	20.71	-0.08	-0.82	24.20	54.00
35-44	34.82	36.15	9.84	28.26	-0.16	-1.21	17.10	49.90
45-54	26.81	26.60	8.53	31.80	-0.19	-0.90	12.00	43.70
55-64	22.49	22.95	7.61	33.85	-0.16	-0.95	8.50	35.70
Females	55.45	55.95	3.47	6.26	-0.51	0.11	47.40	62.40
Unemployment less than Secondary	17.85	16.85	8.40	47.05	0.91	0.67	6.70	41.10
Unemployment with Secondary to Non-Tertiary	10.41	9.10	6.16	59.22	1.55	2.83	3.10	30.20
Unemployment with Tertiary	6.06	4.75	3.95	65.18	2.11	4.87	2.20	20.00

Table 4.6 - Summary Statistics for 2014

	Mean	Median	Standard Deviation	Coefficient of Variation (%)	Skewness Coefficient	Kurtosis Coefficient	Mini mum	Maxi mum
Pre-Primary	3.12	3.00	0.81	25.83	0.30	0.79	1.40	5.20
Primary	6.09	6.15	1.86	30.51	0.85	1.31	3.50	11.60
Lower Secondary	3.76	3.55	0.86	22.98	0.99	0.58	2.60	6.10
Upper Secondary	4.28	3.90	1.15	26.97	1.14	0.59	2.60	6.90
Tertiary	4.16	4.05	0.91	21.96	0.53	-0.04	2.70	6.50
Public	75.23	85.65	27.27	36.25	-1.79	2.16	0.00	100.00
25-34	40.73	41.65	8.08	19.83	-0.15	-0.50	24.80	56.20
35-44	37.24	38.45	9.97	26.77	-0.27	-1.29	20.50	51.90
45-54	28.48	28.75	9.10	31.94	-0.18	-1.04	13.00	45.10
55-64	23.60	24.90	8.06	34.16	-0.24	-1.00	8.70	36.90
Females	55.33	55.70	3.46	6.25	-0.08	0.29	48.20	63.70
Unemployment less than Secondary	15.44	14.60	6.56	42.50	0.70	0.10	4.50	31.60
Unemployment with Secondary to Non-Tertiary	8.60	7.35	5.06	58.77	1.78	4.15	3.10	26.20
Unemployment with Tertiary	5.04	3.95	3.51	69.71	2.13	5.64	1.40	18.00

Table 4.7 - Summary Statistics for 2016

Looking at the evolution between the years 2008 and 2016, the means of the variables enrolled students in Lower and Upper Secondary and Females having Tertiary Education have a decrease along this time. The mean of the Females having Tertiary Education is always higher than 50%, which means that, in percentage, there are more females than males with Tertiary Education. There is an increase in the means of the variables enrolled students in Pre-Primary and Tertiary until the years 2014 and 2010 respectively, and then a decrease until 2016, although with minimum variations. On the other hand, the variable enrolled students in Primary decrease until 2012 and then increase. This variable has always the highest mean when compared to the other levels of education, probably because this level is more accessible to everyone. The mean of the students in Tertiary Education in Public System is quite constant in the period studied. The values for Public System are high, so, consequently, Private System would have lower values due to the reduced number of students enrolled in Private System. There is an increase in the means for the variables related to the age, with the population having Tertiary Education and ages between 25 and 34 with the highest means and the group with ages between 55 and 64 with the lowest values for the mean. That means that, with the evolution of the years, there are more people with Tertiary Education, and the youngest persons are the ones with higher levels of education, as expected. In relation to the unemployment rate, the means increase until 2012 for all the levels of education in study and then decrease until

2016. In the case of the Unemployment Rate for Tertiary Education, the mean only starts the decrease in 2014. This variable, when compared to the Unemployment Rate for the other two levels of education, has the lowest values, which confirms that more years of study implies less unemployment. It is possible to conclude that, over the years, the population is ever more instructed, and the younger population has the highest levels of education. Consequently, with the increase of educated people, the unemployment rates are decreasing in the last years.

The standard deviation (SD) is a measure used to quantify the dispersion or variation of a set of values, but to compare the dispersion between the variables the coefficient of variation (CV) is more adequate, because do not depend on the unit of variable's measurement nor the magnitude of the variable values. If the variables have high values of CV that means that the dispersion is high, the data are spread out over an extensive range of values in relation to the mean and there are more differences between the countries for that variables. The variables that present high variability in relation to the mean are Unemployment Rate less than Secondary, Unemployment Rate with Secondary to Non-Tertiary, Unemployment Rate with Tertiary, Public System and ages between 55-64. On the other hand, when variables have low values of CV, such as Females, that shows that the data values tend to be closer to the mean of the set, the expected value, that means that the percentage of Females having Tertiary Education is more a less the same over the years.

The form of the distribution of each variable was also analysed. "Skewness is a measure of the asymmetry of the distribution of a variable and kurtosis measures the tailedness of the distribution." (Gudivada, 2017). One distribution is considered skewed if one of the tails is longer than the other. When a distribution has a extended tail in the positive direction it is considered positively skewed. On the other hand, when the long tail happens in the negative direction the distribution has a negative skew. The variables with positively skewed distribution for every year are Primary, Lower Secondary, Upper Secondary, Tertiary, Unemployment Rate less than Secondary and Unemployment Rate with Tertiary. Public, 25-34, 35-44, 45-54 and 55-64 are the variables with negatively skewed distribution during all the period in study. The variable Unemployment Rate with Secondary to Non-Tertiary obtain a positive skewness coefficient for all the years except to 2008, and the variable Females is the opposite. The variable Pre-Primary has negative value

for skewness coefficient in 2012 and positive values for the remaining years. In general, the Skewness measure does not have very high values.

As referred before, kurtosis measures if the tails are lighter or heavier relative to normal distribution. When the variables have high values for the kurtosis that means that the distribution is heavy, the tail data exceed the tails of normal distribution and there is a tendency to exist outliers. These variables are Lower Secondary, Upper Secondary and Public. The variables Primary, Unemployment Rate with Secondary to Non-Tertiary and Unemployment Rate with Tertiary obtained, also, high values since 2012, the opposite of Unemployment less than Secondary. When the variables have lower values for the kurtosis, such us 25-34, 35-45, 44-54, 55-64 and Tertiary, for every year, the tail data is generally less extreme than the tails of normal distribution, there is a tendency of lack of outliers and the distribution is flat.

In Figure 4.1 are presented the boxplots for each variable over time.

Through the observation of the boxplots, it is possible to analyse the dispersion of the variables, the existence of outliers and the variability of data. The country with more enrolled students in Pre-Primary, from 2014 to 2016, is Estonia and it is an outlier. On the other hand, Ireland is an outlier for enrolled students in Primary from 2012 to 2016, that means that, in this case and for these years, is the country with more students enrolled. The enrolled students in Lower Secondary and Upper Secondary have different outliers during all the periods. For Lower Secondary, in 2008, Lithuania, Malta and Germany are outliers owe to high values, but in 2010 only Lithuania and Malta remain as outliers. Lithuania is outlier from 2008 to 2016, being the only one for years 2012 and 2016. In 2014 Germany is also an outlier with Lithuania. For Upper Secondary, from 2008 to 2012, there are three outliers, Iceland, Belgium and Finland. In 2014, Iceland and Belgium are the ones that remain as outliers for this variable, but in 2016, Finland return as outlier as well as Iceland, Belgium and United Kingdom. For the variable enrolled students in Tertiary Education there are no outliers, which confirms the tendency verified by the low value of kurtosis. The variable Public System has values highly concentrated around the second quartile (the median). There are five inferior outliers during all the period that are Belgium, Cyprus, Estonia, Latvia and UK. Finland is also an outlier in 2016. These countries have more students enrolled in Tertiary Education in Private System than in Public, probably because

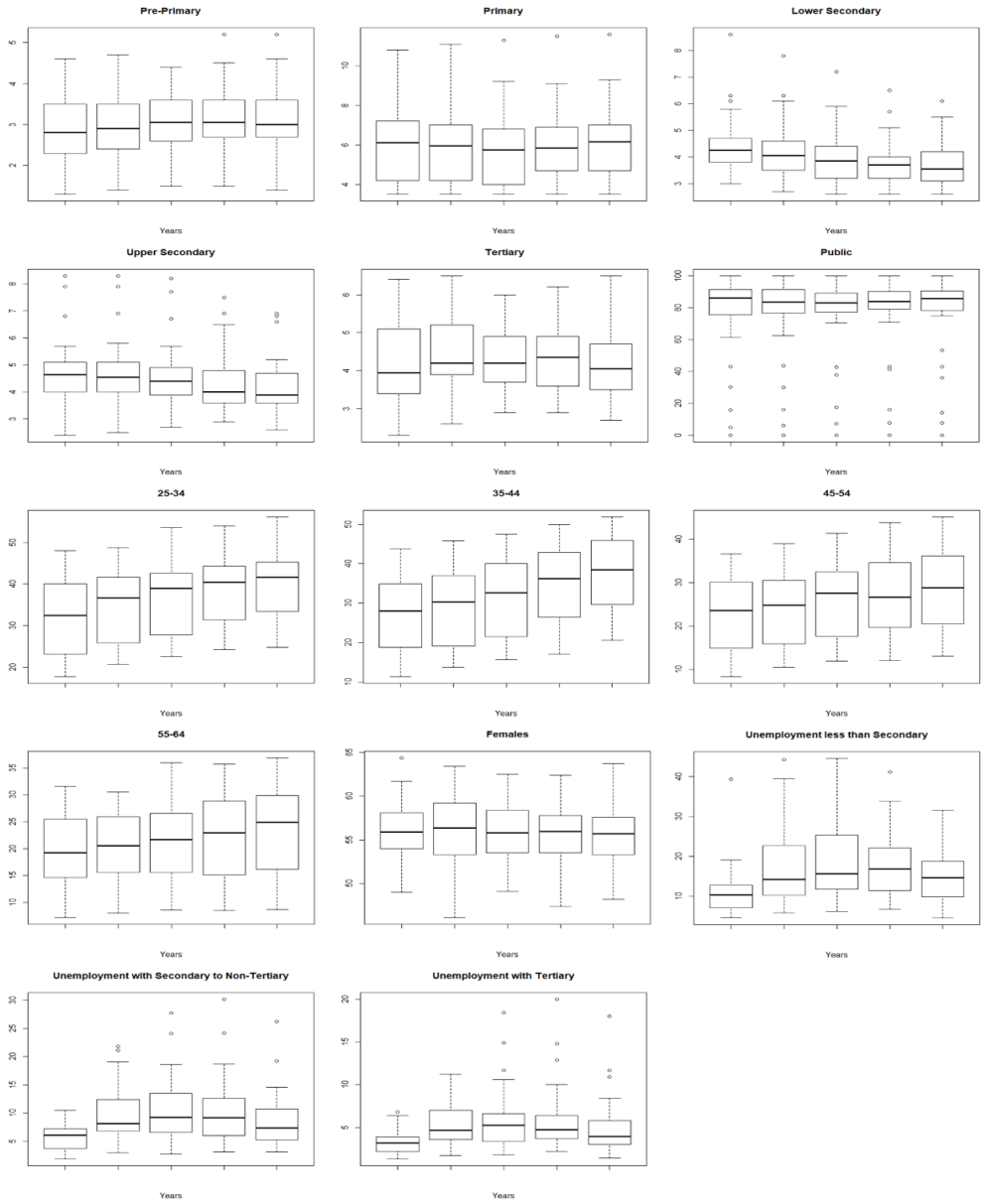

Figure 4.1 – Boxplots of the Variables

they are rich countries or are countries that invest a lot on education, especially in private schools.

In terms of the variables related to the group ages of population having Tertiary Education there are no outliers, as was also expected due to the low values of kurtosis. The medians of these variables, in general, increase from 2008 until 2016, and are higher than the means. That means that since 2008 there is an increase of people's concern about being instructed. Populations want to have more education regardless their age, what justifies the increase, over time, for all age groups. The distributions are negatively skewed since the data constitute higher frequency of high values scores (median is closer to upper quartile).

There is also high dispersion for these variables, the distance between the minimum and the maximum is large.

The case of Females is a different one. Latvia and Iceland are the countries with the highest values for Females with Tertiary Education in 2008, being considered as outliers. In general, the values for median are quite constants during all the period and very similar with the values for mean. The countries analysed are only from Europe, which justifies equal access to education between women and men. If countries from other continents had been analysed, would probably exist cases of discrimination.

The distribution of the variables related to unemployment rate is relatively similar. In general, the medians have an increase until the year 2012 and then a decrease until 2016. In these 3 variables there are several superior outliers. For Unemployment less than Secondary, Slovakia is an outlier in the years 2008, 2010 and 2014. On the other hand, Unemployment with Secondary to Non-Tertiary has two outliers, Greece and Spain, from 2012 to 2016 and two different ones in the year 2010, Latvia and Lithuania. Unemployment with Tertiary has Portugal as outlier in 2008. In 2012, Greece and Spain are also outliers of this variable. From 2014 to 2016, it is possible to find three outliers, Cyprus, Greece and Spain. These outliers are positive, which means that these countries have higher rates of unemployment than the means. Fortunately, Portugal stopped to be an outlier due to the reduction of the unemployment rate. In general, these variables start with a highly concentrated distribution in 2008 and six years later have a positive skewed distribution, where the low values are predominant, meaning that there is a tendency to the reduction of the unemployment over the years. The distributions of these 3 variables also confirm that for high levels of education, the unemployment rates are lower.

4.2.2 Bivariate Analysis

In the case of Bivariate Analysis, it is the analysis of exactly two variables, simultaneously. This analysis is used to discover if there are relationships between two variables or not, to know the strength of the association and to figure out what are the relations. In this way, it becomes easier to predict a possible value for a certain variable. For example, if the correlation between two variables is positive and high, when one of them increases it is probable that the other one will also increase.

	PP	P	LS	US	T	Pb	25-34	35-44	45-54	55-64	F	ULS	US-NT	UT
PP	1.00	0.03	0.00	0.41	0.06	-0.09	0.17	0.23	0.25	0.31	0.44	0.05	-0.18	-0.16
P		1.00	-0.21	0.15	0.00	-0.02	**0.68**	0.64	0.48	0.37	0.06	-0.44	-0.34	-0.07
LS			1.00	-0.32	-0.06	0.26	-0.08	-0.20	-0.17	-0.10	0.08	0.23	0.07	-0.15
US				1.00	0.23	-0.22	0.07	0.22	0.27	0.27	0.37	0.05	-0.28	-0.28
T					1.00	-0.17	0.14	0.16	0.19	0.18	0.48	0.03	0.16	0.14
Pb						1.00	-0.21	-0.16	-0.26	-0.25	-0.24	0.01	-0.05	0.00
25-34							1.00	**0.88**	**0.78**	**0.67**	-0.06	-0.33	-0.19	-0.06
35-44								1.00	**0.95**	**0.86**	-0.05	-0.31	-0.27	-0.16
45-54									1.00	**0.95**	0.00	-0.23	-0.34	-0.28
55-64										1.00	0.10	-0.21	-0.34	-0.34
F											1.00	0.22	-0.11	-0.22
ULS												1.00	0.45	0.11
US-NT													1.00	**0.83**
UT														1.00

PP – Pre-Primary**, P –** Primary**, LS –** Lower Secondary**, US –** Upper Secondary**, T –** Tertiary**, Pb –** Public**, F –** Females**, ULS –** Unemployment less than Secondary**, US-NT -** Unemployment with Secondary to Non-Tertiary**, UT -** Unemployment with Tertiary

Table 4.8 - Linear Correlation Coefficients in 2008

	PP	P	LS	US	T	Pb	25-34	35-44	45-54	55-64	F	ULS	US-NT	UT
PP	1.00	0.07	0.04	0.28	0.06	-0.43	0.13	0.30	0.35	0.44	0.61	0.02	-0.33	-0.39
P		1.00	-0.14	0.41	0.46	-0.11	0.59	**0.71**	0.62	0.44	0.19	-0.30	-0.06	0.02
LS			1.00	-0.02	-0.01	0.12	0.13	0.13	0.08	0.18	-0.15	0.04	-0.14	-0.24
US				1.00	0.24	-0.23	0.12	0.41	0.47	0.35	0.29	-0.40	-0.31	-0.21
T					1.00	-0.01	0.58	0.59	**0.65**	0.57	-0.02	0.06	0.32	0.37
Pb						1.00	-0.26	-0.31	-0.38	-0.46	-0.26	0.08	0.00	0.06
25-34							1.00	**0.87**	**0.73**	**0.69**	0.09	-0.01	0.08	0.09
35-44								1.00	**0.93**	**0.85**	0.08	-0.14	-0.02	0.00
45-54									1.00	**0.95**	0.00	-0.17	-0.06	-0.02
55-64										1.00	0.01	-0.12	-0.11	-0.10
F											1.00	-0.01	-0.30	-0.34
ULS												1.00	0.62	0.50
US-NT													1.00	**0.95**
UT														1.00

PP – Pre-Primary**, P –** Primary**, LS –** Lower Secondary**, US –** Upper Secondary**, T –** Tertiary**, Pb –** Public**, F –** Females**, ULS –** Unemployment less than Secondary**, US-NT -** Unemployment with Secondary to Non-Tertiary**, UT -** Unemployment with Tertiary

Table 4.9 - Linear Correlation Coefficients in 2016

In the Tables above are presented the correlations in 2008 and in 2016 to compare and study the evolution of the correlations during the period in analysis. The coefficient of linear correlation, that measures the strength of relationships between variables, can assume values between -1 and 1.

The variables Pre-Primary, Lower Secondary, Upper Secondary, Public, Females and Unemployment less than Secondary are not highly correlated with none of the other variables. It was considered the existence of a strong positive correlation, that is when one of the variables increases the other one highly correlated will also increase, to coefficients with values higher than 0.65. Therefore, there are strong positive correlations in all the groups of age between each other. That can happen due to the fact that, for example,

parents with Tertiary Education are more propense to demand their children to also have Tertiary Education. There is also strong positive correlation between Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary, probably because there are some countries with values very different that the other ones. In 2008, the variables Tertiary and the four groups of age were not correlated with high positively values, but in 2016, it is possible to see a significant increase of these correlations, all higher than 0.57. That means that there was an evolution, over the years, in which the Tertiary Education increased the enrolled students for all the ages.

When compared these two years, 2008 has a high positive correlation for Primary and 25-34 and 2016 have a high correlation for Primary and 35-44 instead, that reveals a progress in people's concern about being instructed. Over the years, there is an increase of people's willingness to be educated, regardless of their age, which justifies older people wanting to have Primary Education.

In the remaining variables the coefficients were, in general, the same in 2008 and 2016, so there are not many differences in the behaviour of these variables.

4.3 Double Principal Component Analysis – Application of the Method

SPAD was the software used to apply the Double Principal Component Analysis that is presented in this section. The main purposes of this analysis, as previously mentioned, are find a common structure between the different years, identify the differences and similarities over time and analyse the evolution of the variables and countries.

4.3.1 Interstructure

The first phase of the method is the Interstructure, that is the analysis of global evolution by performing a Principal Component Analysis on the clouds' centres of gravity.

In Table 4.10 are presented the eigenvalues of the interstructure and the percentage of explained variance.

The first component has an eigenvalue of 9.06 and explains 64.70% of the variance, the second component explains 29.20% of the variance that corresponds to an eigenvalue of 4.09, and so on.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**9.06**	**64.70**	**64.70**
2	**4.09**	**29.20**	**93.90**
3	0.83	5.96	99.86
4	0.02	0.14	100.00

Table 4.10 – Eigenvalues and % of Inertia of the interstructure

There are some methods to determine the number of factors that should be retained in the analysis. The Kaiser's criterion is one of the most used and consists in retaining the factors that have eigenvalues higher than 1, which means that the principal components are "more informative" than the initial variables. Cattel is another criterion that is based on retaining the eigenvalues for which the variation between two eigenvalues consecutive is relatively large. The Pearson's criterion keeps the number of components that explain approximately 80% of the total inertia.

In this case, using the Kaiser's criterion, the information between the variables can be explained in only two components. These two axes explain 93.90% of the total variance of the initial variables, that meets the requirement of the Pearson's criterion. So, the number of components that should be selected is 2.

It is also important to study the variables that contribute the most to the formation of each axis. The correlations between the variables and the axes, presented in Table 4.11 and also represented in Figure 4.2., in general, are high (since almost every variable are close to the borders of the correlation circle).

Variable	Axis 1	Axis 2
PP	**0.99**	0.03
P	-0.06	**-0.97**
LS	**-0.99**	-0.02
US	-0.92	0.38
T	-0.22	**0.97**
Pb	0.20	-0.47
25-34	**0.98**	-0.17
35-44	**0.97**	-0.24
45-54	**0.96**	-0.23
55-64	**0.97**	-0.22
F	**-0.99**	0.07
ULS	0.66	0.75
US-NT	0.63	0.78
UT	0.73	0.68

Table 4.11- Correlations between the variables and the principal components

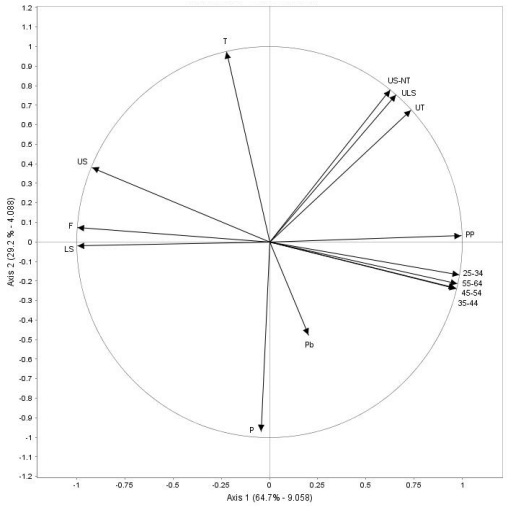

Figure 4.2 - Correlation Circle of the variables in the principal plan 1–2

For the analysis of the Correlation Circle, it is possible to visualize that the variables Pre-Primary, Lower Secondary, Females 25-34, 35-44, 55-64 and 45-54, are the ones that are more correlated with the first principal component and have more impact on the first axis' formation. The variables that contribute the most to the formation of the second component, that means, that have higher coordinates for the second axis, are Primary and Tertiary.

To analyse the principal components it is essential to understand the concepts of Absolute Contribution of the variable/individual i for the kth principal component/axis (CTA_i^k), that measures the contribution of the variable/individual i to the formation of the kth principal component/axis. The Relative Contribution (CTR_i^k) of the component/axis k for the variable/individual i, measures the quality of the representation of the variable/individual i on kth axis. It is considered that one variable/individual is well represented if the relative contribution is higher than 0.5.

By analysing the Table 4.12, it is possible to conclude that years 2008 and 2016 are the ones that contribute the most to the formation of the first axis and are well represented (CTR>0.5). All years, except 2014, have high impact on the second axis.

	Axis 1			**Axis 2**		
Time	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
2008	-4.80	**50.86**	0.83	-2.15	**22.54**	0.17
2010	-1.87	7.72	0.37	2.24	**24.52**	0.53
2012	0.63	0.88	0.08	2.15	**22.63**	0.91
2014	2.76	16.82	0.80	0.24	0.27	0.01
2016	3.28	**23.72**	0.60	-2.48	**30.04**	0.34

Table 4.12 – Coordinates and contributions of individuals (interstructure of DPCA)

The Euclidean image represented in Figure 4.3 reveals some proximity between 2010 and 2012. The years 2008 and 2016 are the years more separated in time and with less similarities. The first axis shows a constant evolution by the variables correlated with this axis with an increase of the coordinates over the years while the second axis has an increase of the coordinates until 2012 and then a decrease.

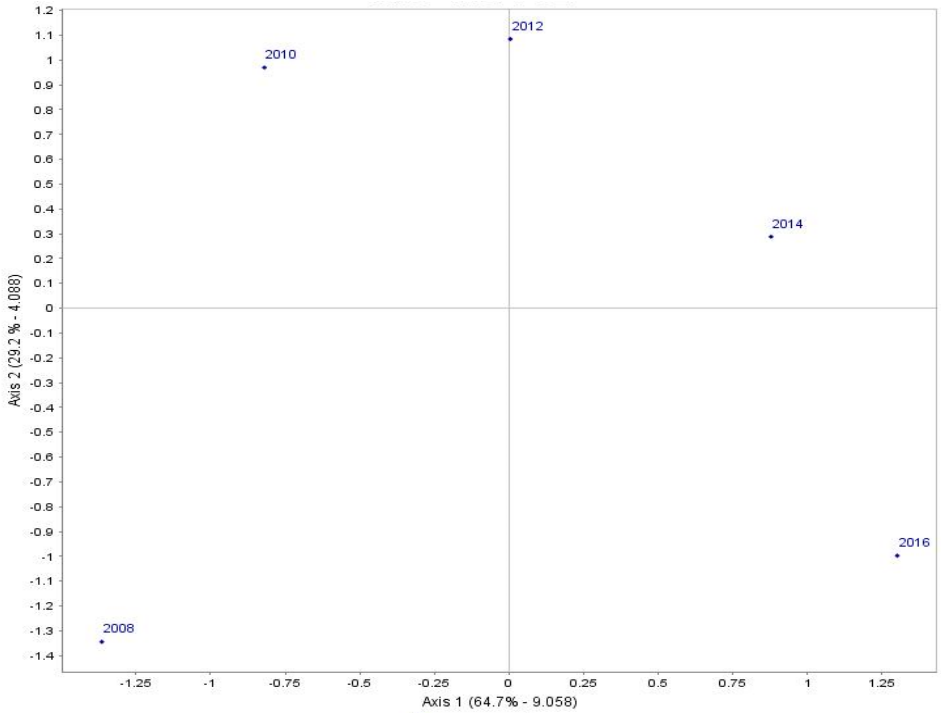

Figure 4.3 - Euclidean Image of the Interstructure

4.3.2 Analysis of the Clouds of Individuals

As referred before, this analysis was performed by SPAD. The analysis consists in a normed PCA, for every year in study, with the same weight for all the countries.

4.3.2.1 Year 2008

Based on Table 4.13, it is possible to see that the first axis has an eigenvalue of 4.67 that corresponds to 33.39% of the explained variance, that means that this axis explains 33.39% of the initial information. The eigenvalue for the second axis is 2.13 and the explained variance is 15.24% and so on. According to Kaiser's criterion ($\lambda_k > 1$), should be retained 4 axes. On the other hand, to Pearson's criterion, the number of axes that should be selected is 6, and to Cattel's criterion is 2. For this analysis, 4 axes were chosen, that correspond to Kaiser's criterion. These 4 axes/components are the ones that explained the most part of the total variance. The remaining ones have lower contributions overall.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.67**	**33.39**	**33.39**
2	**2.13**	**15.24**	**48.62**
3	**1.87**	**13.33**	**61.95**
4	**1.29**	**9.24**	**71.19**
5	0.98	7.03	78.22
6	0.88	6.29	84.51
7	0.72	5.12	89.63
8	0.61	4.38	94.02
9	0.39	2.76	96.78
10	0.23	1.62	98.40
11	0.12	0.83	99.23
12	0.06	0.46	99.69
13	0.04	0.27	99.96
14	0.01	0.04	100.00

Table 4.13 – Eigenvalues and % of Inertia for 2008

After the selection of the axes, the original variables that most contribute to the axes' formation were studied. Since were chosen 4 axes, the study is about axes 1 and 2, 1 and 3 and 1 and 4. The coordinates of the variables in the axes corresponds to their correlations (r) with the principal components, and the correlation coefficient only assumes values between -1 and 1. Therefore, the variables are considered well represented in one axis if the square of r is higher than 0.5. The correlations/coordinates of the variables are presented in Table 4.14 and are also represented in Figure 4.4.

It is possible to analyse that the variables 35-44, 45-54, 55-64, 25-34 and Primary are the ones with the highest negative correlation with the first principal component. The variables 35-44, 45-54, 55-64 and 25-34 are well represented in PC1. On the other hand, the variable Primary is not very well represented in PC1, because its relative contribution is, approximately, 0.45, but is well represented in [PC1, PC2]. According to the weights of each variable in each component it is possible to nominate the first component as Tertiary Education by age groups since the weights of the variables 25-34, 35-44, 45-54 and 55-64 are very high in this component.

Females and Upper Secondary are the variables most correlated with the second component, with Females well represented in PC2 (CTR=0.67) and Upper Secondary that is not well represented in PC2 (CTR=0.37). In terms of the third component, the variables Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary are the most significant ones, with correlations of 0.79 and relative contributions of approximately 0.62.

The fourth component is highly negatively correlated with Lower Secondary, that is well represented in PC4. It is possible to classify the fourth axis as "Lower Secondary Education".

Variable	Axis 1	Axis 2	Axis 3	Axis 4
PP	-0.34	0.56	-0.06	-0.14
P	**-0.67**	-0.30	0.02	0.15
LS	0.25	0.04	-0.24	**-0.83**
US	-0.40	**0.61**	-0.05	0.36
T	-0.19	0.44	0.49	-0.03
Pb	0.28	-0.29	-0.36	-0.30
25-34	**-0.82**	-0.29	0.25	-0.20
35-44	**-0.93**	-0.21	0.20	-0.16
45-54	**-0.93**	-0.07	0.14	-0.18
55-64	**-0.88**	0.04	0.08	-0.24
F	-0.14	**0.82**	0.05	-0.06
ULS	0.42	0.44	0.24	-0.40
US-NT	0.52	-0.06	**0.79**	-0.17
UT	0.39	-0.28	**0.79**	0.09

Table 4.14 - Correlations between the variables and the principal components

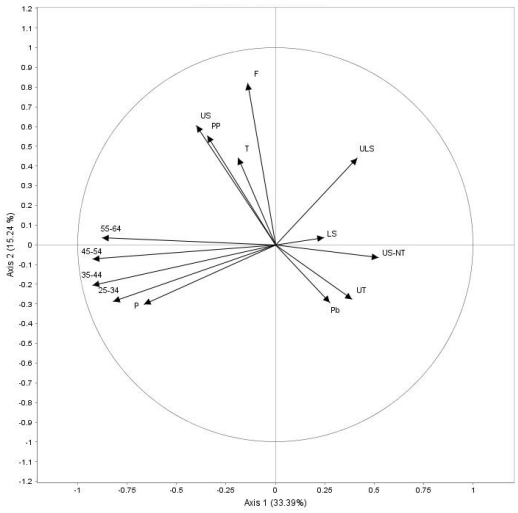

Figure 4.4 - Correlation Circle of the variables in the principal plan 1–2

In Figure 4.4 are represented the correlations of the variables with the first two axes. The correlation circles for the plan 1-3 and plan 1-4 are presented in the Annexes (Figure B.1). The variables with more contribution to the formation of the components are the ones with higher distance from the centre and, consequently, higher coordinates on the axes, that means, higher correlation between the variables and the axes. It is possible to confirm in Figure 4.4 the conclusions obtained from Table 4.14.

On the other hand, it is also possible to analyse the contributions of each individual to the formation of the axes. To the selection of the individuals that contribute the most to the formation of the axes it is commonly used to select the ones with higher absolute contribution until the cumulated sum of absolute contributions be approximately equal to 80%.

To study the countries that contribute the most to the formation of the first axis it is necessary to analyse the absolute contributions that are in Annexes in Table B.1. These countries are described in the Table 4.15. Croatia, Finland, Italy, Norway and Portugal are well represented on the first axis while Belgium, Cyprus, Czech Republic, Denmark, Estonia, Malta, Iceland, Netherlands, Romania, Slovakia are not very well represented on the first axis (CTR<0.5), but are well represented in the plan [1,2], except Belgium and Malta.

Negative Coordinates		Positive Coordinates	
Norway	Denmark	Slovakia	Italy
Iceland	Belgium	Croatia	Romania
Finland	Netherlands	Portugal	Czech Republic
Estonia	Cyprus	Malta	

Table 4.15 - Countries that contribute the most for the 1st Axis (2008)

The same analysis was performed to the second axis and the countries more relevant are presented below. None of them is well represented on the second axis, they are only well represented in the plan [1,2], except Greece, Latvia and Spain.

Negative Coordinates		Positive Coordinates	
Ireland	Cyprus	Slovakia	Estonia
Switzerland	Netherlands	Latvia	Hungary
Greece	Spain	Iceland	

Table 4.16 - Countries that contribute the most for the 2nd Axis (2008)

Once again, the countries that contribute the most to the third axis are in Table 4.17. In this axis, Austria is the only country well represented on the third axis, the other ones are well represented in plans.

Negative Coordinates		Positive Coordinates
Malta	Iceland	Spain
Austria	Denmark	Greece
Czech Republic	Norway	Latvia
Netherlands		Estonia

Table 4.17 - Countries that contribute the most for the 3rd Axis (2008)

In Table 4.18 are the countries with more contribution for the fourth axis and Lithuania is the one with higher quality of representation and the only well represented on this axis.

Negative Coordinates		Positive Coordinates	
Lithuania	Slovakia	Italy	Slovenia
Germany	Spain	Belgium	Romania

Table 4.18 - Countries that contribute the most for the 4th Axis (2008)

In Figure 4.5 are the countries (individuals) in the plan formed by the first two principal axes that represent the positioning of each country in the plan. As well as the correlation circles, the representations of the plans 1-3 and 1-4 are in Annexes (Figure B.2 and B.3).

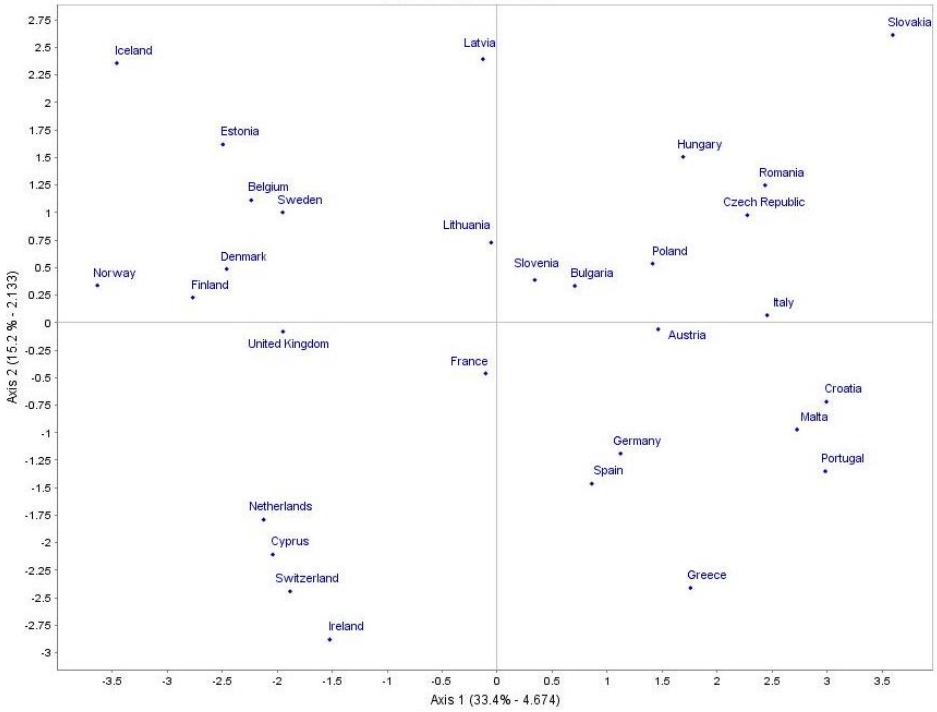

Figure 4.5 – Individuals' representation in the principal plan [1,2] for 2008

The first axis opposes the countries Slovakia, Croatia, Portugal, Malta, Italy, Romania and Czech Republic to the countries Norway, Iceland, Finland, Estonia, Denmark, Belgium, Netherlands and Cyprus, because the first group of countries has low values for Primary, 25-34, 35-44, 45-54 and 55-64. While for the second group of countries it is the opposite, since these countries have high values for the variables Primary, 25-34, 35-44, 45-54 and 55-64. Therefore, the first axis opposes countries with high percentage of population having Tertiary Education by age groups to countries with low percentages.

In the case of the second axis, Slovakia, Latvia, Iceland, Estonia and Hungary have high values for the variable Females and Upper Secondary, unlike Ireland, Switzerland, Greece, Cyprus, Netherlands and Spain, that have low values for this variable.

The variables Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary have high values for Spain, Greece, Latvia and Estonia and low values for Malta, Austria, Czech Republic, Netherlands, Iceland, Denmark and Norway, in terms of the third axis.

Lastly, the fourth axis, denominated before as Lower Secondary Education, opposes Italy, Belgium, Slovenia and Romania that have low percentage of students

enrolled with education inferior to Secondary to Lithuania, Germany, Slovakia and Spain that have high percentage of the students with this level of education.

4.3.2.2 Year 2010

For 2010 is presented an analysis similar to the one performed for 2008. In this case, according to Kaiser's criterion it is also considered 4 axes, but if it was considered Pearson's criterion (% of the explained inertia higher than 80%) 5 axes should be selected.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.35**	**31.07**	**31.07**
2	**3.00**	**21.40**	**52.48**
3	**1.97**	**14.09**	**66.57**
4	**1.28**	**9.12**	**75.69**
5	0.95	6.80	82.48
6	0.70	4.99	87.48
7	0.60	4.29	91.77
8	0.39	2.80	94.57
9	0.35	2.53	97.10
10	0.24	1.71	98.82
11	0.10	0.73	99.54
12	0.04	0.29	99.83
13	0.01	0.10	99.93
14	0.01	0.07	100.00

Table 4.19 - Eigenvalues and % of Inertia for 2010

To analyse the variables and their correlations with the axes is presented the Table 4.20. The variables more correlated with the first axis are the ones with higher coordinates, that are 35-44, 45-54, 55-64, 25-34 and Primary. All these variables are well represented in PC1, except Primary (r^2=0.41) that is only well represented in [PC1, PC2]. In terms of the second axis, Unemployment Secondary to Non-Tertiary, Unemployment less than Secondary and Unemployment with Tertiary are the variables with higher negative coordinates. The 3 variables are well represented in PC2. The second axis can be nominated as the axis that describes the unemployment rates. Also, with negative correlations, the variables Upper Secondary, Females and Pre-Primary are the ones with higher impact on the third component and the ones that contribute the most to the formation of the third component. Upper Secondary and Females are well represented in PC3, but Pre-Primary is only well represented in a plan. Relatively to the fourth component the conclusion is the same that in 2008. Lower Secondary is the variable with higher negative correlation with the fourth axis and is well represented in this component. The same conclusions are in Figure 4.6, where the correlations of the variables with the first

two axes are represented. It is possible to conclude that the variables described before as more contributors are the ones with higher distance to the centre.

Variable	Axis 1	Axis 2	Axis 3	Axis 4
PP	0.30	-0.10	**-0.64**	-0.29
P	**0.64**	0.31	0.16	-0.01
LS	-0.24	-0.10	0.24	**-0.83**
US	0.42	0.19	**-0.73**	0.10
T	0.40	-0.51	-0.26	0.04
Pb	-0.35	0.30	0.12	-0.55
25-34	**0.87**	-0.03	0.30	-0.12
35-44	**0.95**	0.15	0.19	-0.10
45-54	**0.93**	0.07	0.13	-0.05
55-64	**0.88**	0.09	0.07	-0.11
F	0.11	-0.45	**-0.73**	-0.20
ULS	0.01	**-0.85**	-0.03	-0.22
US-NT	0.09	**-0.95**	0.25	0.00
UT	0.09	**-0.80**	0.35	0.24

Table 4.20 - Correlations between the variables and the principal components

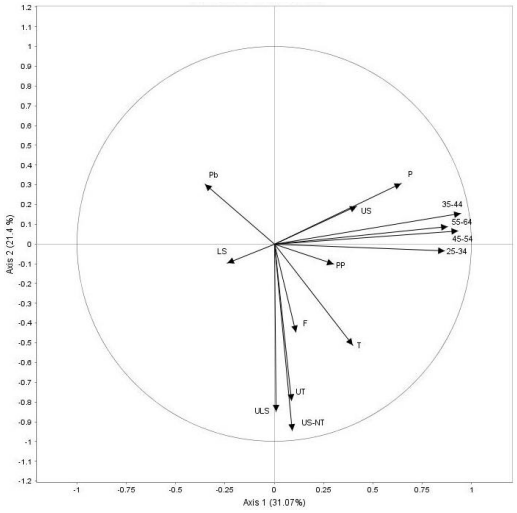

Figure 4.6 - Correlation Circle of the variables in the principal plan 1–2

Looking at Table B.2 in the Annexes and analysing the absolute contributions, the individuals that contribute the most to the formation of the first axis are in Table 4.21. All the countries are well represented on the first axis, except Belgium, Estonia, Iceland, Ireland, Norway and Slovakia. These countries are only well represented in the plan [1,2], excluding Iceland and Ireland.

Negative Coordinates		Positive Coordinates	
Malta	Czech Republic	Finland	Norway
Italy	Croatia	Belgium	Ireland
Romania	Slovakia	Iceland	Sweden
Portugal		Estonia	

Table 4.21 - Countries that contribute the most for the 1st Axis (2010)

In the same way, the Table 4.22 represents the countries with higher contributions to the formation of the second axis. Latvia, Netherlands and Switzerland are the ones well represented on this axis. The other ones are only well represented in plan [1,2], except Spain.

Negative Coordinates		Positive Coordinates
Latvia Lithuania Estonia	Slovakia Spain	Switzerland Netherlands Norway

Table 4.22 - Countries that contribute the most for the 2nd Axis (2010)

The countries that influenced more the formation of the third axis are presented below. None of these countries is well represented on this axis.

Negative Coordinates		Positive Coordinates	
Iceland Belgium Sweden	Romania Czech Republic	Ireland Greece Cyprus	Switzerland Spain Netherlands

Table 4.23 - Countries that contribute the most for the 3rd Axis (2010)

The highest absolute contributions for the fourth axis were obtained by the countries in Table 4.24. Once more, none of these countries has relative contribution higher than 0.5, so, the countries are only well represented in plans.

Negative Coordinates		Positive Coordinates	
Lithuania Germany Malta	Denmark France	Latvia Greece Cyprus United Kingdom	Slovenia Italy Estonia

Table 4.24 - Countries that contribute the most for the 4th Axis (2010)

In Figure 4.7, the countries are represented in the plan [1,2]. The correlation circles and the representations for the plans 1-3 and 1-4 are in Annexes (Figures B.4, B.5 and B.6).

Finland, Belgium, Iceland, Estonia, Norway, Ireland and Sweden have values higher than the mean of the variables Primary, 25-34, 35-44, 45-54 and 55-64, that are highly correlated with the first component. These countries oppose to Malta, Italy, Romania, Portugal, Czech Republic, Croatia and Slovakia that have lower values than the mean of these variables, in the principal plan.

In terms of the second axis, the one that resumes the unemployment rates, the countries with low values to Unemployment Secondary to Non-Tertiary, Unemployment less than Secondary and Unemployment with Tertiary are Switzerland, Netherlands and Norway, that are the opposite of Latvia, Lithuania, Estonia, Slovakia and Spain that have high values. That means that in Switzerland, Netherlands and Norway the unemployment rates are low for all the levels of education. On the other hand, Latvia, Lithuania, Estonia,

Slovakia and Spain, in 2010, have a lot of unemployment regardless to the level of education.

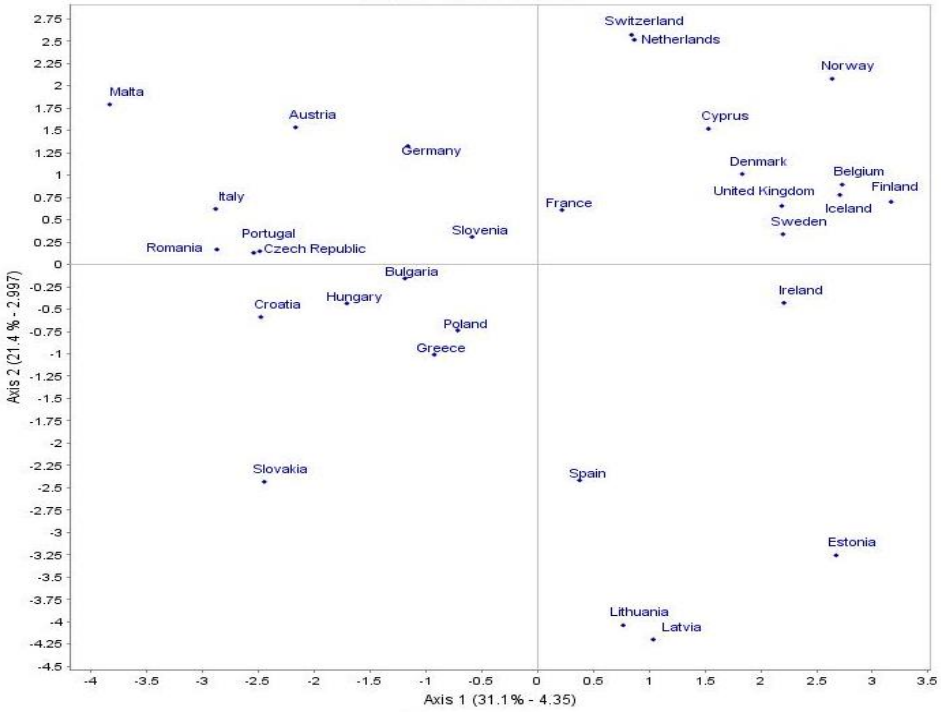

Figure 4.7 - Individuals' representation in the principal plan [1,2] for 2010

The third axis opposes the countries Iceland, Belgium, Sweden, Romania and Czech Republic, that have high values to Pre-Primary, Upper Secondary and Females to Ireland, Greece, Cyprus, Switzerland, Spain and Netherlands that have low values to these variables.

Lithuania, Germany, Malta, Denmark and France have high values for the variable Lower Secondary that is the one with higher negative correlation with the fourth axis. On the other hand, Latvia, Greece, Cyprus, United Kingdom, Slovenia, Italy and Estonia have few people with education Lower than Secondary. As can be confirmed in Figure B.6 in Annexes, Lithuania is, by far, the country with more correlation with this axis, that means that is the country with more students enrolled on this level of education.

4.3.2.3 Year 2012

In this case, the Kaiser's and Pearson criteria are coincident, therefore, the number of axes selected is 5, that explain 82.74% of the total variance. (Table 4.25)

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.62**	**33.03**	**33.03**
2	**2.76**	**19.71**	**52.74**
3	**1.80**	**12.87**	**65.61**
4	**1.31**	**9.35**	**74.96**
5	**1.09**	**7.78**	**82.74**
6	0.67	4.78	87.52
7	0.52	3.71	91.24
8	0.42	2.97	94.21
9	0.36	2.54	96.75
10	0.21	1.52	98.27
11	0.16	1.13	99.40
12	0.04	0.31	99.71
13	0.03	0.22	99.93
14	0.01	0.07	100.00

Table 4.25 - Eigenvalues and % of Inertia for 2012

For 2012, the variables more negatively correlated with the first component are 35-44, 45-54, 55-64, 25-34 and Primary, where all the variables are well represented in PC1 (relative contribution higher than 0.5), except Primary. The same conclusions were obtained for 2008. The second component has Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary with the highest coordinates. These 2 variables are the ones with more impact on the formation of the second component and are well represented in PC2. Females and Pre-Primary are the variables more negatively correlated with the third component, and only Females is well represented in this component. For the third axis, the variable Females is the one with higher weight in this component, so Females with Tertiary Education is what better describes this axis. The fourth and fifth components have Lower Secondary and Public as the variables with more negatively impact, respectively, in these components. Lower Secondary is well represented in PC4, but Public is not well represented in PC5, only in plans. Once more, the conclusion to the fourth component is the same that in 2008 and 2010.

In Figure 4.8 it is possible to see the same conclusions for the first two axes, where the variables with higher distance to the centre for each axis (higher coordinates) are the ones with more impact on the formation of the component 1 and 2. As in the other years, the correlation circles for the plan 1-3, plan 1-4 and plan 1-5 are presented in Figure B.7 in Annexes.

Variable	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
PP	-0.35	-0.35	**-0.60**	-0.15	0.07
P	**-0.68**	0.16	0.24	0.16	-0.45
LS	0.08	0.14	0.08	**-0.90**	0.15
US	-0.49	-0.55	-0.19	0.14	-0.36
T	-0.47	0.28	-0.47	-0.16	-0.39
Pb	0.38	-0.03	0.27	-0.51	**-0.66**
25-34	**-0.79**	0.41	0.04	-0.04	0.07
35-44	**-0.93**	0.25	0.18	-0.05	-0.01
45-54	**-0.93**	0.22	0.08	-0.06	0.13
55-64	**-0.87**	0.18	0.08	-0.13	0.24
F	-0.19	-0.38	**-0.78**	0.00	-0.09
ULS	0.28	0.56	-0.58	-0.20	0.13
US-NT	0.30	**0.88**	-0.24	0.13	-0.09
UT	0.33	**0.81**	-0.10	0.26	-0.20

Table 4.26 - Correlations between the variables and the principal components

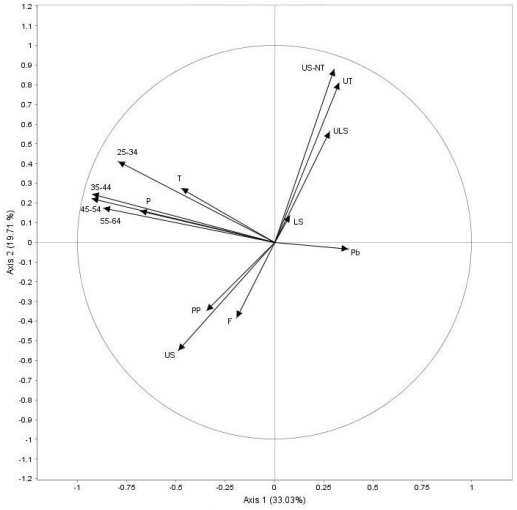

Figure 4.8 - Correlation Circle of the variables in the principal plan 1–2

The Table B.3 that is in Annexes presents the coordinates, absolute contributions and relative contributions in 2012 for 5 axes. The countries that contribute the most to the formation of the first axis are in Table 4.27. Croatia, Denmark, Finland, Italy, Norway, Portugal, Romania and Sweden are well represented on the first axis, because their relative contributions are higher than 0.5. Belgium, Iceland and Malta are well represented in plan [1,2].

Negative Coordinates		Positive Coordinates	
Iceland	United Kingdom	Croatia	Slovakia
Finland	Estonia	Italy	Portugal
Norway	Sweden	Romania	Malta
Belgium	Denmark		

Table 4.27 - Countries that contribute the most for the 1st Axis (2012)

The Table 4.28 represents the countries that more contribute to the second axis. The countries well represented on the second axis are Greece and Spain, the remaining ones are only well represented in plan [1,2], except Lithuania.

Negative Coordinates		Positive Coordinates	
Malta	Iceland	Greece	Ireland
Romania	Austria	Spain	Lithuania

Table 4.28 - Countries that contribute the most for the 2nd Axis (2012)

In a similar way, the countries with higher absolute contributions are presented below. Only Switzerland is well represented on the third axis. The other ones are well represented in plans.

Negative Coordinates		Positive Coordinates	
Latvia	Iceland	Switzerland	Netherlands
Slovakia	Lithuania	Ireland	Malta
Estonia		Germany	

Table 4.29 - Countries that contribute the most for the 3rd Axis (2012)

The fourth axis has Latvia, Cyprus, United Kingdom, Belgium, Greece, Italy, Lithuania, Germany and Slovakia as the ones with more contribution for the formation of the fourth axis, and Lithuania is the only well represented on this axis.

Negative Coordinates	Positive Coordinates	
Lithuania	Latvia	Belgium
Germany	Cyprus	Greece
Slovakia	United Kingdom	Italy

Table 4.30 - Countries that contribute the most for the 4th Axis (2012)

In terms of the axis 5, the countries below are the ones with higher contribution of the formation of this axis. None of them is well represented on this axis. The countries are only well represented in a plan.

Negative Coordinates		Positive Coordinates	
Iceland	Denmark	Estonia	Latvia
Greece	Ireland	United Kingdom	Cyprus
		Germany	

Table 4.31 - Countries that contribute the most for the 5th Axis (2012)

Based on Figure 4.9, it is possible to see the representation of the individuals in the plan formed by the first two axes. The same representation is also in Annexes in Figures B.8, B.9 and B.10 for the plan 1-3, plan 1-4 and plan 1-5.

The first axis opposes Croatia, Italy, Romania, Slovakia, Portugal and Malta, that have lower values than the mean for Primary, 25-34, 35-44, 45-54 and 55-64, that are negatively correlated with the first axis to the countries Iceland, Finland, Norway, Belgium, United Kingdom, Estonia, Sweden and Denmark, that have high values for these variables.

On the second axis, the variables Greece, Spain, Ireland and Lithuania have high values for Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary. On the other hand, Malta, Romania, Iceland and Austria are the opposite with low values for these two variables, that means low unemployment rates.

Latvia, Slovakia, Estonia, Iceland and Lithuania are the countries that have higher values for Pre-Primary and Females, that are the variables more negatively correlated with the third axis. Switzerland, Ireland, Germany, Netherlands and Malta, on the other hand,

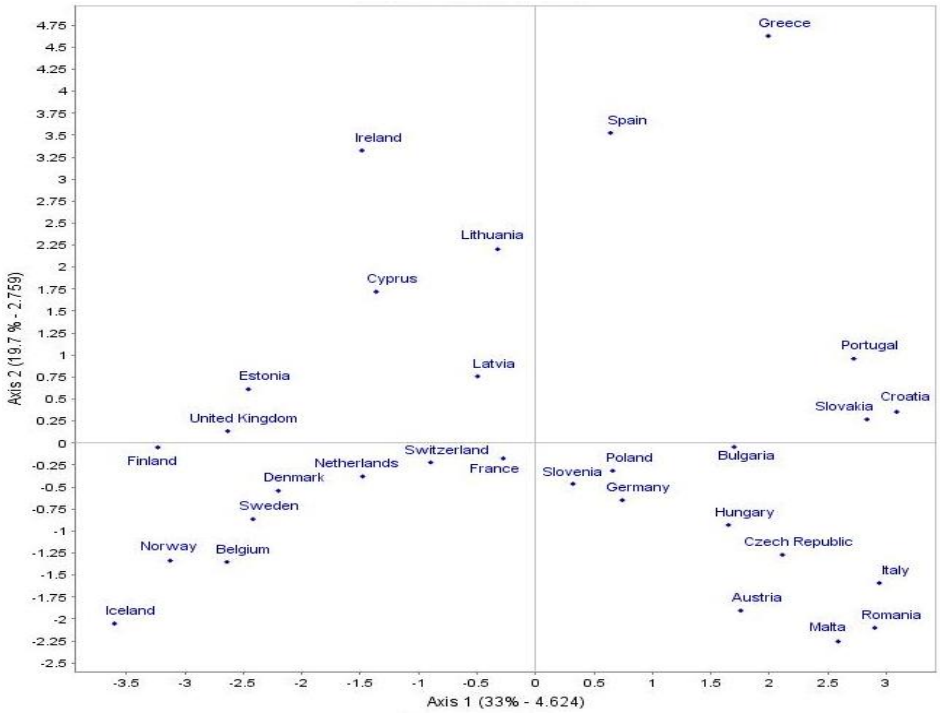

Figure 4.9 - Individuals' representation in the principal plan [1,2] for 2012

have the lowest values for these variables. Therefore, it is possible to say that the third axis opposes countries with high percentage of Females with Tertiary Education to countries with low percentage of Females in this level.

In terms of the fourth axis, Lithuania, Germany and Slovakia are the countries with more percentage of students enrolled in Lower Secondary Education, while Latvia, Cyprus, United Kingdom, Belgium, Greece and Italy are the ones with less percentage.

In terms of the fifth axis, Iceland, Greece, Denmark and Ireland have high values for Public System of Education, so it is possible to conclude that these countries have a high number of students enrolled in Tertiary Education studying in a Public System. On the other hand, Estonia, United Kingdom, Germany, Latvia and Cyprus are the countries with higher percentage of students enrolled in a Private System, which can be a consequence of the wealth of these countries and/or the progressive investment in systems of education. It is interesting to note that this last group of countries, specially Estonia, United Kingdom, Latvia and Cyprus were considered as negative outliers of variable Public in section 4.2.1, which confirms these conclusions.

4.3.2.4 Year 2014

Once more, the Kaiser and Pearson's criteria are coincident, so, the number of axes that is selected is 5, where the eigenvalues are higher than 1 and 83.47% of the variance are explained. The other 9 components have lower contributions globally.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.93**	**35.22**	**35.22**
2	**2.73**	**19.50**	**54.72**
3	**1.66**	**11.85**	**66.57**
4	**1.31**	**9.38**	**75.95**
5	**1.05**	**7.52**	**83.47**
6	0.67	4.80	88.26
7	0.56	3.98	92.24
8	0.35	2.53	94.77
9	0.33	2.39	97.16
10	0.23	1.63	98.79
11	0.10	0.75	99.54
12	0.03	0.24	99.78
13	0.02	0.13	99.91
14	0.01	0.09	100.00

Table 4.32 - Eigenvalues and % of Inertia for 2014

In Table 4.33 and in Figure 4.10 it is possible to see the same conclusions for the first two axes. The table contains the coordinates/correlations of the variables and in the figure are represented the variables with higher distance to the centre for each axis (higher coordinates), that are the ones with more influence on the formation of the component 1 and 2. The correlation circles for the plan 1-3, plan 1-4 and plan 1-5 are presented in Figure B.11 in the Annexes, as on the other years.

As in years 2008 and 2012, the variables that most contribute negatively to the formation of the first component are the variables related to the group ages and Primary. Once more, all these variables are well represented in PC1, except Primary. In terms of the second component, negatively correlated and well represented in the second component, Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary are the variables with higher impact on the formation of this component. In 2014, Females and Pre-Primary are the more contributors to the third component, although, only Females is well represented in PC3. Lower Secondary is, like in the other years, the variable that contributes the most to the formation of the fourth component and is well represented. On the contribution for the fifth component, and not so well represented in PC5, there is

the Public System of Education that stands out. It is possible to conclude that the fifth axis can be described as the Tertiary Education in Public System.

Variable	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
PP	-0.45	0.37	**0.62**	-0.24	0.05
P	**-0.67**	-0.19	-0.21	0.33	0.29
LS	0.01	0.04	-0.30	**-0.87**	0.02
US	-0.56	0.35	0.01	0.24	0.38
T	-0.58	-0.41	0.14	-0.14	0.44
Pb	0.39	-0.08	-0.46	-0.25	**0.67**
25-34	**-0.77**	-0.40	-0.03	-0.15	-0.10
35-44	**-0.91**	-0.25	-0.14	0.00	-0.07
45-54	**-0.93**	-0.22	-0.12	-0.05	-0.13
55-64	**-0.89**	-0.13	-0.04	-0.16	-0.22
F	-0.23	0.34	**0.77**	-0.04	0.30
ULS	0.37	-0.56	0.47	-0.37	-0.01
US-NT	0.32	**-0.90**	0.22	0.06	0.03
UT	0.30	**-0.86**	0.15	0.24	0.09

Table 4.33 - Correlations between the variables and the principal components

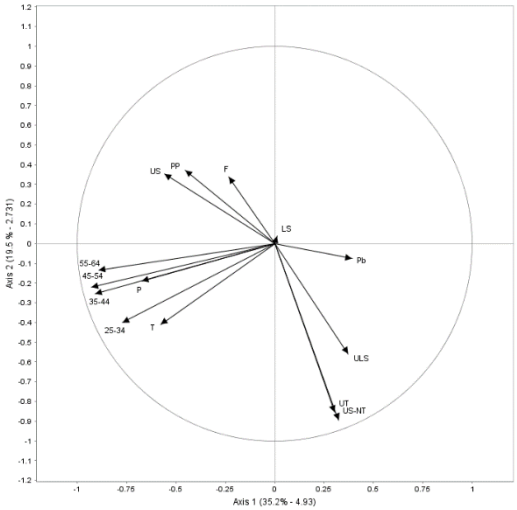

Figure 4.10 - Correlation Circle of the variables in the principal plan 1–2

In Table B.4, in the Annexes, are shown the individuals and their respective coordinates and absolute and relative contributions. For the first axis, the individuals that contribute the most to its formation are the ones in Table 4.34. Different of the other ones, Estonia, Malta and United Kingdom are not well represented on the first axis. From these 3 countries, Malta is the only one that is well represented in plan [1,2].

Negative Coordinates		Positive Coordinates	
Iceland	Belgium	Romania	Croatia
Norway	Sweden	Italy	Malta
Finland	Estonia	Slovakia	Portugal
United Kingdom			

Table 4.34 - Countries that contribute the most for the 1st Axis (2014)

It is also possible to see in Table 4.35, the countries that contribute the most to the formation of the second axis. Greece and Spain are the only ones with relative contributions higher than 0.5 for this axis. The remaining ones have only high relative contributions in a plan.

Negative Coordinates		Positive Coordinates
Greece	Ireland	Malta
Spain	Cyprus	Romania

Table 4.35 - Countries that contribute the most for the 2nd Axis (2014)

In Table 4.36 are the countries more contributors of the third axis. Only Latvia and Switzerland are well represented on this axis, the other ones are only well represented in a plan.

Negative Coordinates		Positive Coordinates	
Germany Switzerland Ireland	Netherlands Malta	Latvia Estonia	Slovakia Poland

Table 4.36 - Countries that contribute the most for the 3rd Axis (2014)

In a similar way, the fourth axis has the following countries with the higher absolute contributions. Only Lithuania is well represented on this axis.

Negative Coordinates		Positive Coordinates	
Lithuania Germany	Slovakia France	Malta Italy	United Kingdom

Table 4.37 - Countries that contribute the most for the 4th Axis (2014)

Finally, there are the countries that contribute the most to the formation of the fifth axis. In this case, none of them is well represented. These countries are only well represented in a plan.

Negative Coordinates		Positive Coordinates
United Kingdom Estonia Latvia	Switzerland Germany Cyprus	Iceland Denmark

Table 4.38 - Countries that contribute the most for the 5th Axis (2014)

In Figure 4.11 is possible to see the individuals represented in the plan formed by axes 1 and 2. The Figures with the representations of the plan 1-3, plan 1-4 and plan 1-5 are also in Annexes in Figures B.12, B.13 and B.14.

The first axis opposes Iceland, Norway, Finland, United Kingdom, Belgium, Sweden and Estonia, that are the countries with many students with Tertiary Education in the groups of age 25-34, 35-44, 45-54 and 55-64 to the countries Romania, Italy, Slovakia, Croatia, Malta and Portugal, that are the ones with less students enrolled in Tertiary Education by age groups.

Greece, Spain, Ireland and Cyprus have values higher than the mean for Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary that are highly negatively correlated to the second component. On the other hand, Malta and Romania have low values for the unemployment rates, as in 2012.

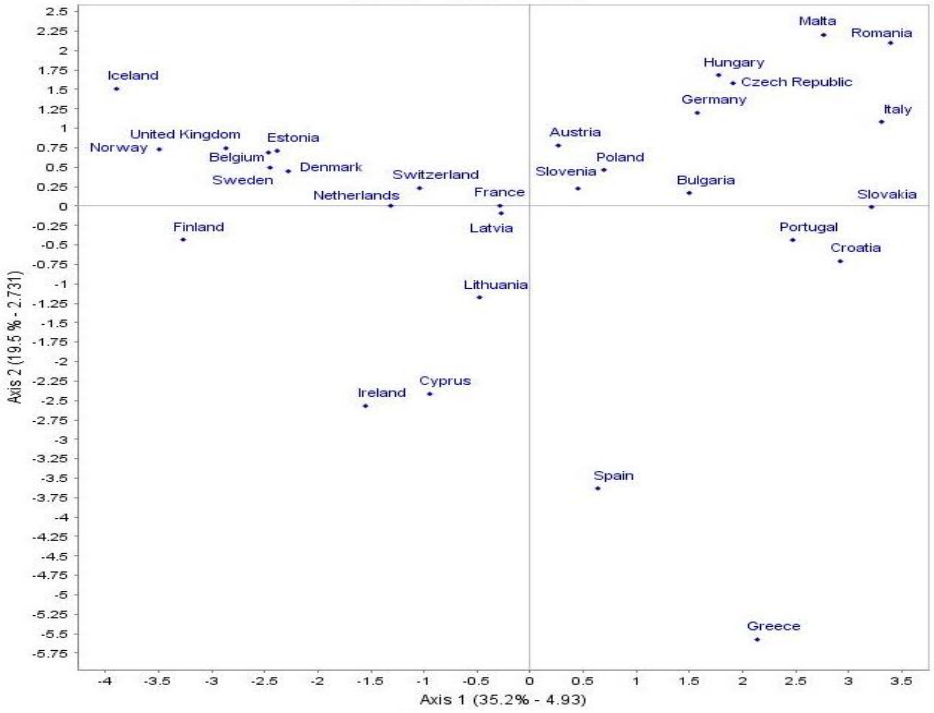

Figure 4.11 - Individuals' representation in the principal plan [1,2] for 2014

Pre-Primary and Females are variables that, for the third component, have high values for Latvia, Estonia, Slovakia and Poland, and low values for Germany, Switzerland, Ireland, Netherlands and Malta. Since the third axis is denominated by Females with Tertiary Education, Latvia, Estonia, Slovakia and Poland are the ones with more females with higher level of education.

In terms of the fourth axis, Lithuania, Germany, Slovakia and France are the countries with higher values to the Lower Secondary Education, contrarily to Malta, Italy and United Kingdom that have the lowest values. That means, that these 3 last countries have high levels of education.

Based on the fifth axis, the conclusions are similar to the year 2012. The Public System has more students enrolled in Tertiary Education in Iceland and Denmark. Through the Figure B.14 in Annexes it is possible to note that Iceland is by far the country that uses more the Public System. The Private System is more usual in United Kingdom, Estonia, Latvia, Switzerland, Germany and Cyprus, as was expected through the analysis of the outliers in Univariate Analysis.

4.3.2.5 Year 2016

Based on Table 4.39, it is possible to see that the first axis has an eigenvalue of 4.99 that corresponds to 35.65% of the explained variance, the eigenvalue for the second axis is 2.97 and the explained variance is 21.20%, and so on. According to the Kaiser's criterion ($\lambda_k > 1$), the number of axes that should be selected is 4 axes.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.99**	**35.65**	**35.65**
2	**2.97**	**21.20**	**56.85**
3	**1.57**	**11.22**	**68.07**
4	**1.31**	**9.39**	**77.46**
5	0.90	6.42	83.88
6	0.73	5.22	89.10
7	0.55	3.90	93.00
8	0.37	2.62	95.62
9	0.27	1.91	97.52
10	0.18	1.30	98.82
11	0.09	0.67	99.50
12	0.04	0.29	99.79
13	0.02	0.16	99.95
14	0.01	0.05	100.00

Table 4.39 - Eigenvalues and % of Inertia for 2016

Based on Table 4.40 and on Figure 4.12, it is possible to obtain the information about the coordinates/correlations of the axes. The variables with more impact on the formation of the first component are 35-44, 45-54, 55-64, 25-34, Primary and Tertiary, which still confirms that the first axis should be described as Tertiary Education by age groups. All these variables are well represented in PC1, except Tertiary. This variable is only well represented in [PC1, PC2]. Similarly, to 2010 and 2014, the variables Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary are the variables more correlated, negatively, with the second component. The variables are well represented in this component. Contrarily to 2010 and 2012, where Females and Pre-Primary have the highest negative coordinates of the third component, in 2014 and 2016, these variables are also the ones with higher contributions, but with positive coordinates. The variables are not well represented in PC3, only in a plan. All the years in study have the same conclusion to the fourth component, that is that Lower Secondary is the one with higher contribution in this component. This variable is also well represented in PC4. Summing up, the correlations between the variables and the axes are the expected based on the previous years.

In Figure 4.12 are represented the correlations of the variables with the first two axes. Once more, the variables more important to the axes are the ones with higher distance to the centre.

Variable	Axis 1	Axis 2	Axis 3	Axis 4
PP	-0.45	0.46	**0.61**	-0.22
P	**-0.71**	-0.08	-0.20	0.43
LS	-0.08	0.09	-0.33	**-0.80**
US	-0.54	0.34	-0.06	0.36
T	**-0.63**	-0.50	-0.07	0.08
Pb	0.44	-0.14	-0.52	0.02
25-34	**-0.80**	-0.30	-0.06	-0.14
35-44	**-0.95**	-0.16	-0.10	-0.07
45-54	**-0.95**	-0.13	-0.08	-0.07
55-64	**-0.90**	-0.06	0.00	-0.27
F	-0.22	0.48	**0.62**	0.19
ULS	0.24	-0.57	0.53	-0.36
US-NT	0.15	**-0.91**	0.27	0.05
UT	0.11	**-0.92**	0.16	0.21

Table 4.40 - Correlations between the variables and the principal components

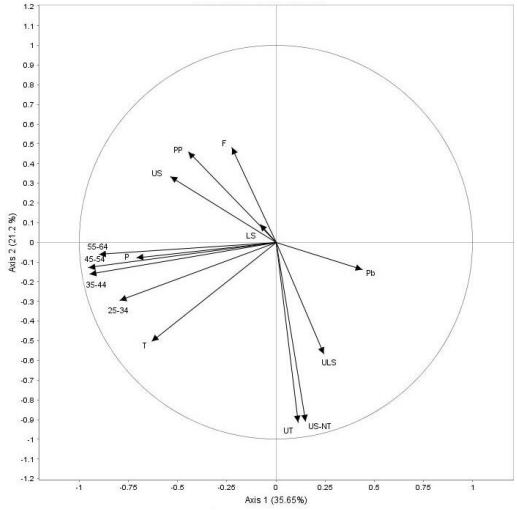

Figure 4.12 - Correlation Circle of the variables in the principal plan 1–2

According to the Table B.5 in Annexes, it is possible to analyse the individuals that the contribute the most to the formation of the four axes. In Table 4.41 are the countries that contribute the most to the formation of the first axis. All these countries are well represented on the first axis, except United Kingdom.

Negative Coordinates		Positive Coordinates	
Iceland	United Kingdom	Romania	Croatia
Finland	Sweden	Italy	Hungary
Norway	Belgium	Malta	Portugal
		Slovakia	Czech Republic

Table 4.41 - Countries that contribute the most for the 1st Axis (2016)

In terms of the second axis, Greece, Spain, Cyprus, Iceland, Malta and Ireland are the countries more relevant. Only Greece and Spain are well represented on this axis. The remaining ones are only well represented in a plan.

Negative Coordinates		Positive Coordinates
Greece	Cyprus	Iceland
Spain	Ireland	Malta

Table 4.42 - Countries that contribute the most for the 2nd Axis (2016)

The individuals with higher absolute contributions are presented below. On this axis, the third one, only Latvia and Netherlands are well represented.

Negative Coordinates		Positive Coordinates
Ireland Germany Switzerland	Netherlands Denmark	Latvia Estonia Slovakia

Table 4.43 - Countries that contribute the most for the 3rd Axis (2016)

The countries that contribute the most to the formation of the fourth axis are also in Table 4.44. Lithuania is well represented on the fourth axis. The other ones are only well represented in a plan.

Negative Coordinates		Positive Coordinates	
Lithuania Germany	France Slovakia	Iceland Malta	Italy

Table 4.44 - Countries that contribute the most for the 4th Axis (2016)

In Figure 4.13 is represented the plan formed by the first two axes. The correlation circles and the individuals' representations for the remaining plans are in Figures B.15, B.16 and B.17 of the Annexes.

The first axis opposes Iceland, Finland, Norway, United Kingdom, Sweden and Belgium, that are the countries with the highest values of the variables Primary, Tertiary, 25-34, 35-44, 45-54 and 55-64 to Romania, Italy, Malta, Slovakia, Croatia, Hungary, Portugal and Czech Republic, that have the lowest values for these variables, and consequently less population in Tertiary Education by age groups.

Greece, Spain, Cyprus and Ireland are the countries with higher values for Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary, that are the variables that contribute the most to the formation of the second component. On the other hand, Iceland and Malta are the countries with lower values for these variables. As it is possible to see in Figure 4.13, Greece is the country with high values to the unemployment rates, that means that in 2016 Greece is still a country with high unemployment.

Females and Pre-Primary are the variables more correlated with the third axis. This axis opposes Latvia, Estonia and Slovakia, that have high values for these two variables, to

Ireland, Germany, Switzerland, Netherlands and Denmark, that have low values. The last countries have more percentage of males in Tertiary Education than females.

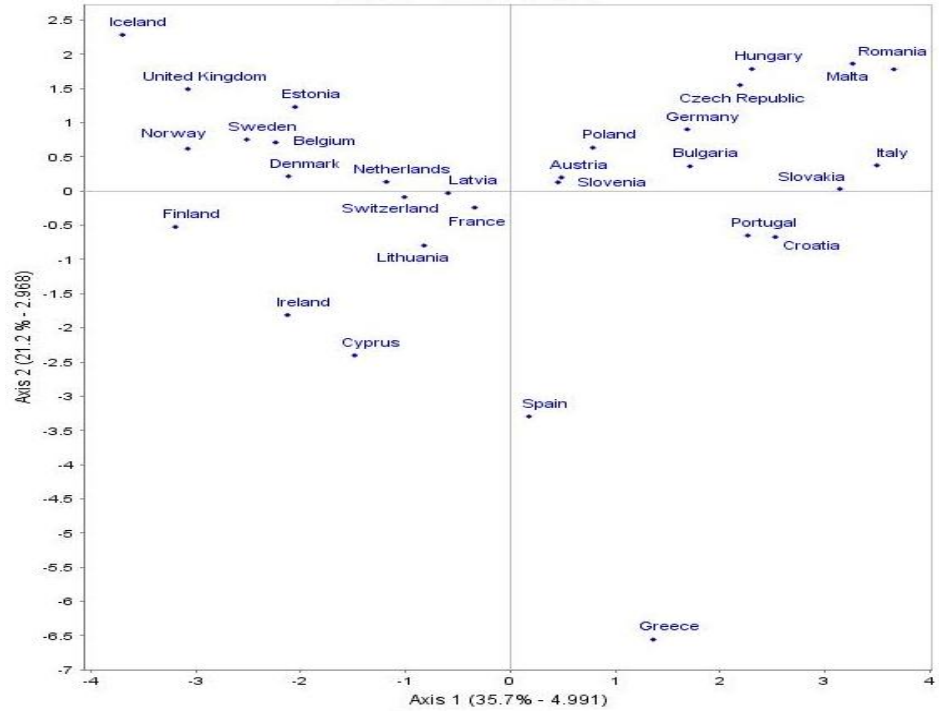

Figure 4.13 - Individuals' representation in the principal plan [1,2] for 2016

Finally, on the fourth axis, Lithuania, Germany, France and Slovakia are the countries with most people with education Lower than Secondary, that is what defines this axis. Iceland, Malta and Italy are the ones with less people with education Lower than Secondary.

In general, the behaviour of the data tables was similar when analysed separately.

4.3.3 Intrastructure

The last phase of the DPCA method is the Intrastructure. The principal objective is the representation of individuals in a common space over time. To perform the trajectories of the variables and individuals, it is necessary the selection of the new system of axes. As mentioned in the section 3.5.1, there are 4 criteria for choosing the best axis system. In this part are used the first two criteria. The first one is based on the system where the loss of inertia is minimum. In this case, 2012 was the system of axes selected, once it is the one with the minimum loss. Therefore, the compromise matrix of the variables is r_{12}.

In this case, as in section 4.3.2.3 and according to Kaiser and Pearson's criteria, five axes are selected, that explain 82.74% of the total variance (Table 4.25).

Variable	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
PP	-0.35	-0.35	**-0.60**	-0.15	0.07
P	**-0.68**	0.16	0.24	0.16	-0.45
LS	0.08	0.14	0.08	**-0.90**	0.15
US	-0.49	-0.55	-0.19	0.14	-0.36
T	-0.47	0.28	-0.47	-0.16	-0.39
Pb	0.38	-0.03	0.27	-0.51	**-0.66**
25-34	**-0.79**	0.41	0.04	-0.04	0.07
35-44	**-0.93**	0.25	0.18	-0.05	-0.01
45-54	**-0.93**	0.22	0.08	-0.06	0.13
55-64	**-0.87**	0.18	0.08	-0.13	0.24
F	-0.19	-0.38	**-0.78**	0.00	-0.09
ULS	0.28	0.56	-0.58	-0.20	0.13
US-NT	0.30	**0.88**	-0.24	0.13	-0.09
UT	0.33	**0.81**	-0.10	0.26	-0.20

Table 4.45 - Correlations between the variables and the principal components – 1st Criterion

As described before, the variables more negatively correlated with the first axis are 35-44, 45-54, 55-64, 25-34 and Primary. In terms of the second axis, Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary have the highest correlations. The variables more negatively correlated with the third component are Females and Pre-Primary. Lower Secondary and Public are, respectively, the variables more important to the fourth and fifth components.

The second criterion chosen is based on the maximization of inertia explained. The first five axes explain 80.58% of the inertia. According to Pearson and Kaiser's criteria, the number of axes selected is 5.

Axes	Variance of the axis (eigenvalue)	% of explained variance	Cumulated % of explained variance
1	**4.48**	**32.02**	**32.02**
2	**2.71**	**19.36**	**51.38**
3	**1.82**	**13.03**	**64.41**
4	**1.21**	**8.66**	**73.08**
5	**1.05**	**7.50**	**80.58**
6	0.71	5.04	85.62
7	0.62	4.44	90.06
8	0.51	3.62	93.68
9	0.40	2.87	96.54
10	0.21	1.51	98.05
11	0.18	1.29	99.34
12	0.05	0.33	99.67
13	0.03	0.22	99.89
14	0.02	0.11	100.00

Table 4.46 - Eigenvalues and % of Inertia for the 2nd Criterion

The variables more correlated with the first axis are 45-54, 35-44, 55-64, 25-34 and Primary. Unemployment with Secondary to Non-Tertiary, Unemployment with Tertiary and Unemployment less than Secondary are, once more, the variables that contribute the

most to the formation of the second axis. In terms of the third axis, Females is the variable more relevant and Lower Secondary has, as expected, the highest negative correlation with the fourth axis. Analysing the fifth axis, Public is the variable with highest coordinates for this axis.

In Table 4.47 and in Figure 4.14 it is possible to see the conclusions described for the first two axes. The correlation circles for the plan 1-3, plan 1-4 and plan 1-5 are presented in Figure C.1 in Annexes.

Variable	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
PP	0.40	-0.24	**0.61**	-0.18	-0.19
P	**0.67**	-0.03	-0.27	0.18	0.40
LS	-0.17	-0.08	-0.08	**-0.88**	0.05
US	0.43	-0.43	0.37	0.24	0.33
T	0.45	0.30	0.35	-0.08	0.51
Pb	-0.37	-0.03	-0.34	-0.35	**0.61**
25-34	**0.82**	0.29	-0.19	-0.12	-0.06
35-44	**0.93**	0.15	-0.21	-0.07	-0.03
45-54	**0.94**	0.10	-0.13	-0.09	-0.10
55-64	**0.89**	0.05	-0.07	-0.18	-0.21
F	0.15	-0.23	**0.82**	-0.05	0.17
ULS	-0.18	**0.68**	0.45	-0.27	-0.12
US-NT	-0.13	**0.95**	0.15	0.05	0.04
UT	-0.11	**0.91**	0.02	0.22	0.11

Table 4.47 - Correlations between the variables and the principal components

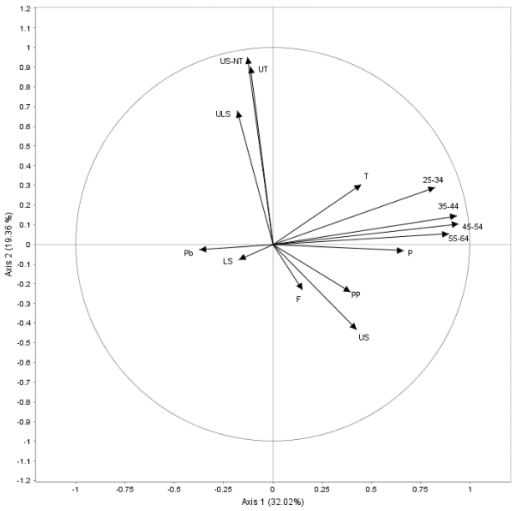

Figure 4.14 - Correlation Circle of the variables in the principal plan 1–2 – 2nd Criterion

The trajectories are essential on this type of analysis. It allows the description of the evolution of each variable/individual along time in the image of compromise, the evaluation of the proximities and oppositions between the individuals relative to the variables correlated with a specific axis, the stability of the variables over the years and other relevant aspects.

4.3.3.1 Trajectories of the variables

In this section it is possible to see the trajectories of the variables according to the first criterion, where the trajectories are obtained through the projection of the variables, in this case, in the principal components of 2012. The coordinates of the variables for the first criterion are presented in Annexes (Table D.1 to Table D.5).

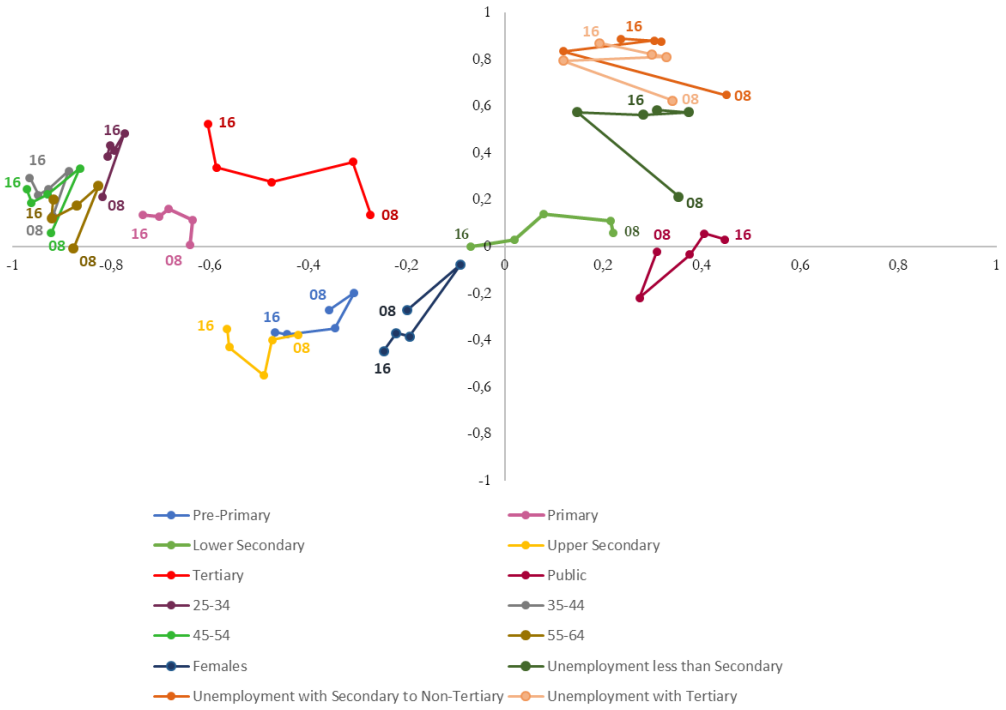

Figure 4.15 – Trajectories of the variables in plan 1-2 according to the 1st Criterion

The first axis explains 33.03% of the total inertia and the variables more negatively correlated are 35-44, 45-54, 55-64, 25-34 (correlations between -0.77 and -0.97) and, although moderately, Primary (correlations between -0.63 and -0.74). These variables are the ones more deviated to the left, as expected. It is possible to say that the first axis is, once more, essentially associated to the Tertiary Education by age groups, and opposes countries with high percentage of population with Tertiary Education by age groups (Iceland, Finland and Norway that have a total contribution of 23.94% to the axis' formation) to countries with low percentage of population with Tertiary Education on these age groups (Croatia, Italy and Romania that have a total contribution of 19.19% to the axis' formation). The trajectories of these variables are similar, where the year 2016 has lower values for the first axis and higher values for the second when compared to the year 2008. In terms of this axis, all the variables present significant variations over the years, except 25-34, and all the variables decrease their values for the coordinates in this axis, except Public and 25-34. That means that, in general, there is a progress in populations having Tertiary Education by age groups.

The second axis that explains 19.71% of the total inertia and has Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary with the highest correlations (correlations between 0.63 and 0.89). It still makes sense that the second axis is

associated to unemployment rates and opposes countries with high unemployment rates for all the levels of education (Greece, Spain and Ireland with a total contribution of 54.26% to the axis' formation) to countries with low unemployment rates (Malta, Romania, Iceland and Austria with a total contribution of 20.95%). The trajectories of the variables in this axis present some differences for the different years in study. Overall, all the variables increase their values for the second axis except Pre-Primary, Lower Secondary and Females. Lower Secondary is the variable more stable over time in terms of this axis.

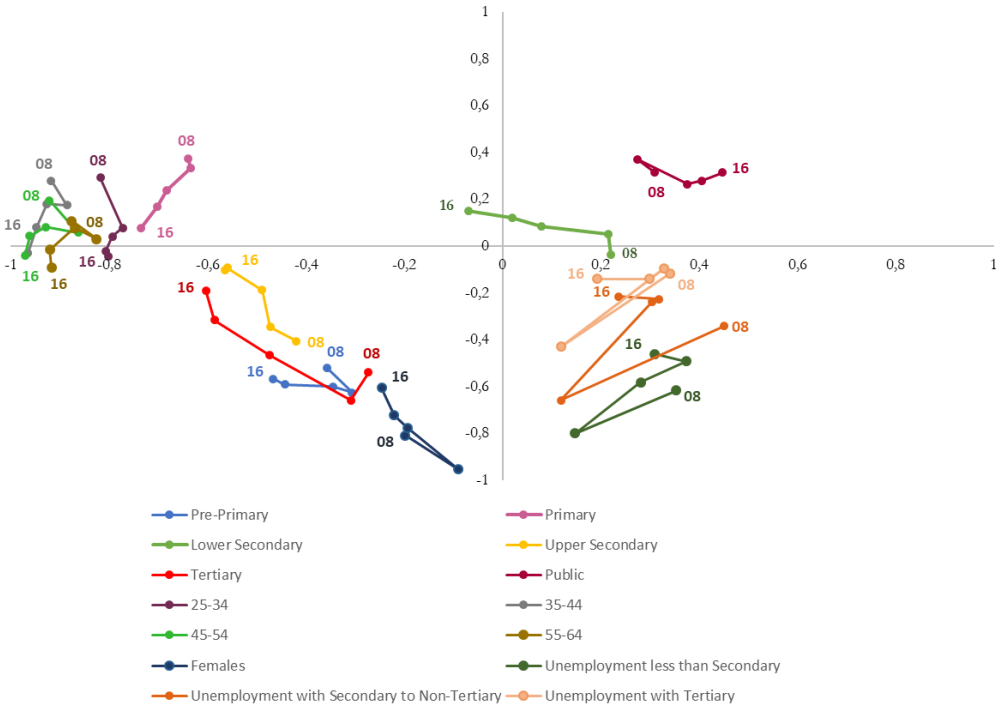

Figure 4.16 - Trajectories of the variables in plan 1-3 according to the 1st Criterion

In terms of the third axis, where 12.87% of total inertia is explained, the variable more correlated with this axis is Females, that has correlations between -0.61 to -0.95. The variables Pre-Primary (correlations between -0.52 and -0.63) and Unemployment less than Secondary (correlations between -0.46 and -0.80) are also negatively correlated, moderately, with this axis. Therefore, it is possible to conclude that the third axis is essentially associated to the Females with Tertiary Education and opposes countries with high percentage of Females in Tertiary Education (Latvia, Slovakia and Estonia that have a total contribution of 25.71% to the axis' formation) to countries with low percentage of Females in Tertiary Education (Switzerland, Ireland and Germany that have a total contribution of

36.82% to the axis' formation). In general, all the variables are broad and present variations for the different years in study.

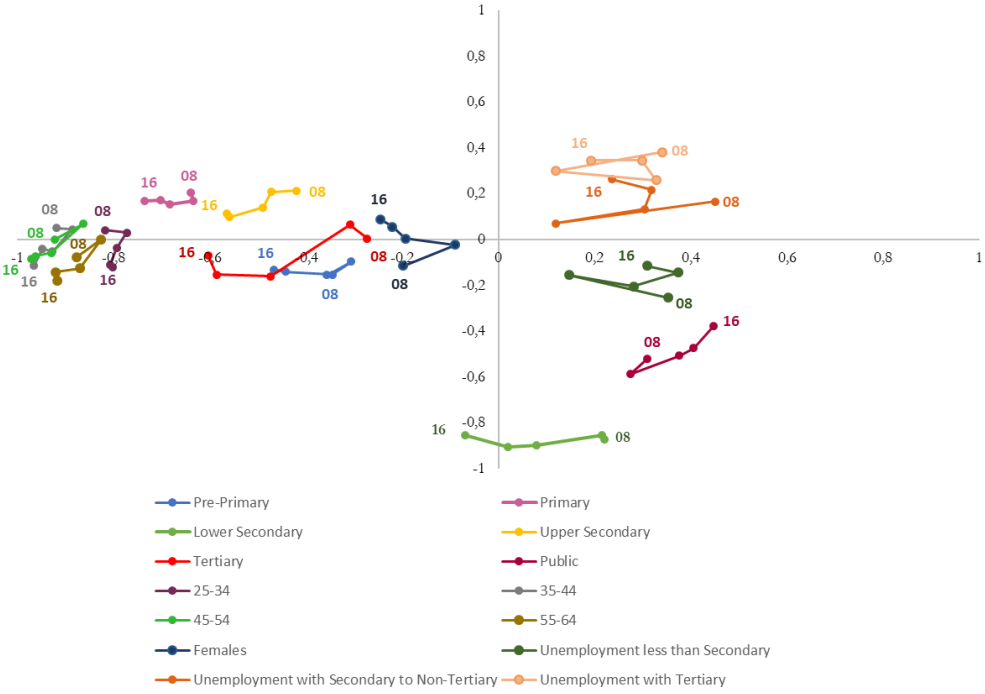

Figure 4.17 - Trajectories of the variables in plan 1-4 according to the 1st Criterion

The fourth axis explains 9.35% of the total inertia and has Lower Secondary with the highest correlations (correlations between -0.85 and -0.91). The fourth axis opposes countries with high percentage of students with education Lower than Secondary (Lithuania and Germany with a total contribution of 50.16% to the axis' formation) to countries with low percentage of students with Lower Secondary Education (Latvia, Cyprus and United Kingdom with a total contribution of 18.18%). In Figure 4.17 it is possible to see that the variable Lower Secondary, as expected, is the closest to the limit of -1 for the fourth axis. In general, the fourth axis do not present many variations in the trajectories of the variables over time, which means that there is some stability in the variables for this axis.

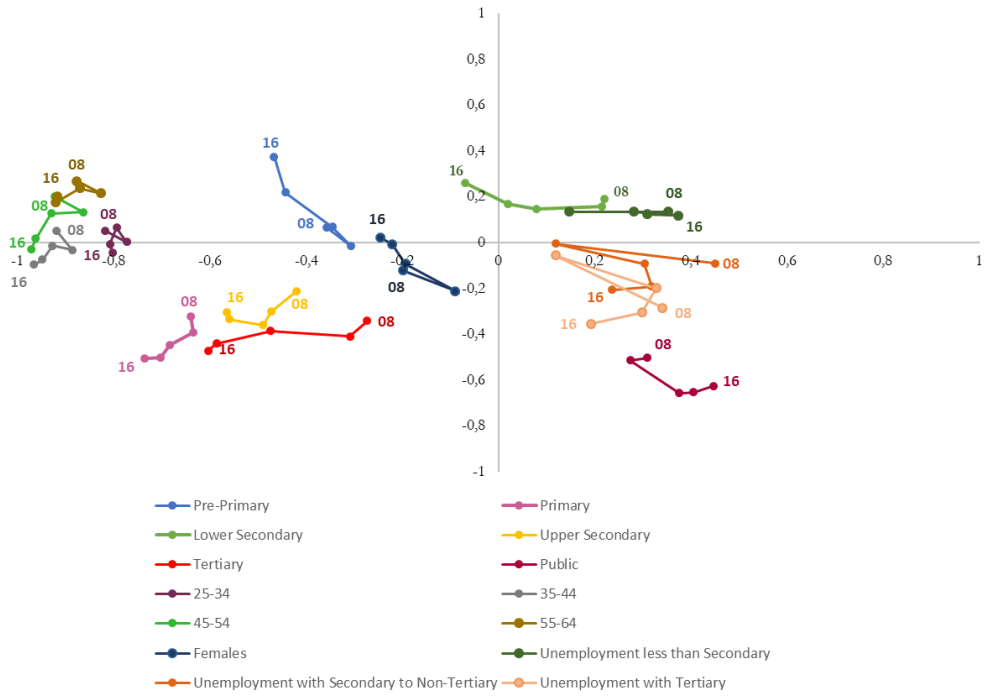

Figure 4.18 – Trajectories of the variables in plan 1-5 according to the 1st Criterion

In Figure 4.18 are represented the trajectories of variables in plan 1-5. The fifth axis explains 7.78% of the total inertia and the variable Public is the one more correlated with this axis (correlations between -0.50 and -0.66). It is possible to say that the fifth axis is, once more, essentially associated to the Tertiary Education in Public System, and opposes countries with high percentage of students with Tertiary Education in Public System (Iceland and Greece that have a total contribution of 28.54% to the axis' formation) to countries with high percentage of students with Tertiary Education in Private System (Estonia, United Kingdom, Germany and Latvia that have a total contribution of 39.10% to the axis' formation). In the fifth axis the variables more stable over time, that is, that do not present significant modifications over time are Unemployment less than Secondary, Lower Secondary and 25-34. The trajectory of variable 25-34 can be considered a closed trajectory. Closed trajectories happen when the points that correspond to the different years are near, that means a certain stability to the years in study. Overall, the variables tend to have descendent trajectories (decrease the values for the coordinates of the fifth axis), except Pre-Primary, Lower Secondary and Females.

4.3.3.2 Trajectories of the countries

Once more, according to the first criterion, the trajectories of the countries are obtained through its projection in the principal axes of 2012. In Figure 4.19 is represented the trajectories of the countries in plan 1-2, composed by the first two axes of the compromise Euclidean image according to the first criterion. The coordinates of the trajectories are presented in Annexes in Tables E.1 to E.5.

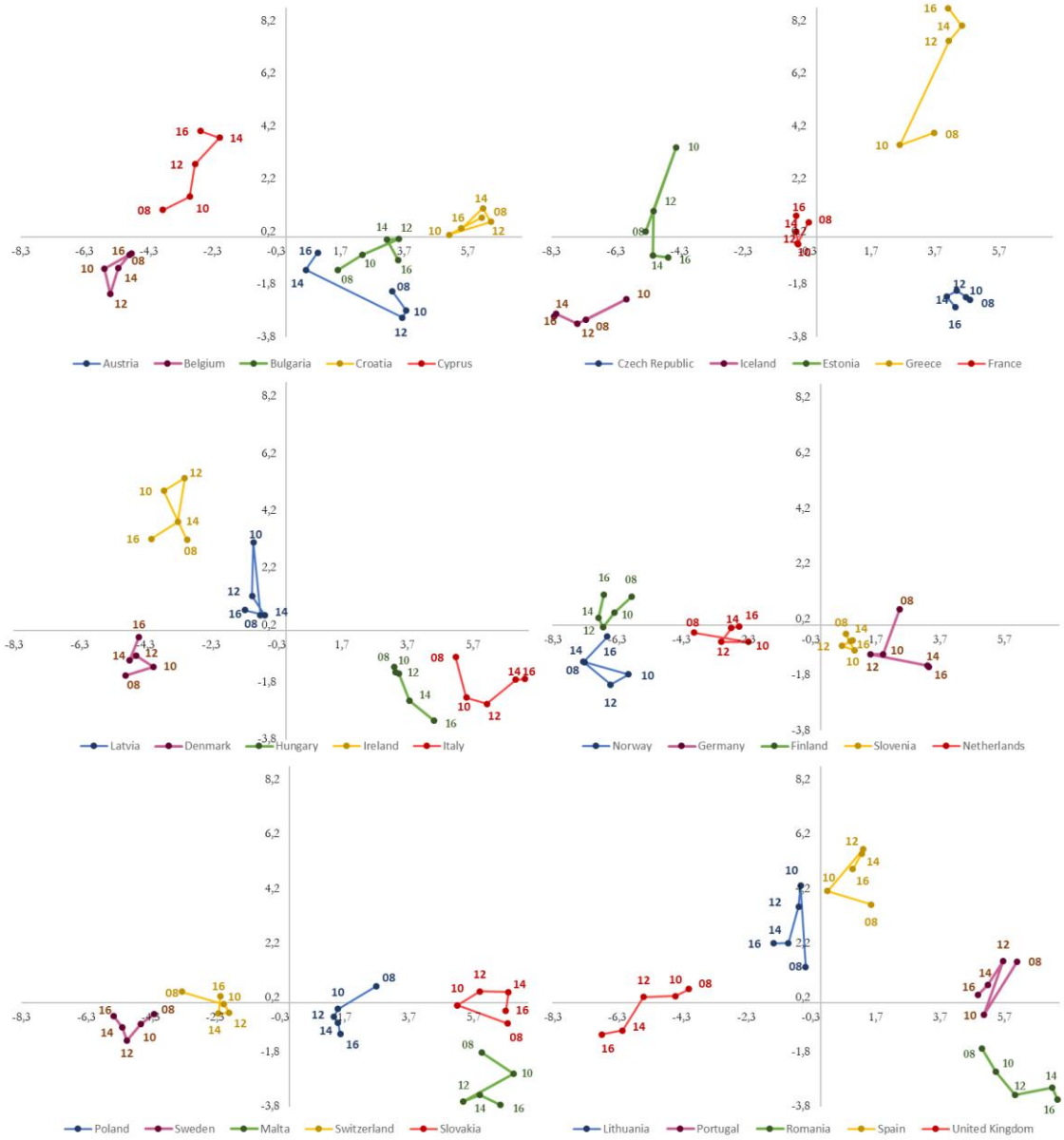

Figure 4.19 - Trajectories of the countries in plan 1-2 according to the 1st Criterion

According to the classifications given to the compromise axes, trajectories with a deviation to the left along axis 1 mean that for these countries there is an increase of the

Tertiary Education by age groups in relation to the average while trajectories with deviation to the right in terms of the first axis mean exactly the opposite. When the deviation of the trajectories is ascending in plan 1-2 means that the unemployment rates in these countries increase in relation to the unemployment rates in all countries. When the deviation on trajectories is descending in plan 1-2 means that these countries have decreases of the unemployment rates in relation to the unemployment rates for all countries.

Analysing the trajectories with the first criterion, Czech Republic, Denmark, France and Slovenia have closed trajectories. That means that these countries do not present significant modifications for the two axes along the years, therefore, there is a stability of these countries during the period in study. Although the trajectories of Croatia, Netherlands and Switzerland present slight variations in terms of the first axis, they can also be considered closed trajectories. With Belgium is the same case, but with higher variations in second axis instead of the first.

When the first axis is analysed, it is possible to see that Austria, Bulgaria, Croatia, Czech Republic, Germany, Greece, Hungary, Italy, Malta, Poland, Portugal, Romania, Slovakia, Slovenia and Spain are countries that have all the trajectory with positive coordinates. According to the Table 4.45, Primary, 25-34, 35-44, 45-54 and 55-64 are negatively correlated with this axis, that means that the countries referred before have values of these variables inferior to the means of these variables. The remaining 15 countries (Belgium, Cyprus, Denmark, Estonia, Finland, France, Iceland, Ireland, Latvia, Lithuania, Netherlands, Norway, Sweden, Switzerland and United Kingdom) have all coordinates negative for the first axis. These countries have higher values than the means for Primary, 25 -34, 35-44, 45-54 and 55-64. The course along the years for Hungary, Italy and Romania shifts in a positive way, that implies a decrease on the values of Primary, 25 -34, 35-44, 45-54 and 55-64 in relation to the means of these variables (when all the countries are considered) over time, that can be also resumed in a decrease of Tertiary Education by age groups in relation to the average. On the other hand, Austria, Iceland, Lithuania, Poland, Portugal, Sweden and United Kingdom, in general, have a left deviation over the first axis, so, between 2008 and 2016, there is an increase of the values of Primary, 25 -34, 35-44, 45-54 and 55-64 in relation to the means of the variables (when all the countries are considered), that also means an increase of Tertiary Education by age groups in relation to the average for these 6 countries.

Croatia, Cyprus, Greece, Ireland, Latvia, Lithuania and Spain have positive coordinates for the second axis, as can be confirmed by the Table E.2 in Annexes. Therefore, these countries have higher values of Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary than the mean. Cyprus has an increase of the values for these variables in relation to the means of these variables (considering all the countries) in all period considered, that means an increase of the unemployment rates in relation to the unemployment rates for all countries. Lithuania has an increase of the values of Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary in relation to the means of these variables (when all the countries are considered) from 2008 to 2010 and then has a decrease. With Greece happens the opposite, there is a decrease of the values of these two variables in relation to the means of these variables (when all the countries are considered) between 2008 and 2010 and then there is an increase. In the case of Ireland and Spain, the values of these variables increase in relation to the means of these variables (when all the countries are considered) until 2012 and then decrease.

Still about the second axis, Germany, Hungary, Malta, Poland, Romania and United Kingdom have descendent trajectories, that means that the coordinates, for this axis, are decreasing from 2008 until 2016, in a general way. It is possible to say that these countries have decreases of the unemployment rates in relation to the unemployment rates for all countries.

Austria is a country where the values of the Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary in relation to the means of these variables (when all the countries are considered) decrease until 2012 and then increase until 2016. In terms of the means of the first axis, there is an increase of the Tertiary Education by age groups in relation to the average, when it is compared 2008 with 2016.

The trajectory of Bulgaria does not have significant variations in terms of the second axis. That can happen due to a stability of the variables correlated with the second component. But there is a decrease of Tertiary Education by age groups in relation to the average along the period, except from 2012 to 2014.

In terms of global panorama, 2008 versus 2016, Portugal increases the values of Primary, 25 -34, 35-44, 45-54 and 55-64 in relation to the means of these variables (when all the countries are considered), that means an increase of the Tertiary Education by age

groups in relation to the average, as referred before. Portugal decreases the values of Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary in relation to the means of these variables (when all the countries are considered), that means a decrease of the unemployment rates in relation to the unemployment rates for all countries. The values of the variables highly negatively correlated with the first axis suffer an increase from 2008 to 2010, then a decrease until 2012, followed by an increase until 2016. On the other hand, there is a decrease of the unemployment rates until 2010, followed by an increase from 2010 to 2012 and then a decrease until the end of the period in study.

In terms of the third axis, that represents the percentage of Females with Tertiary Education, when the deviations of the trajectories are ascending in plan 1-3 means that the percentage of Females with Tertiary Education in the countries decreases in relation to the percentage of Females with Tertiary Education in all countries. This happens due to the negative correlation between the variable Females and the third axis. When the deviations on trajectories are descending in plan 1-3 means that the countries have increases of the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries.

It is possible to see in Figure 4.20 that Austria, Denmark, Germany, Greece, Ireland, Malta, Netherlands, Norway, Portugal, Switzerland and United Kingdom are countries that have positive coordinates for the third axis for all the period. This means that these countries have lower values of Females with Tertiary Education than the mean (when all the countries are considered). Switzerland is the country with lower values of this variable, and consequently higher coordinates. On the other hand, Belgium, Czech Republic, Estonia, Hungary, Iceland, Latvia, Lithuania, Poland, Slovakia, Slovenia, Spain and Sweden are the countries with negative coordinates since 2008. These countries have higher values than the mean of Females with Tertiary Education. It is possible to conclude that Latvia and Slovakia are the countries more developed in terms of feminism with the highest levels of Females in Tertiary Education when compared to the other countries.

Belgium, Finland, Slovenia, Sweden and Portugal do not present significant variations in terms of third axis, what means that the percentage of Females with Tertiary Education is quite constant over the years for these countries.

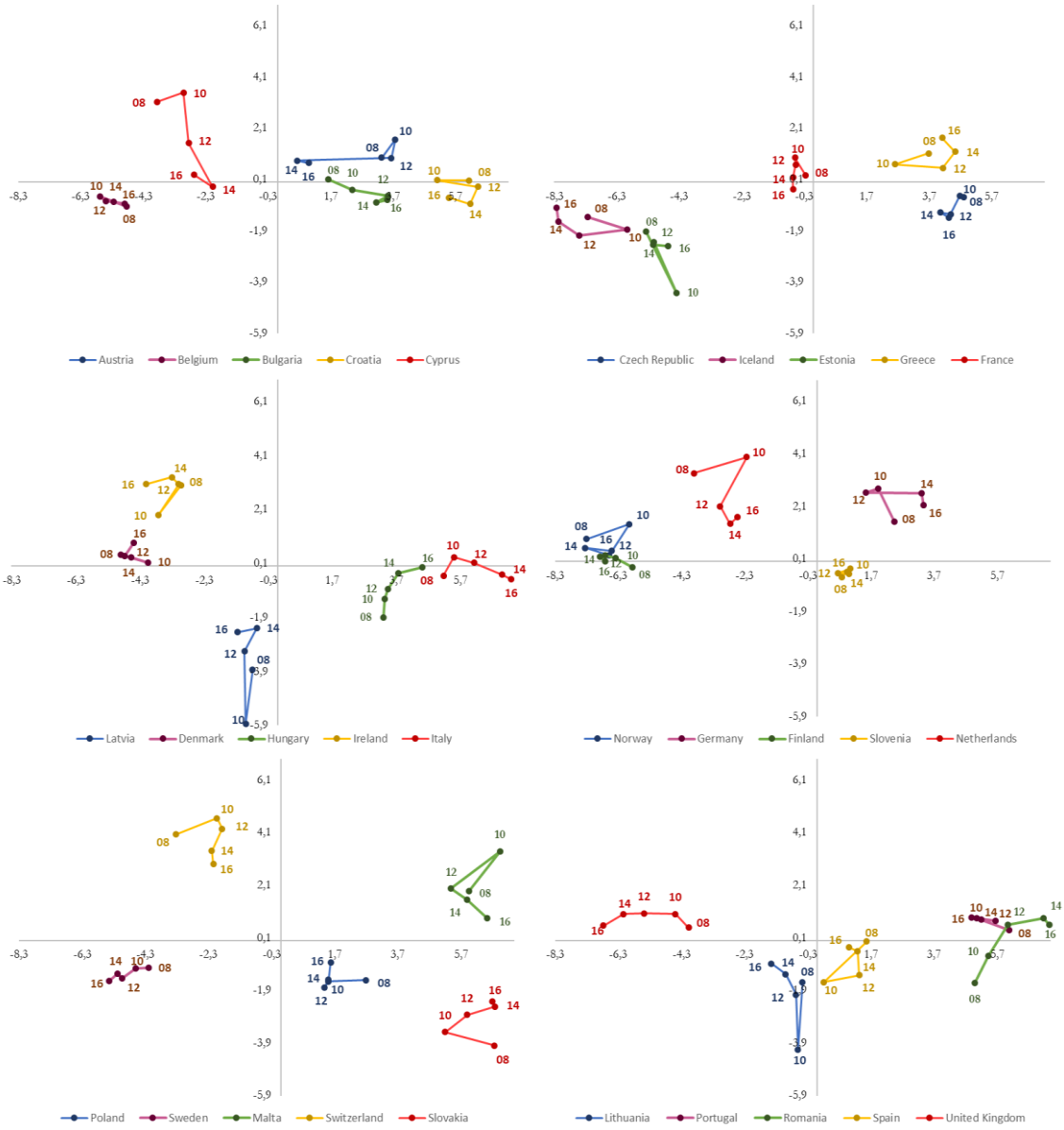

Figure 4.20 - Trajectories of the countries in plan 1-3 according to the 1st Criterion

Hungary is the only country that increases the coordinates of the third axis in all period from -1.899 to -0.042, that means a significant negative evolution in terms of the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries.

The trajectories of Austria, France, Germany, Italy, Malta and Switzerland are similar since these countries decrease their values of Females in relation to the mean of this variable (when all the countries are considered) until 2010, and then increase until the end of the study. Contrarily, Denmark, Lithuania and Spain have exactly the opposite trajectories. They have increases in variable Females in relation to the means of these

variables (when all the countries are considered) from 2008 to 2010 and then decrease these values.

Romania has an ascending trajectory in plan 1-3 until 2014, that represents a decrease of the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries and then an increase of this percentage until 2016, due to a decrease in the coordinates (Table E.3). With Bulgaria happens the opposite.

Croatia, Cyprus and Netherlands, in terms of the third axis, have a decrease of the values of Females in relation to the mean of this variable (when all the countries are considered) until 2010, followed by an increase until 2014 and a decrease until the end of the period, as well as Slovenia (although with minimum variations).

According to the classification attributed before, the fourth axis is represented by Lower Secondary Education. The trajectories for plan 1-4 are in Figure 4.21.

The variable Lower Secondary is the one with highest negative correlation with the fourth axis, that implies that when the deviations of the trajectories are ascending in plan 1-4, there is a decrease in the percentage of students enrolled in Lower Secondary Education in these countries in relation to the students enrolled in Lower Secondary Education in all countries. In the same way, when the deviations on trajectories are descending in plan 1-4 that means that the countries have increases of the percentage of students enrolled in Lower Secondary Education in relation to the average.

Denmark, France, Germany, Hungary, Lithuania, Netherlands, Norway, Slovakia and Sweden have all coordinates negative for the fourth axis. This can be comproved in Table E.4 in Annexes. These countries have higher values of Lower Secondary in relation to the mean of this variable (when all the countries are considered). Lithuania is, as expected, the country located more negatively in terms of the fourth axis.

Contrarily to these previous countries, Belgium, Cyprus, Estonia, Greece, Italy, Latvia, Poland, Portugal, Romania, Slovenia and United Kingdom have all the trajectories above 0, that means that these countries have lower values of Lower Secondary in relation to the mean of this variable (when all the countries are considered). These countries can be considered the ones with higher percentage of students with education higher than secondary in all period.

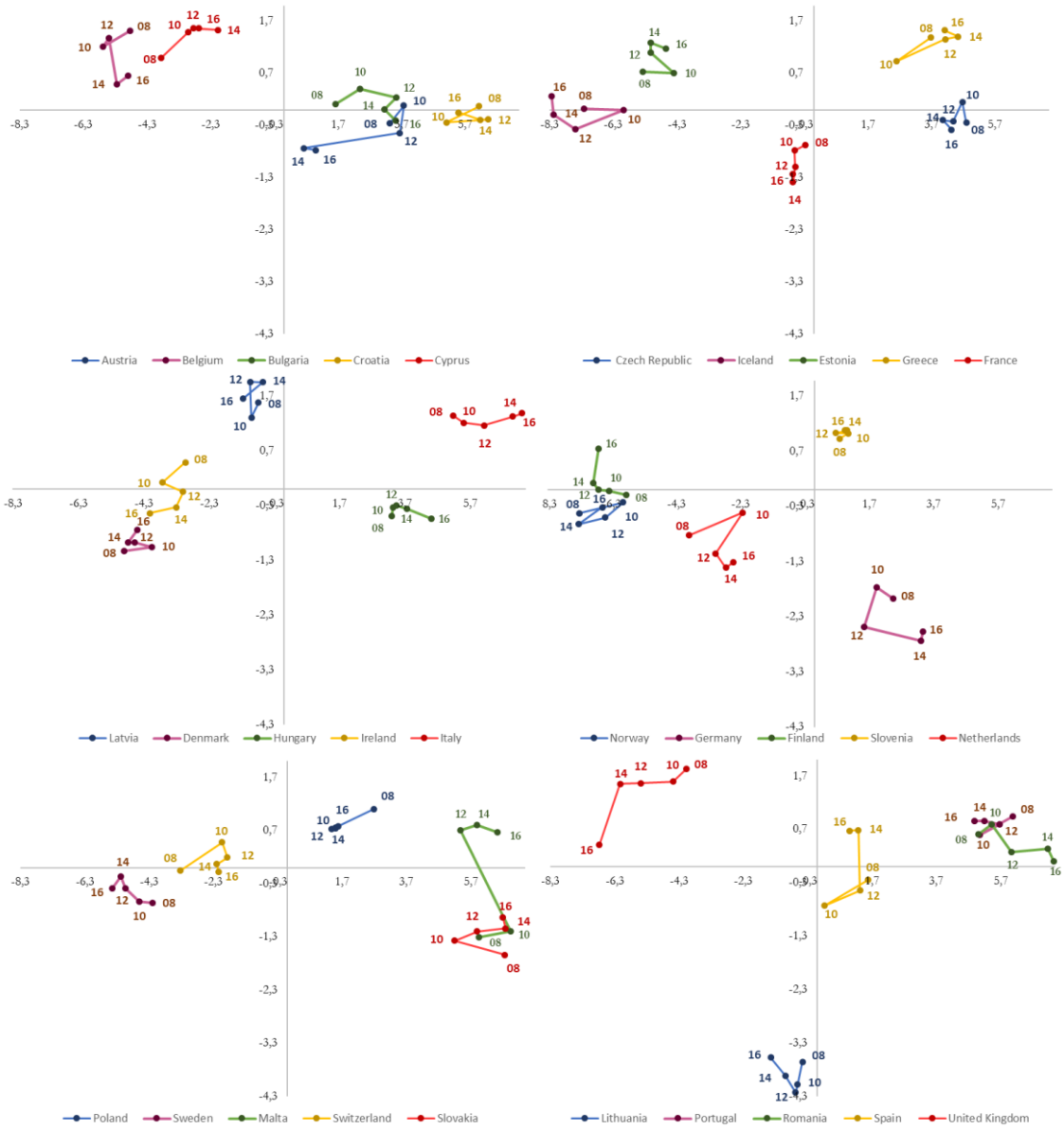

Figure 4.21 - Trajectories of the countries in plan 1-4 according to the 1st Criterion

Ireland and United Kingdom have increases of Lower Secondary Education in relation to the average during all the period. On the other hand, Denmark, Finland and Slovakia are decreasing the values of Lower Secondary Education in relation to the average over the years, that means an increase of the level of education in these countries.

The trajectories of the countries Austria, Bulgaria, Czech Republic, Hungary, Ireland, Latvia, Malta, Romania, Slovenia, Spain, Sweden, Switzerland and United Kingdom have in common a decrease in the coordinates from 2014 to 2016. This decrease, that means an increase of Lower Secondary Education in relation to the average in the last

years, incomes the idea that these countries are increasing their levels of education and are having more students in Tertiary Education.

In the same way, the fifth axis is described by the variable Public that is the one with highest negative correlation with the fifth axis. This means that when the deviations of the trajectories are ascending in plan 1-5, there is a decrease in the percentage of students in Tertiary Education in Public System in relation to the average. When the deviations on trajectories are descending in plan 1-5 that means that the countries have increases in the percentage of students in Tertiary Education in Public System in relation to the average.

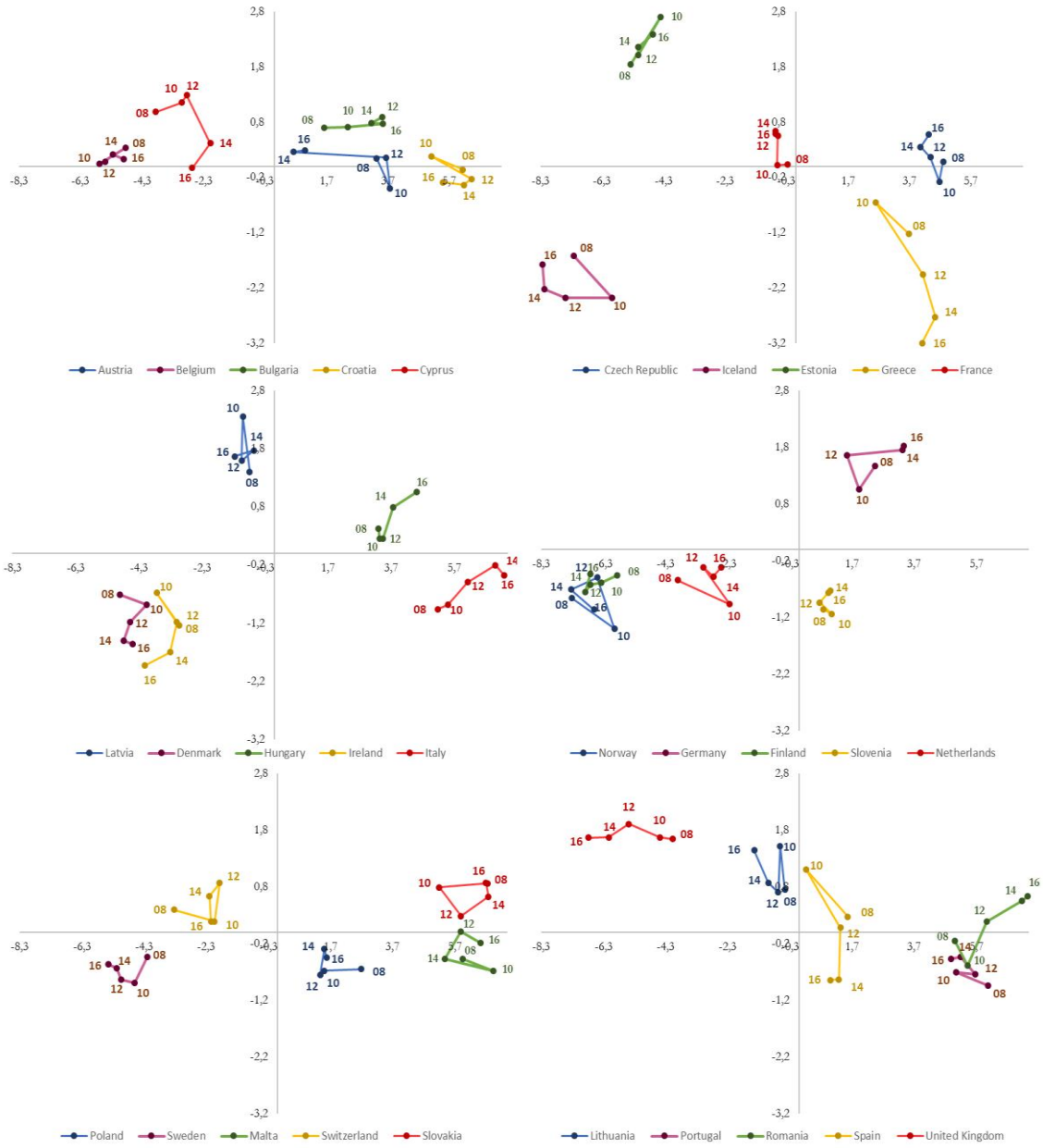

Figure 4.22 - Trajectories of the countries in plan 1-5 according to the 1st Criterion

Denmark, Finland, Greece, Iceland, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Slovenia and Sweden are countries that have negative coordinates for all the years in study. This means that these countries have high values of Public in relation to the mean of this variable (when all the countries are considered). Therefore, these 12 countries have more percentage of students in Tertiary Education in Public System than in Private System. On the other hand, Belgium, Bulgaria, Estonia, France, Germany, Hungary, Latvia, Lithuania, Slovakia, Switzerland and United Kingdom have positive trajectories over the years, that implies lower values of Public in relation to the mean of this variable (when all the countries are considered), and consequently higher percentage of students studying in Private Systems (in Tertiary Education).

There are some countries that increase the percentage of students studying in Public System in relation to the average from 2008 to 2016. The countries with an increase more significant are Cyprus, Denmark, Greece, Ireland and Spain.

Contrarily, Austria, Bulgaria, Czech Republic, Estonia, Finland, France, Germany, Hungary, Italy, Latvia, Lithuania, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia and United Kingdom are all the countries that decrease their values of Public in relation to the mean of this variable (when all the countries are considered) when compared 2008 to 2016. As it is possible to see, these countries are the majority, so, in general, the countries are investing more in Private Systems of education to improve the levels of education. Romania is the country with the highest increase of coordinates.

The trajectories based on the second criterion are obtained by a normed PCA on the overlap of the data tables of the years.

The Figure 4.23 presents the trajectories of the countries according to the second criterion for the plan 1-2. The coordinates of the trajectories for the five axis are described from Table E.6 to Table E.10 in Annexes. In terms of the first compromise axis, according to the second criterion, trajectories with a deviation to the right along the first axis means that, for these countries, there is an increase of the Tertiary Education by age groups in relation to the average, while trajectories with deviation to the left means exactly the opposite. In this criterion the classification given to the second compromise axis is the same that on the first criterion, that means that when the deviations of the trajectories are ascending in plan 1-2, means that the unemployment rates in these countries increases in

relation to the unemployment rates in all countries. When the deviations on trajectories are descending in plan 1-2, means that these countries have decreases of the unemployment rates in relation to the unemployment rates for all countries.

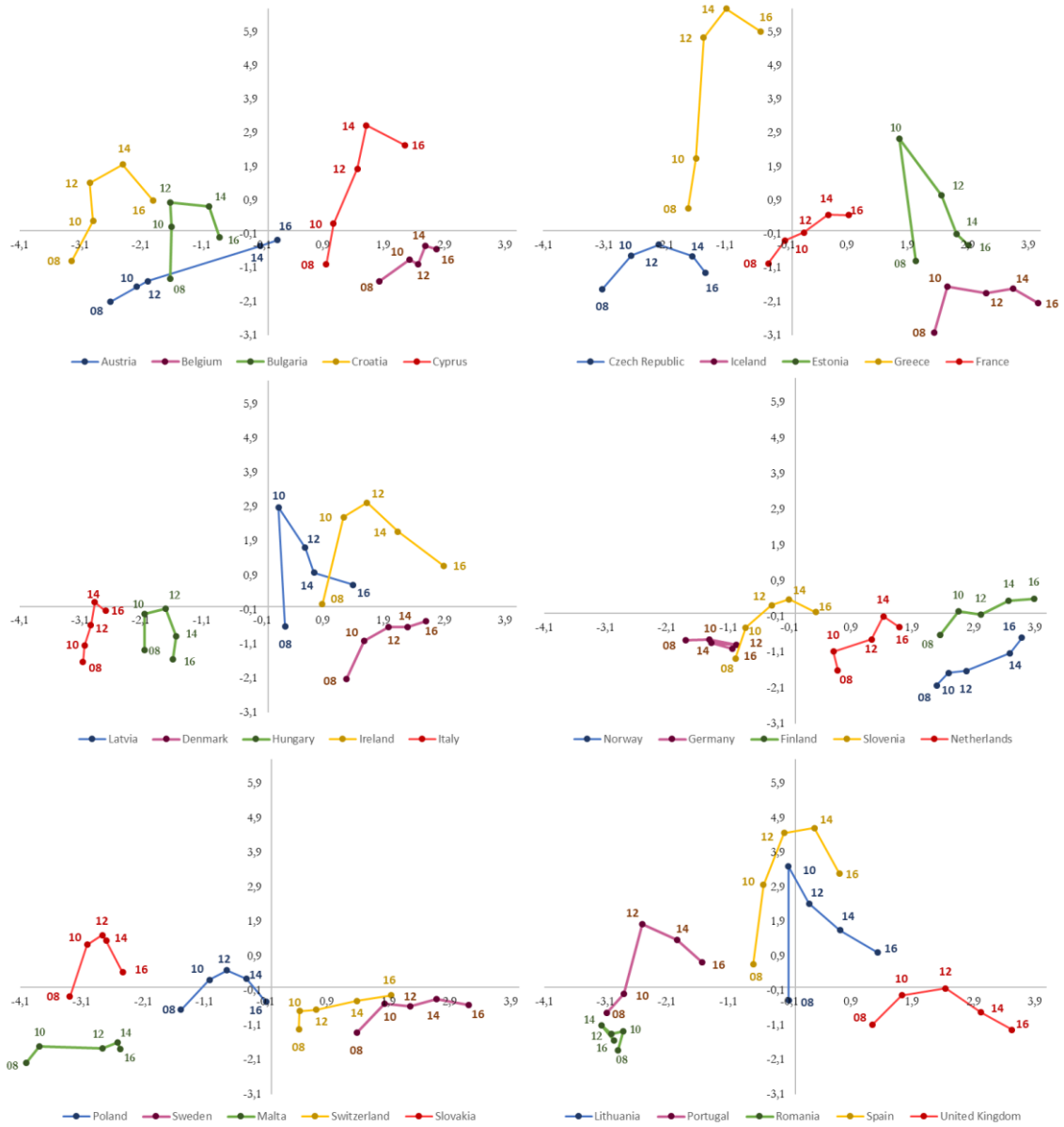

Figure 4.23 - Trajectories of the countries in plan 1-2 according to the 2nd Criterion

The countries Austria, Belgium, Cyprus, Czech Republic, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Malta, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland and United Kingdom present similar trajectories in terms of the first axis. These countries have an increase of the values of variables Primary, 25 -34, 35-44, 45-54 and 55-64, the ones more correlated with the first axis (Table 4.47), in relation to the means of these variables (when all the countries are considered) over the period in

study. This means that there is an increase of the Tertiary Education by age groups in relation to the average. From these countries, Austria, Norway and Switzerland are countries that also have an increase, from 2008 until 2016, in the values of Unemployment less than Secondary, Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary, that are the variables that contribute the most to the formation of the second axis, in relation to the means of these variables (when all the countries are considered). Cyprus, France, Greece, Italy, Slovenia and Spain have an increase of their values of the variables related to the unemployment in relation to the means of these variables (when all the countries are considered) until 2014 and then a decrease. Also, from the countries with an increase of the coordinates of the first axis over time, there are also countries that increase their values of unemployment rates in relation to the means of these variables (when all the countries are considered) since 2008 until 2012 and then decrease until 2016. These countries are Czech Republic, Ireland, Poland, Portugal, Slovakia and United Kingdom and reflect the decrease in the unemployment rates in relation to the unemployment rates for all countries, since 2012. Belgium, Iceland, Malta and Sweden have inconstant trajectories in relation to the second axis. From 2008 to 2010 these countries have an increase of the values of the variables related with unemployment rates in relation to the means followed by a decrease until 2012, an increase until 2014 and, lastly a decrease from 2014 until 2016. Finland has a trajectory similar to the one described before, but instead of a decrease from 2014 to 2016, it suffers an increase.

Germany and Romania are countries with closed trajectories. As presented before, this type of trajectories does not prepresent significant modifications over time. Although this reveals certain stability along the years, in the case of Germany there is more variability in terms of the first axis, and Romania has more variability in terms of the second axis.

Estonia, Latvia and Lithuania have similarities on their second axis, once these 3 countries increase their values of Unemployment less than Secondary, Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary in relation to the means of these variables (when all the countries are considered) from 2008 to 2010 and then decrease until the end of the period studied. On the other hand, for the first axis, the coordinates of these countries decrease between 2008 and 2010 and then increase until 2016.

Croatia and Netherlands present an increase of the variables more correlated with the second axis from 2008 to 2014 and then decrease. Bulgaria and Hungary present also an increase for these variables, but only until 2012. These four countries, in terms of the first axis, when compared to the global panorama, year 2008 versus 2016, have an increase of Tertiary Education by age groups in relation to the average.

The trajectories to the third axis, that, once more, is represented by the percentage of Females with Tertiary Education, are presented in Figure 4.24. For the second criterion, the correlation between Females and this axis is positive, that means that when the deviations of the trajectories are ascending in plan 1-3, the percentage of Females with Tertiary Education in the countries increases in relation to the percentage of Females with Tertiary Education in all countries. On the other hand, when the deviations on trajectories are descending in plan 1-3, the countries have decrease of the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries.

The countries with all coordinates positive for the third axis, that means higher values of Females in relation to the mean of this variable (when all the countries are considered), are Belgium, Czech Republic, Estonia, Iceland, Italy, Latvia, Poland, Slovakia and Sweden. It is interesting to note that all these countries, except Italy, have negative coordinates for this axis in all period, according to the first criterion.

Switzerland is the country with the lowest coordinates for this axis, that means that this country has lower values of Females in relation to the mean of this variable (when all the countries are considered), also as concluded in the first criterion.

According to the Figure 4.24, it is possible to conclude that Denmark, Estonia, Hungary, Iceland, Ireland, Latvia, Lithuania, Slovakia, Slovenia and Sweden have the percentage of Females with Tertiary Education in these countries with an increase in relation to the percentage of Females with Tertiary Education in all countries until 2010 and then have a decrease until the end of the period. Similarly to these trajectories but with the increase of Females until 2012 instead of 2010 are the trajectories of the countries Austria, Belgium, Bulgaria, Czech Republic, Greece, Italy, Poland, Portugal and Spain.

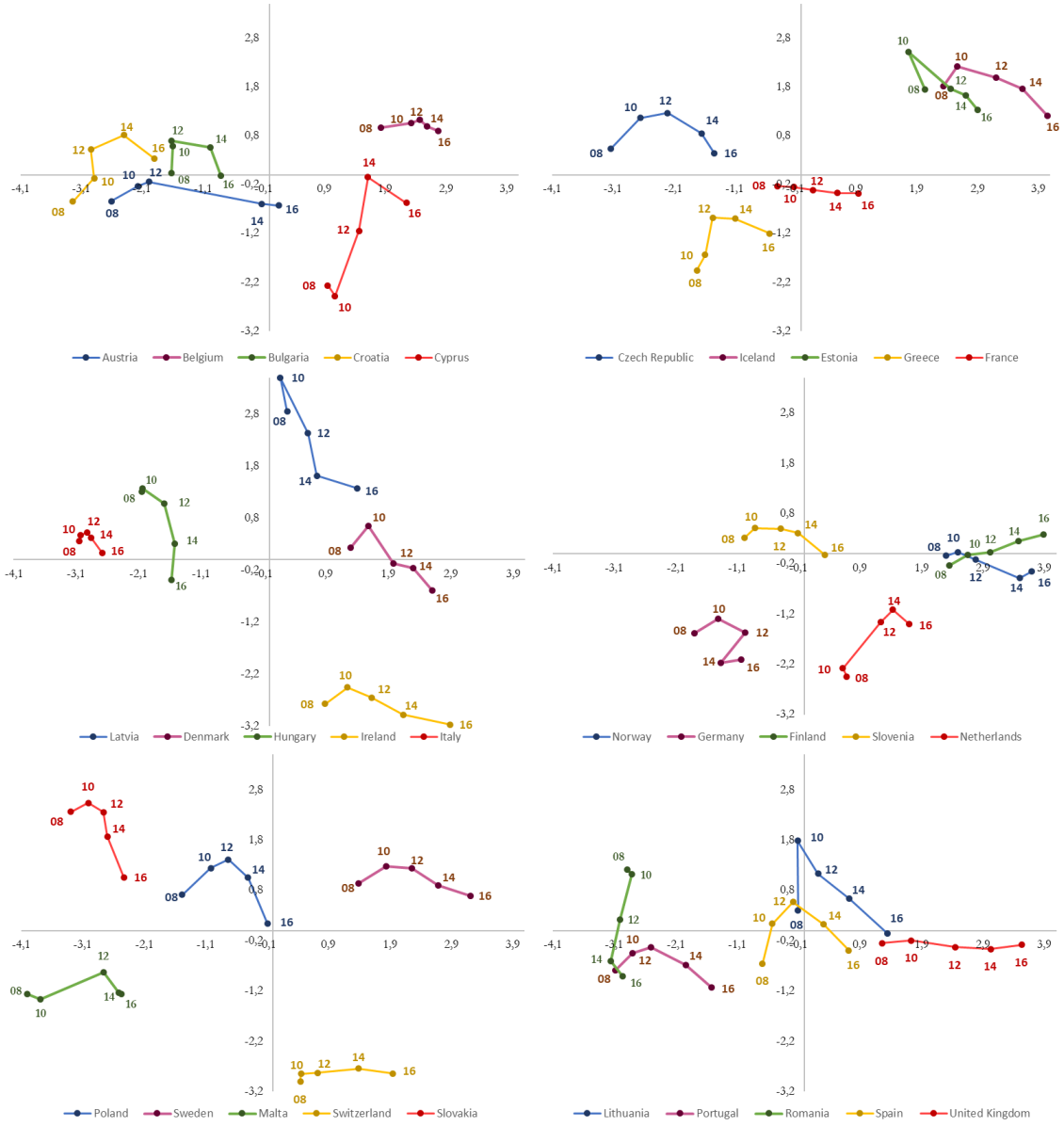

Figure 4.24 - Trajectories of the countries in plan 1-3 according to the 2nd Criterion

Finland is the only country that increases the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries during all the period in study. This country is the opposite of France and Romania that decrease the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries over the years.

The trajectories of Germany, Norway and United Kingdom (although with minimum variations) are ascending in plan 1-3 from 2008 until 2010, that means higher values of Females in relation to the mean of this variable (when all the countries are considered), then decrease until 2014, and then increase again until 2016. The trajectory of Cyprus is the contrary.

Although with different variations, Croatia, Netherlands and Switzerland increase the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries until 2014, finishing the trajectory with a decrease until the end of the period.

In terms of the fourth axis, the Figure 4.25 represents the trajectories in plan 1-4. As in the first criterion, ascending deviations in plan 1-4, means that there is a decrease in the percentage of students enrolled in Lower Secondary Education in relation to the average, and descending deviations means the opposite.

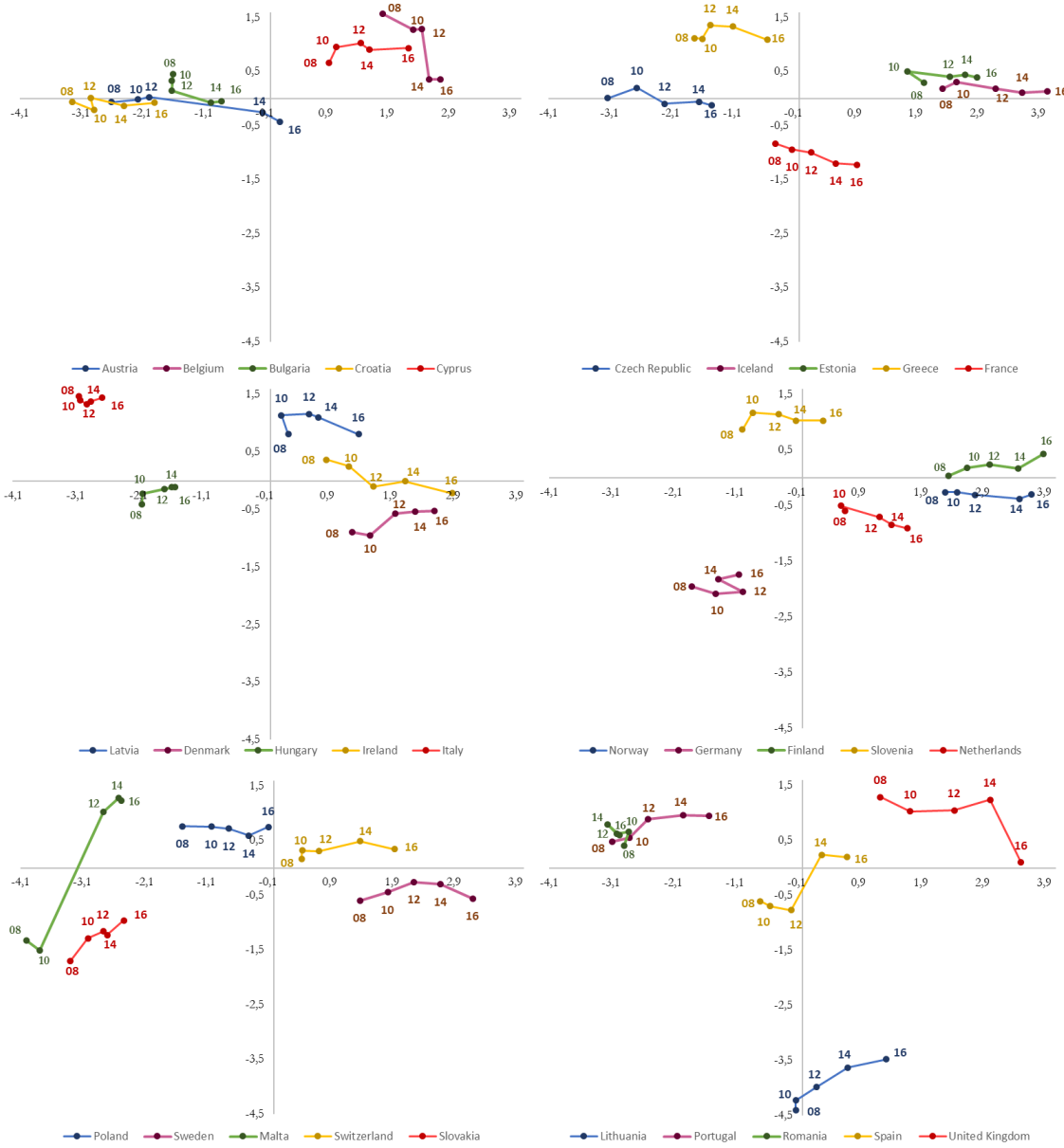

Figure 4.25 - Trajectories of the countries in plan 1-4 according to the 2nd Criterion

Belgium, Cyprus, Estonia, Finland, Greece, Iceland, Italy, Latvia, Poland, Portugal, Romania, Slovenia, Switzerland and United Kingdom have positive coordinates since 2008 until 2016, what means lower values of Lower Secondary in relation to the mean of this variable (when all the countries are considered), and consequently higher values for levels of education higher than secondary. In other words, these countries are more developed in terms of education.

The trajectory of France is descending in plan 1-4 during all the period that means increase of the values of Lower Secondary in relation to the mean of this variable (when all the countries are considered). On the other hand, Lithuania is ascending in plan 1-4. As in the first criterion, this country is the one with higher values of Lower Secondary Education. It is also interesting to note that this country was considered positive outlier of Lower Secondary during the entire period in section 4.2.1.

The countries with higher variations for the fourth axis are Belgium, Lithuania, Malta, Slovakia, Spain and United Kingdom. From these countries, Lithuania, Malta, Slovakia and Spain decrease the values of Lower Secondary in relation to the mean of this variable (when all the countries are considered) when compared 2008 to 2016. In opposition, Belgium and United Kingdom increase.

Netherlands has a decrease of the values of Lower Secondary in relation to the mean of this variable (when all the countries are considered) until 2010 and then increase until the end of the period in study. With Denmark happens the opposite.

In general, the trajectories in plan 1-4 according to the second criterion do not have many significant variations. Croatia, Czech Republic, Estonia, Germany, Greece, Hungary, Iceland, Italy, Norway, Poland, Slovenia, Sweden and Switzerland are the countries more stable over time. These countries have maximum amplitude of 0,35, that means that there are not significant modifications in terms of the Lower Secondary Education over time in these countries.

According to the second criterion the variable Public is the one that contributes the most to the formation of the fifth axis, that implies that when the deviations of the trajectories are ascending in plan 1-5, there is an increase in the percentage of students in Tertiary Education in Public System in relation to the average. On the other hand, when the deviations on trajectories are descending in plan 1-5 that means that the countries have

decreases in the percentage of students in Tertiary Education in Public System in relation to the average.

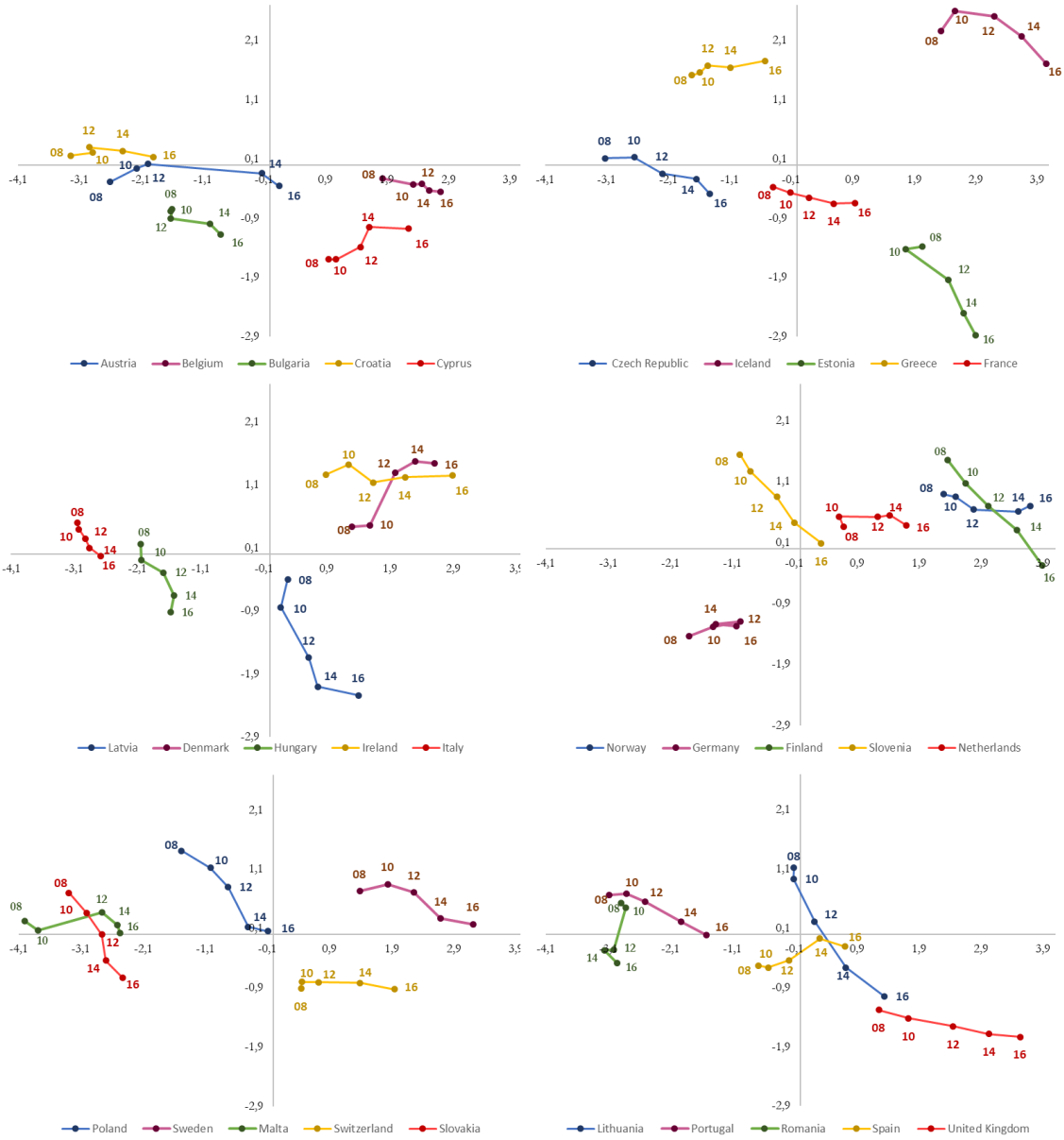

Figure 4.26 - Trajectories of the countries in plan 1-5 according to the 2nd Criterion

Croatia, Denmark, Greece, Iceland, Ireland, Malta, Netherlands, Norway, Poland, Slovenia and Sweden are countries that have all the coordinates positive for variable Public. As it is possible to confirm, these countries, except Croatia and Malta, are the same that the ones described before with negative coordinates for the first criterion. These countries have high values of Public in relation to the mean of this variable (when all the countries

are considered), that means that these countries have more percentage of students enrolled in Tertiary Education in Public Systems instead of Private Systems.

In general, the countries decrease a lot their values of Public for all the period in study, with significant variations. This means that, over years, there is a tendency of higher investment in Private Systems of Tertiary Education over the Public Systems.

Cyprus, Denmark, Germany, Greece, Netherlands and Spain are the only countries that, when compared 2008 to 2016, have an increase of the percentage of students in Tertiary Education in Public System in relation to the average.

The countries with more amplitude, that implies more differences between the coordinates maximum and minimum, are Denmark, Estonia, Finland, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia. From these, only Denmark has an ascendant trajectory, as referred before. The remaining ones have significant modifications in terms of the increase of the percentage of students in Tertiary Education in Private System in relation to the average.

Overall, as can be seen in Figure 4.26, Iceland is the country with higher values of Public in relation to the mean of this variable (when all the countries are considered), while Estonia is the opposite, have lower values of Public in relation to the mean of this variable (when all the countries are considered), which means high percentage of students enrolled in Tertiary Education in Private Systems.

5. Conclusions

The main purpose of this dissertation was to study the evolution of education in European countries between 2008 and 2016 by performing a Double Principal Component Analysis. Through the analysis was possible to obtain a common structure for the different data tables.

The application of the method allowed us to conclude that the years 2008 and 2016 were the years with less similarities between them, what makes sense because are the years most distant and means that there was an evolution between the period studied.

In relation to the analysis of the principal axes and analysing the correlations between the variables and the components retained, in general, the results obtained were quite similar along the period in study.

The variables of percentage of population with ages between 25-34, 35-44, 45-54, 55-64 having Tertiary Education and the percentage of students enrolled in Primary Education were the ones with higher, mostly negative, correlations with the first principal component. It allowed the classification of the first axis as the Tertiary Education by age groups that opposed countries with high percentage of people with Tertiary Education by ages groups to countries with low percentage of people with Tertiary Education by age groups.

The second principal component was highly correlated with unemployment rates for people with Upper Secondary Education to Post-secondary Non-tertiary Education and Tertiary Education, except for 2008. In the analysis, this axis was mainly associated with unemployment rates, which allowed the opposition between countries with high unemployment rates regardless the level of education and countries with low unemployment rates.

Overall, the percentage of Females having Tertiary Education and the percentage of students enrolled in Pre-Primary Education (although with moderate correlations) were the variables with higher significance for the third principal component. These variables obtained a positive correlation with the principal component in 2014 and 2016 and negative one in 2010 and 2012. In general, the third axis was classified as Females with

Tertiary Education opposing countries with high percentage of Females in Tertiary Education to countries with low percentage of Females in Tertiary Education.

Also with negative correlations, the percentage of students enrolled in Lower Secondary Education was the variable that contributed the most to the formation of the fourth principal component during all the period. As expected, this variable represented the fourth axis and opposed countries with high percentage of students that have levels of education lower than Secondary to countries with low percentage of students in the same situation.

Although the analysis of the fifth principal component was made only for 2012 and 2014, the variable percentage of students enrolled in Tertiary Education in Public System obtained the highest negative correlation with the principal component in 2012 and in 2014 the positive one. This variable is what better described the fifth axis, that is the opposition between countries where the Public System is very used for Tertiary Education and countries where Private System is more used. It was possible to conclude that Denmark and Iceland were the countries with more percentage of students enrolled in Tertiary Education in a Public System, while Cyprus, Estonia, Germany, Latvia and United Kingdom had the highest contributions to students enrolled in a Private System.

Concerning the countries, it is interesting to note that Finland, Iceland, Norway and United Kingdom obtained high values for the variables Primary, 25-34, 35-44, 45-54 and 55-64 for all the period in study, according to the first criterion. Through the second criterion, these countries still had high values for these variables. It is possible to confirm that by the trajectories. In first criterion the trajectories of these countries are on the left side (the referred variables are negatively correlated with the first axis) and in second criterion the trajectories are on the right side, because the variables are positive correlated with the first axis. On the other hand, Croatia, Italy, Malta, Portugal, Romania and Slovakia had lower values for these variables all the time, as can also be confirmed through the trajectories of the two criteria.

In terms of the second axis, Greece was the country with the highest values of Unemployment with Secondary to Non-Tertiary and Unemployment with Tertiary in relation to the means of these variables (when all the countries are considered). For the two criteria, the trajectories of this country had a very significant increase over the years. This

means that Greece is increasing the unemployment rates regardless the level of education. On the other hand, Iceland and Malta have lower unemployment rates in relation to the unemployment rates in all countries according the two criteria.

Austria is a country that had an increase of Tertiary Education by age groups in relation to the average for the two criteria, which means a deviation to the left in the first criterion and a deviation to the right in the second one. The trajectories of this country are quite broad being ascending in plan 1-2, that means that the unemployment rates increase significantly in relation to the unemployment rates in all countries.

In the same way, although with different variations, Czech Republic, France, Ireland, Latvia, Lithuania, Poland, Spain and Sweden were countries that had an increase of Tertiary Education by age groups in relation to the average, for the two criteria, when compared the first year in study to the last, that means a decrease of the coordinates for the first axis in the first criterion and an increase in the second one. When analysed in terms of the second axis, Lithuania as well as Spain, had unemployment rates that increased significantly in relation to the unemployment rates in all countries.

Belgium, Croatia, Czech Republic, Denmark, France, Netherlands, Slovenia and Switzerland were countries that, according to first criterion, were considered, in a certain way, as having closed trajectories. Although, in the second criterion, comparing 2008 versus 2016, all these countries had positive evolutions in terms of the second axis, which means that the unemployment rates in these countries increased in relation to the unemployment rates in all countries, as well as in terms of the first axis, that means that had also an increase of the Tertiary Education by age groups in relation to the average.

Portugal, overall and comparing the first year in analysis and the last, had an increase of the values of Primary, 25 -34, 35-44, 45-54 and 55-64 in relation to the means of these variables (when all the countries are considered) according to the two criteria. When the second axis is analysed, Portugal presented a decrease of the unemployment rates in relation to the unemployment rates for all countries, from 2012 to 2016 to the two criteria. Since Portugal was in crisis in 2012, this decrease can be justified due to the high number of people that had to emigrate at that time to find a job, what led to decreases in the unemployment rates since 2012.

Hungary was a country that decreased the percentage of Females with Tertiary Education in relation to the percentage of Females with Tertiary Education for all countries during all the period, according to the first criterion, in other words, the coordinates increased for the third axis. On the other hand, in the second criterion, was France and Romania that decreased their values for Females in relation to the mean of this variable (when all the countries are considered) for all the period. Besides that, the two criteria were coincident in terms of the country with the lowest values on the third axis, that was Switzerland. So, it is possible to conclude that Switzerland is the country that has less women enrolled in Tertiary Education.

It is also interesting to analyse that, in terms of the fourth axis, Lithuania was the country that had the highest negative contribution for this axis, according to the two criteria. This country had high values for Lower Secondary, that means that the level of education in Lithuania is not very high. In other words, in this country, most people have a level of education inferior to secondary. In the individuals' representations in the principal plan [1,4], it is also possible to conclude that Lithuania was always the country with higher negative coordinates in relation to the fourth axis, that justifies the high value for the variable Lower Secondary. Another note is that Lithuania was a positive outlier during all the period in study for the variable Lower Secondary, which is coincident with these conclusions.

To the fifth axis, according to the second criterion, Cyprus, Denmark, Germany, Greece, Netherlands and Spain were the only countries that increased their values of Public in relation to the mean of this variable (when all the countries are considered). All these countries, except Germany and Netherlands, were considered as the countries with the increases more considerable of this variable in the first criterion. This means that there is an increase in the percentage of students in Tertiary Education in Public System in relation to the average for these countries. In the first criterion, Romania was the country with the highest decrease of the values of Public in relation to the mean of this variable (when all the countries are considered), according to the second criterion, Estonia was the country with the lowest value for this variable. However, Iceland obtained high percentage of students in Tertiary Education in Public System in relation to the average for the two criteria, that means that this country does not invest much in Private Systems.

In general, the behaviour of the data tables, when analysed separately, was quite similar over the years. That can be confirmed by the variables that remained more correlated with the principal components over the years and by the stability of some countries that maintained its higher contributions to formation of the axes. It is also possible to conclude that, overall, the two criteria for the trajectories were coincident most of times.

This study enables us to conclude that there is a higher tendency to improve the education systems over the years. Regardless the age, populations are more instructed than ever and with higher levels of education. When compared to 2008, there was an evolution of people's awareness about the need of being instructed. It is also possible to note, that the youngest population are the one with higher levels of education, which means that having Tertiary Education after Secondary Education is even more a decision for the teenagers.

Another conclusion that is possible to take is that, overall, the unemployment rates had a decrease since 2008, which confirms that higher levels of education lead to less unemployment. With more education, populations become wiser and more capable to make decisions that can cause evolution on their countries and, consequently, the authorities are understanding the intensification of the need to invest in education.

Future Work

This dissertation is merely a study that cannot represent the end of a project. There is some analysis and suggestions that could also be made that would improve the results of the analysis and could help more investigations. One interesting additional analysis would be analysing not only European countries, but also countries from different continents. If this analysis was made, certainly the results would present more variations between the countries, due to the different levels of development of the continents. Another interesting idea would be comparing the results obtained by the Double Principal Component Analysis with the results obtained with other methods, such as STATIS or Multiple Factorial Analysis. In this dissertation were mainly used variables related to Tertiary Education, it also would be interesting to explore different indicators to better understand the evolution of education over time.

Bibliography

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Comp Stat, 2*.

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics, 5*(2).

Ambapour, S. (2001). *Statis : une méthode d'analyse conjointe de plusieurs tableaux de données.* Bureau D'Application Des Methodes Statistiques Et Informatiques, Brazzaville.

Babin, B., Hair, J., Anderson, R., & Black, B. (2015). *Multivariate Data Analysis* (7 ed.). Pearson.

Balbi, S., Stawinoga, A., & Triunfo, N. (2012). Text Mining tools for extracting knowledge from Firms Annual Reports. *JADT*.

Big Data Analysis - What is a Tensor. (2017). Retrieved from Big Data Mining & Machine Learning: http://www.big-data.tips/what-is-a-tensor

Bouatta, A., & Bensaid, L. (2011). *Application des méthodes de l'analyse de données sur l'évolution du parc automobile algérien.* Université des Sciences et de la Technologie Houari Boumediene, Argelia.

Bouroche, J. M. (1975). Analyse des donnés ternaires: la double analyse en composantes principales. *Thèse de 3ème cycle.* Université de Paris VI.

Bryant, F., & Yarnold, P. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In *Principal-components analysis and exploratory and confirmatory factor analysis* (pp. 99-136). Washington, DC: American Psychological Association.

Carvalho, F. R. (2013). *Análise Factorial.* FCTUC, Coimbra.

Chamberlain, G. (1982). *Multivariate regression models for panel data. Journal of econometrics.*

Dazy, F., & Le Barzic, J. (1996). *L'Analyse des Données Évolutives, methods et applications.* Éditions Technip.

Dillon, W. R., & Goldstein, M. (1985). *Multivariate Analysis: Methods and Applications.* Wiley, John & Sons, Incorporated.

Escofier, B., & Pagès, J. (1994). *Computational Statistics & Data Analysis* (1 ed., Vol. 18). Elsevier.

Europeia, C. (2014). *A Modernização do Ensino Superior na Europa: Acesso, Retenção e Empregabilidade*. Lisbon: Direção-Geral de Estatísticas da Educação e Ciência.

Figueiredo, A. (2004). Aplicação do método Statis a dados económicos. *Estatística com acaso e necessidade*, pp. 235-247.

Figueiredo, A., Figueiredo, F., Monteiro, N., & Straume , O. (2012). *Restructuring in privatised firms: A Statis approach. Structural Change and Economic Dynamics* (1 ed., Vol. 23). Elsevier.

Franco, G. D. (2016). Multiple correspondence analysis: one only or several techniques? (Springer, Ed.) *Quality & Quantity: International Journal of Methodology, 50*(3).

Fundación BBVA. (2013). Multivariate Statistics. *Multivariate Statistics' Project*.

Gorsuch, R. L. (1983). *Factor Analysis* (Vol. Second Edition). Hillsdale, New Jersey London: Lawrence Erlbaum Associates .

Gudivada, V. N. (2017). *Data Analytics for Intelligent Transportation Systems*. United States: Elsevier.

Guess, A. (2018). *Tensor Methods for Machine Learning*. Retrieved from Dataversity: https://www.dataversity.net/tensor-methods-for-machine-learning/#

Guirao, H. (1998). *Analyse des données : méthodes multivariées (factorielles) sur tableaux longitudinaux*. Université Claude Bernard Lyon I: enssib.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis with readings* (4th ed.). Upper Saddle River: NJ: Prentice-Hall.

Hamilton, L. D. (2014). *Introduction to Principal Component Analysis (PCA)*.

Hsiao, C. (2003). *Analysis of Panel Data* (2 ed.). United Kingdom: University Press, Cambridge.

ISCED. (2011). Source definition. *Tertiary education*. UNESCO Institute for Statistics.

ISCED. (2011). Source definition. *Education;*. UNESCO Institute for Statistics.

Jolliffe, I. T. (1986). *Principal Component Analysis* (1 ed.). New York: Springer.

Jolliffe, I. T. (2002). *Principal Component Analysis* (2 ed.). Springer Science & Business Media.

Jöreskog, K. G. (1978). *An econometric model for multivariate panel data. In Annales de l'INSEE.* Institut national de la statistique et des études économiques.

Kassambara, A. (2017). *Practical Guide to Principal Component Methods in R* (1 ed.). CreateSpace Independent Publishing Platform.

Keita, N. N. (2011). Méthode STATIS. *Tableaux multiples et Données Evolutives, 1.*

Keita, N. N., & Saporta, G. (2011). *Tableaux multiples et données.*

Khattree, R., & Naik, D. N. (2000). *Multivariate Data Reduction and Discrimination with SAS Software.* SAS Publishing.

Kolda, T. G., & Bader, B. W. (2009). Tensor Decompositions and Applications. *SIAM Review, 51*(3).

Kotu, V., & Deshpande, B. (2019). *Time Series Forecasting* (2 ed.). Data Science.

Kroonenberg, P. M. (2008). *Applied Multiway Data Analysis.* John Wiley & Sons.

Lawley, D., & Maxwell, A. E. (1962). Factor Analysis as a Statistical Method. *Journal of the Royal Statistical Society. Series D (The Statistician), 12*, 209-229.

Le Barzic, J.-F., Dazy, F., Lavallard, F., & Saporta, G. (2000). *L'Analyse des Données Évolutives: Méthodes et applications.* Editions Technip.

Lera, L., Pérez, R., & Boquet, A. (2006). El doble análisis en componentes principales para datos categóricos y su aplicación en un estudio de migración. *Revista Colombiana de Estadística, 29*(1).

Lima, C., Tapadas, N., & Masson, S. (2018). *O que é o PCA – Principal Component Analysis e como aplicá-lo a um conjunto de dados*. Retrieved from bi4all: https://www.bi4all.pt/noticias/blog/o-que-e-o-pca-principal-component-analysis-e-como-aplica-lo-a-um-conjunto-de-dados/

Morrison, D. F. (1976). *Multivariate statistical methods* (2 ed.). New York: Mc Graw-Hill.

PORDATA. (2018). Retrieved from https://www.pordata.pt/en/Europe

Porro-Muñoz, D., Duin, R. P., Talavera, I., & Orozco-Alzate, M. (2011). *Signal Processing* (11 ed., Vol. 91). Elsevier.

R Development Core Team. (2006). R: A Language and Environment for Stastitical Computing, R Foundation for Statistical Computing. Vienna, Austria.

Rocha, L. F. (2018). Análise Estatística Multivariada. Secretary of State for Environment and Sustainability - SEMAS/PA.

Roser, M., & Ortiz-Ospina, E. (2018). *Tertiary Education.* Retrieved from Our World in Data.

Sartorio, S. D. (n.d.). Aplicações de técnicas de análise multivariada em experimentos agropecuários usando o software R. *08/07/2008.* Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

Smith, L. I. (2002). *A tutorial on Principal Components Analysis.* Department of Computer Science, University of Otago, New Zealand.

Soares, R. M. (1999). Revista de Estatística. *Factorial Correspondence Analysis: an Application to a Three-Dimensional Contingency Table*. Gabinete de Estudos e Conjuntura do Instituto Nacional Estatística.

Statista. (2017). *Education in Europe - Statistics & Facts.* The Statistics Portal.

STHDA: Statistical Tools for High-throughput Data Analysis. (2017, September). Principal Component Methods in R: Practical Guide. *MFA - Multiple Factor Analysis in R: Essentials*.

Vicini, L., & Souza, A. M. (2005). *Análise multivariada da teoria à prática.* UFSM, CCNE, Santa Maria, RS, Brazil.

Vinzi, V. E., Russolillo, G., Saporta, G., Trinchera, L., & Abdi, H. (2014). *The Multiple Facets of Partial Least Squares and Related Methods.* PLS, Paris, France: Springer.

World Bank Group. (2018). Education.

XLSTAT. (n.d.). *Multiple Factor Analysis (MFA).*

Annexes

Table A.1 – Database – 2008

	Ano 2008													
	Enrolled students as % of resident population by level of Education					Students enrolled in Tertiary Education by sub-system of Education (%)	Population having Tertiary Education as a % of the population by age group				Females in total of enrolled students (%) having Tertiary Education (F)	Unemployment rate by level of Education completed		
	Pre-Primary Education (PP)	Primary Education (P)	Lower secondary Education (LS)	Upper secondary Education (US)	Tertiary Education (T)	Public (Pb)	25-34	35-44	45-54	55-64		Less than primary, primary and lower secondary Education (ULS)	From upper secondary Education to post-secondary non-tertiary Education (US-NT)	Tertiary Education (UT)
AT - Austria	2,7	4,1	4,6	4,7	3,4	84,1	19,2	19,2	17,6	14,7	53,3	8,4	3,6	1,9
BE - Belgium	3,9	6,8	3,1	8,3	3,8	43,1	42,3	35,3	28,7	22,3	55,0	12,4	6,9	3,6
BG - Bulgaria	2,8	3,5	3,5	4,6	3,5	79,2	26,0	23,6	22,4	19,2	55,3	14,6	4,5	2,3
CH - Switzerland	2,0	6,6	3,9	4,0	2,9	82,6	38,5	36,4	31,5	27,3	49,3	6,2	3,4	1,9
CY - Cyprus	2,6	7,2	4,2	4,2	3,3	30,3	48,0	35,2	30,2	19,3	49,0	4,9	3,6	3,0
CZ - Czech Republic	2,8	4,4	4,2	4,7	3,8	86,5	17,7	14,4	14,8	10,6	55,5	19,1	3,7	1,7
DE - Germany	2,9	3,9	6,1	3,6	2,7	90,2	23,9	26,7	25,9	24,4	49,7	15,3	7,2	3,3
DK - Denmark	4,6	7,5	4,5	4,8	4,2	98,2	⊥ 36,4	⊥ 33,9	⊥ 29,8	⊥ 25,7	58,0	⊥ 5,4	⊥ 2,8	⊥ 2,2
EE - Estonia	3,7	5,6	3,6	4,3	5,1	15,8	35,8	34,8	34,3	31,6	61,7	12,1	5,7	2,9
ES - Spain	3,6	6,0	4,3	2,4	3,9	85,9	40,0	32,7	23,3	15,9	54,0	15,3	10,5	6,3
FI - Finland	2,8	6,7	3,8	6,8	5,8	89,3	38,3	43,8	36,6	28,5	54,2	12,5	6,4	3,3
FR - France	Pro 4,0	Pro 6,4	Pro 5,0	Pro 4,2	Pro 3,4	81,9	40,6	30,5	20,0	17,3	55,2	11,1	6,6	4,3
GR - Greece	1,3	5,8	3,1	3,1	5,8	100,0	27,7	25,8	20,7	14,6	50,1	7,5	8,9	6,4
HR - Croatia	2,2	4,2	4,8	4,4	3,3	94,1	19,9	15,6	14,1	15,2	54,6	9,9	9,2	5,0
HU - Hungary	3,2	3,9	4,5	5,4	4,1	84,2	24,1	19,0	17,1	16,2	58,0	18,9	7,1	2,7
IE - Ireland	1,4	10,8	3,9	3,3	4,0	95,3	46,5	38,6	28,2	19,5	54,2	10,7	6,9	3,9
IS - Iceland	3,9	9,4	4,4	7,9	5,2	78,9	32,8	36,2	30,3	23,6	64,4	4,6	1,8	1,6
IT - Italy	2,8	4,9	3,0	4,8	3,4	93,0	19,9	15,2	12,0	9,8	57,4	8,4	6,1	4,5
LT - Lithuania	2,7	4,2	8,6	3,5	6,4	88,5	41,5	29,2	26,5	22,7	59,9	12,5	6,7	2,9
LV - Latvia	3,1	5,4	3,8	4,6	5,9	4,9	28,8	24,8	23,7	21,4	64,4	14,7	7,8	4,1
MT - Malta	s 2,0	s 6,5	s 6,3	s 2,8	s 2,3	100,0	23,1	14,5	8,3	7,2	57,9	8,2	4,2	u 1,3
NL - Netherlands	2,4	7,8	4,7	4,4	3,7	100,0	39,8	32,5	30,5	26,5	51,7	4,6	2,4	1,7
NO - Norway	3,5	9,0	4,0	5,0	4,5	86,0	43,7	38,4	31,5	27,7	60,8	5,9	1,8	1,3
PL - Poland	2,3	6,2	3,9	4,9	5,7	66,8	32,1	18,8	13,0	12,0	57,6	12,1	7,5	3,8
PT - Portugal	2,5	7,3	5,0	4,0	3,6	75,4	23,2	14,5	10,0	7,6	53,5	7,9	7,8	6,8
RO - Romania	s 3,2	s 4,2	s 4,5	s 4,9	s 5,1	61,4	18,8	11,3	10,8	8,9	56,3	7,1	6,0	2,7
SE - Sweden	4,0	7,2	4,3	5,7	4,4	91,4	40,9	32,7	28,4	26,4	60,3	12,8	5,4	3,5
SI - Slovenia	2,2	4,9	3,3	5,1	5,7	88,5	30,0	23,7	19,8	16,1	58,1	6,2	4,4	3,4
SK - Slovakia	2,7	4,2	5,8	5,4	4,3	89,3	18,4	13,9	14,2	10,9	60,3	39,3	8,1	3,5
UK - United Kingdom	s 1,8	s 7,2	s 4,0	s 5,5	s 3,8	0,00	38,6	33,0	30,1	25,5	57,2	10,0	5,5	2,8

Table A.2 – Database – 2010

	Ano 2010													
	Enrolled students as % of resident population by level of Education					Students enrolled in Tertiary Education by sub-system of Education (%)	Population having Tertiary Education as a % of the population by age group				Females in total of enrolled students (%) having Tertiary Education (F)	Unemployment rate by level of Education completed		
	Pre-Primary Education (PP)	Primary Education (P)	Lower secondary Education (LS)	Upper secondary Education (US)	Tertiary Education (T)	Public (Pb)	25-34	35-44	45-54	55-64		Less than primary, primary and lower secondary Education (ULS)	From upper secondary Education to post-secondary non-tertiary Education (US-NT)	Tertiary Education (UT)
AT - Austria	2,9	3,9	4,3	4,6	4,2	83,7	20,7	20,8	18,4	16,3	53,1	9,2	4,4	2,5
BE - Belgium	4,0	6,7	3,1	7,9	4,1	43,5	43,8	39,4	30,9	25,6	55,2	15,3	8,1	4,5
BG - Bulgaria	2,9	3,5	3,1	4,2	3,9	78,0	27,5	24,1	22,4	19,4	55,5	22,7	9,7	4,6
CH - Switzerland	1,9	6,3	3,7	4,0	3,2	82,6	37,4	37,0	32,7	27,6	49,2	8,0	4,8	3,1
CY - Cyprus	2,5	6,6	3,8	4,0	3,9	29,9	48,1	37,5	29,9	22,4	46,1	7,1	6,4	5,7
CZ - Czech Republic	3,0	4,4	3,6	4,5	4,2	84,8	22,6	16,3	15,9	11,5	56,8	25,0	6,9	2,8
DE - Germany	2,9	3,8	6,1	3,4	3,2	87,5	26,0	28,3	26,7	25,4	51,3	14,7	6,8	3,0
DK - Denmark	4,7	7,3	4,5	5,1	4,3	98,2	⊥ 37,6	⊥ 36,8	⊥ 31,2	⊥ 27,8	58,1	⊥ 11,1	⊥ 6,8	⊥ 4,7
EE - Estonia	3,6	5,5	3,2	4,0	5,2	16,2	38,2	33,8	38,6	30,5	60,9	30,5	19,1	9,4
ES - Spain	3,9	6,0	4,3	2,5	4,0	85,2	40,3	35,7	25,3	17,6	53,9	27,1	19,0	11,2
FI - Finland	2,9	6,5	3,6	6,9	5,7	81,2	39,2	45,8	38,9	30,1	53,8	16,1	8,9	4,4
FR - France	Pro 3,9	Pro 6,4	Pro 5,0	Pro 4,0	Pro 3,5	80,3	42,7	33,7	21,7	18,3	55,0	14,5	8,4	5,2
GR - Greece	1,4	5,8	3,2	3,3	5,8	100,0	30,6	25,9	22,2	15,6	49,8	12,7	14,6	10,0
HR - Croatia	2,3	3,9	4,9	4,3	3,5	93,5	25,8	16,7	14,9	17,0	56,3	13,0	12,4	8,4
HU - Hungary	3,3	3,9	4,2	5,4	3,9	83,4	26,1	19,1	18,3	16,3	56,5	25,1	10,6	4,6
IE - Ireland	1,4	11,1	4,0	3,5	4,3	95,5	48,8	43,2	30,4	21,8	52,4	21,9	16,9	8,3
IS - Iceland	4,0	9,3	4,2	8,3	5,7	80,3	36,2	38,8	30,5	22,5	63,4	10,5	7,5	3,6
IT - Italy	2,8	4,8	3,1	4,8	3,3	91,5	20,8	15,8	12,1	10,7	57,6	10,3	7,9	5,7
LT - Lithuania	2,8	4,0	7,8	3,7	6,5	88,2	46,3	32,2	27,5	23,1	59,4	39,4	21,8	7,7
LV - Latvia	3,4	5,4	3,2	4,5	5,4	6,2	34,7	26,4	24,2	22,0	62,7	32,8	21,1	10,6
MT - Malta	s 2,1	s 6,0	s 6,3	s 2,6	s 2,6	100,0	24,3	16,9	10,5	8,0	56,3	9,6	5,5	u 1,7
NL - Netherlands	2,3	7,8	4,6	4,4	3,9	100,0	40,3	32,7	29,8	25,9	51,8	7,2	3,9	2,8
NO - Norway	3,5	8,7	3,9	5,0	4,6	85,8	45,0	40,8	33,5	27,4	60,8	7,3	2,9	1,8
PL - Poland	2,6	5,9	3,5	4,7	5,6	67,1	37,1	23,1	15,0	12,9	59,2	17,5	10,5	5,0
PT - Portugal	2,6	7,1	4,8	4,6	3,6	76,6	25,5	16,2	10,5	8,9	53,3	11,8	11,4	7,0
RO - Romania	s 3,3	s 4,2	s 4,3	s 4,7	s 4,9	62,5	20,7	13,7	10,9	8,5	56,4	5,9	8,0	4,8
SE - Sweden	4,3	7,2	3,9	5,8	4,9	91,3	42,3	36,8	29,3	27,2	59,4	17,0	8,1	4,7
SI - Slovenia	2,4	5,2	2,7	4,7	5,6	86,6	31,3	26,7	20,0	16,3	57,8	11,7	7,5	4,3
SK - Slovakia	2,7	3,9	5,2	5,2	4,3	83,5	24,0	15,9	14,6	12,7	59,7	44,2	14,1	5,8
UK - United Kingdom	s 1,9	s 7,0	s 4,1	s 5,5	s 3,9	0,00	41,6	37,4	32,9	27,5	56,6	13,8	8,2	4,1

Table A.3 – Database – 2012

	Ano 2012													
	Enrolled students as % of resident population by level of Education					Students enrolled in higher Education by sub-system of Education (%)	Population having tertiary Education as a % of the population by age group				Females in total of enrolled students (%)having tertiary Education (F)	Unemployment rate by level of Education completed		
	Pre-Primary Education (PP)	Primary Education (P)	Lower secondary Education (LS)	Upper secondary Education (US)	Tertiary Education (T)	Public (Pb)	25-34	35-44	45-54	55-64		Less than primary, primary and lower secondary Education (ULS)	From upper secondary Education to post-secondary non-tertiary Education (US-NT)	Tertiary Education (UT)
AT - Austria	2,9	3,9	4,1	4,4	4,5	83,2	22,8	21,4	18,4	16,4	53,4	9,8	4,4	2,4
BE - Belgium	4,1	6,7	3,0	7,7	4,3	42,7	43,0	40,2	32,4	25,3	55,5	14,0	7,8	3,9
BG - Bulgaria	3,1	3,5	3,2	3,9	3,9	81,5	27,2	25,0	22,1	21,5	54,6	28,0	11,7	5,8
CH - Switzerland	1,9	6,0	3,6	4,0	3,4	81,9	39,3	39,7	33,7	27,6	49,3	7,8	4,4	3,0
CY - Cyprus	2,6	6,3	3,4	3,8	3,7	37,7	53,6	41,0	32,0	24,1	53,1	13,4	12,8	10,2
CZ - Czech Republic	3,3	4,5	3,5	4,0	4,2	85,9	27,8	18,6	17,6	12,6	57,2	28,5	6,5	2,9
DE - Germany	2,9	3,7	5,9	3,3	3,7	87,2	28,9	29,8	27,9	26,5	50,1	12,2	5,2	2,3
DK - Denmark	3,5	8,4	4,4	5,5	4,9	97,8	⊥ 40,2	⊥ 38,7	⊥ 32,3	⊥ 28,7	57,4	⊥ 11,7	⊥ 6,8	⊥ 4,8
EE - Estonia	4,0	5,6	2,9	3,6	5,1	17,4	40,1	36,8	37,5	36,0	59,2	23,1	10,5	6,1
ES - Spain	4,1	6,2	4,4	2,7	4,2	84,8	40,4	39,0	28,1	19,0	53,6	33,7	24,1	14,9
FI - Finland	3,1	6,4	3,4	6,7	5,7	74,2	39,7	47,5	41,3	31,4	53,7	16,0	8,2	3,9
FR - France	Pro 3,9	Pro 6,4	Pro 5,0	Pro 4,0	Pro 3,5	79,5	42,6	37,5	23,7	19,5	54,8	15,2	9,4	5,4
GR - Greece	1,5	5,8	3,0	3,4	6,0	100,0	34,5	26,5	23,3	19,2	49,1	26,0	27,7	18,4
HR - Croatia	2,4	3,7	4,6	4,4	3,7	92,9	23,6	17,7	16,0	17,0	56,7	18,6	17,3	10,6
HU - Hungary	3,4	3,9	4,0	5,3	3,8	83,0	30,5	22,9	19,4	15,6	55,5	24,8	10,8	4,5
IE - Ireland	1,6	11,3	4,1	3,3	4,2	97,8	49,5	46,4	32,6	25,0	50,9	25,3	18,6	8,1
IS - Iceland	4,1	9,2	4,0	8,2	6,0	82,2	38,4	41,7	34,3	24,9	62,5	10,3	4,8	2,9
IT - Italy	2,8	4,8	3,1	4,7	3,2	90,6	22,5	17,5	12,4	11,5	57,5	13,6	10,0	6,6
LT - Lithuania	3,0	3,7	7,2	3,3	5,9	89,2	48,6	34,7	28,6	25,2	58,4	34,7	16,4	5,6
LV - Latvia	3,8	5,6	2,8	3,9	4,8	7,3	38,7	28,6	27,1	22,0	59,6	26,6	17,6	6,4
MT - Malta	s 2,2	s 5,6	s 3,3	s 4,3	s 2,9	100,0	27,7	21,1	12,8	8,6	56,1	9,5	4,1	u 2,6
NL - Netherlands	3,1	7,6	4,8	4,9	4,7	86,6	41,4	35,5	29,4	26,6	51,4	9,3	5,5	3,4
NO - Norway	3,6	8,4	3,8	4,8	4,7	85,2	45,0	43,6	34,6	29,9	60,1	6,2	2,7	1,8
PL - Poland	3,0	5,7	3,2	4,4	5,3	70,4	40,8	26,4	16,5	12,6	59,9	19,5	11,0	5,6
PT - Portugal	2,6	6,8	4,2	3,9	3,7	79,8	29,0	20,3	13,7	10,9	53,5	16,5	17,7	11,7
RO - Romania	s 3,4	s 4,0	s 4,1	s 4,5	s 3,5	77,1	23,6	15,6	11,9	9,3	54,3	6,7	7,3	5,1
SE - Sweden	4,4	7,5	3,5	5,7	4,8	91,2	43,5	40,0	30,5	28,7	59,7	17,6	7,1	4,4
SI - Slovenia	2,7	5,3	2,6	4,6	5,1	86,4	35,3	30,5	23,0	17,2	57,7	14,8	9,1	6,1
SK - Slovakia	2,8	3,9	4,9	4,7	4,1	81,8	27,0	17,3	16,3	13,7	59,6	44,6	13,5	6,8
UK - United Kingdom	s 2,0	s 7,1	s 3,9	s 5,6	s 3,9	0,00	45,1	42,7	34,7	30,8	56,3	13,9	8,6	4,3

Table A.4 – Database – 2014

	Ano 2014													
	Enrolled students as % of resident population by level of Education					Students enrolled in Tertiary Education by sub-system of Education (%)	Population having Tertiary Education as a % of the population by age group				Females in total of enrolled students (%) having Tertiary Education (F)	Unemployment rate by level of Education completed		
	Pre-Primary Education (PP)	Primary Education (P)	Lower secondary Education (LS)	Upper secondary Education (US)	Tertiary Education (T)	Public (Pb)	25-34	35-44	45-54	55-64		Less than primary, primary and lower secondary Education (ULS)	From upper secondary Education to post-secondary non-tertiary Education (US-NT)	Tertiary Education (UT)
AT - Austria	2,8	3,8	3,9	4,2	4,9	83,5	38,4	33,2	26,5	21,3	53,3	11,4	5,0	4,0
BE - Belgium	4,1	6,9	3,9	6,9	4,4	42,9	44,2	42,2	34,5	26,3	55,9	16,2	8,8	4,7
BG - Bulgaria	3,3	3,6	3,2	4,0	3,9	83,6	31,3	29,5	24,3	22,9	54,3	28,3	10,7	5,1
CH - Switzerland	2,0	5,9	3,2	4,3	3,5	83,1	44,2	42,8	36,2	29,6	49,5	8,7	4,7	3,5
CY - Cyprus	2,7	6,2	3,3	3,6	4,0	41,5	54,0	44,1	31,4	25,6	57,4	19,4	18,2	12,9
CZ - Czech Republic	3,5	4,9	3,5	3,9	4,0	87,7	29,9	21,2	19,6	15,1	57,1	22,1	6,0	2,8
DE - Germany	2,7	3,5	5,7	3,2	3,6	91,7	28,4	28,5	26,2	25,5	47,4	11,8	4,6	2,5
DK - Denmark	3,4	8,3	4,3	5,5	5,3	97,7	⊥ 42,7	⊥ 40,9	⊥ 32,7	⊥ 28,9	57,3	⊥ 10,3	⊥ 5,9	⊥ 4,7
EE - Estonia	5,2	5,9	2,8	3,1	4,6	16,2	40,6	38,7	35,1	35,7	58,7	13,0	8,1	4,8
ES - Spain	3,1	6,4	3,5	3,6	4,3	82,8	41,5	42,8	30,2	21,2	53,3	33,8	24,2	14,8
FI - Finland	3,6	6,4	3,3	6,5	5,6	70,8	40,3	49,9	43,7	34,2	53,6	17,2	9,4	5,1
FR - France	Pro 3,9	Pro 6,3	Pro 5,1	Pro 3,9	Pro 3,6	79,2	44,3	41,3	26,2	21,4	54,5	17,1	10,7	6,3
GR - Greece	1,5	5,8	2,9	3,2	6,2	100,0	38,7	27,1	25,6	21,1	48,7	28,2	30,2	20,0
HR - Croatia	2,6	3,8	4,3	4,4	3,9	92,5	31,9	23,9	15,6	14,8	56,8	25,7	18,7	9,6
HU - Hungary	3,3	4,0	3,9	4,8	3,3	85,2	32,1	24,5	20,3	16,9	55,2	18,5	7,4	3,1
IE - Ireland	1,7	11,5	4,0	3,4	4,4	96,8	50,8	49,4	34,6	24,5	50,5	19,9	14,4	6,8
IS - Iceland	4,3	9,1	3,9	7,5	6,1	80,0	40,8	42,0	36,4	29,0	62,4	7,0	4,3	3,7
IT - Italy	2,7	4,7	3,0	4,6	3,1	90,3	24,2	19,4	12,9	12,3	56,8	16,6	11,9	7,9
LT - Lithuania	3,4	3,7	6,5	2,9	5,1	89,8	52,6	37,6	30,2	27,6	58,1	29,8	13,5	4,2
LV - Latvia	3,9	5,8	2,7	3,3	4,5	7,7	39,4	31,4	26,7	23,0	57,5	24,0	11,7	5,6
MT - Malta	s 2,1	s 5,5	s 2,9	s 4,0	s 2,9	98,8	29,9	24,8	13,3	9,2	54,9	9,0	3,7	u 2,6
NL - Netherlands	3,0	7,3	4,8	4,7	5,0	84,8	44,3	37,6	30,1	26,9	52,3	12,1	7,5	4,1
NO - Norway	3,6	8,3	3,7	4,9	5,1	84,4	50,2	49,2	36,6	31,9	57,8	7,5	3,1	2,2
PL - Poland	3,4	5,7	3,0	3,9	4,6	73,5	42,6	31,9	18,2	13,6	59,4	19,0	10,1	4,7
PT - Portugal	2,6	6,5	3,7	3,7	3,5	83,3	31,4	26,4	16,6	12,8	53,5	15,3	15,3	10,0
RO - Romania	s 2,9	s 4,8	s 4,0	s 4,0	s 2,9	83,8	25,4	17,1	12,0	8,5	53,8	6,7	7,2	5,8
SE - Sweden	4,5	7,8	3,3	5,2	4,4	91,0	46,0	46,1	32,3	29,9	59,4	19,1	7,0	4,4
SI - Slovenia	2,9	5,4	2,6	4,4	4,5	86,9	38,0	34,7	24,3	17,9	58,3	15,4	10,4	6,3
SK - Slovakia	2,9	3,9	4,8	3,8	3,7	83,0	29,8	20,7	15,3	13,9	59,7	41,1	12,6	6,4
UK - United Kingdom	s 2,2	s 7,0	s 3,7	s 6,5	s 3,6	0,00	45,8	45,8	36,7	33,1	56,0	11,3	6,9	3,2

Table A.5 – Database – 2016

	Ano 2016													
	Enrolled students as % of resident population by level of Education					Students enrolled in Tertiary Education by sub-system of Education (%)	Population having Tertiary Education as a % of the population by age group				Females in total of enrolled students (%) having Tertiary Education (F)	Unemployment rate by level of Education completed		
	Pre-Primary Education (PP)	Primary Education (P)	Lower secondary Education (LS)	Upper secondary Education (US)	Tertiary Education (T)	Public (Pb)	25-34	35-44	45-54	55-64		Less than primary, primary and lower secondary Education (ULS)	From upper secondary Education to post-secondary non-tertiary Education (US-NT)	Tertiary Education (UT)
AT - Austria	2,9	3,8	3,9	4,0	4,9	82,3	39,7	35,1	28,0	22,9	53,0	12,7	5,8	3,6
BE - Belgium	4,1	7,0	3,8	6,8	4,5	43,0	44,3	43,2	35,1	27,5	55,7	15,9	8,1	4,2
BG - Bulgaria	3,3	3,7	3,1	3,7	3,7	85,7	32,8	29,7	24,8	23,5	53,9	22,2	6,7	3,4
CH - Switzerland	2,1	6,0	3,1	4,2	3,5	85,1	48,8	46,1	37,9	31,6	49,6	9,4	4,9	3,3
CY - Cyprus	2,8	6,4	3,1	3,4	4,7	36,0	56,2	45,9	32,7	27,2	54,7	15,6	14,3	10,9
CZ - Czech Republic	3,5	5,3	3,5	3,7	3,5	89,4	32,6	23,6	19,6	15,5	57,3	20,5	3,5	1,9
DE - Germany	2,8	3,5	5,5	3,1	3,7	90,6	30,5	29,8	27,1	26,3	48,2	10,1	3,7	2,2
DK - Denmark	3,2	8,2	4,2	5,2	5,5	98,0	⊥ 45,3	⊥ 43,4	⊥ 34,4	⊥ 29,9	56,4	⊥ 9,1	⊥ 5,2	⊥ 4,9
EE - Estonia	5,2	6,3	2,9	3,2	3,9	14,3	41,2	40,2	37,5	36,2	58,7	12,7	7,8	3,9
ES - Spain	2,9	6,5	3,5	3,7	4,2	81,2	41,0	44,1	32,6	23,2	53,3	28,0	19,2	11,7
FI - Finland	3,8	6,5	3,2	6,6	5,4	53,3	40,7	50,4	45,1	36,9	53,3	16,7	9,5	5,8
FR - France	Pro 3,9	Pro 6,4	Pro 5,0	Pro 4,0	Pro 3,7	79,4	44,0	43,5	29,3	22,1	54,4	18,0	10,7	5,7
GR - Greece	1,4	6,0	3,0	3,2	6,5	100,0	41,0	31,3	28,2	20,9	48,6	26,4	26,2	18,0
HR - Croatia	2,7	3,9	4,2	4,2	3,9	92,3	32,8	26,7	17,2	16,2	56,9	17,4	14,6	7,8
HU - Hungary	3,2	4,0	3,9	4,3	3,0	86,9	30,4	26,6	20,6	17,2	54,3	13,2	4,8	1,8
IE - Ireland	1,7	11,6	4,1	3,5	4,6	96,3	51,9	51,9	37,9	28,0	51,5	14,7	9,8	5,0
IS - Iceland	4,0	9,3	3,7	6,9	5,5	78,3	43,4	48,3	39,9	28,9	63,7	4,5	3,1	1,7
IT - Italy	2,6	4,7	2,9	4,7	3,0	89,1	25,6	20,5	14,0	12,4	56,0	15,7	11,1	6,8
LT - Lithuania	3,6	3,8	6,1	2,6	4,7	90,9	54,9	44,6	30,5	30,7	56,9	25,1	10,4	2,9
LV - Latvia	3,8	6,1	2,9	3,1	4,3	7,8	42,1	36,3	28,0	27,0	59,1	20,6	11,3	4,3
MT - Malta	s 2,1	s 5,5	s 2,8	s 3,6	s 2,9	97,3	34,0	22,7	13,0	8,7	55,7	7,6	3,4	u 1,4
NL - Netherlands	3,0	7,0	4,8	4,7	4,9	85,6	45,2	40,9	31,9	27,4	51,9	9,8	6,1	3,5
NO - Norway	3,5	8,4	3,5	4,8	5,3	85,3	49,2	48,8	40,4	33,0	58,2	9,9	4,1	3,0
PL - Poland	3,0	6,5	2,9	3,6	4,2	74,9	43,5	35,3	20,5	13,9	59,0	14,5	6,9	3,3
PT - Portugal	2,5	6,2	3,6	3,8	3,3	83,5	35,0	30,5	18,0	13,2	53,1	12,1	12,3	8,4
RO - Romania	s 2,7	s 4,8	s 3,9	s 3,8	s 2,7	86,9	24,8	20,6	14,0	9,6	53,9	7,6	6,2	3,1
SE - Sweden	4,6	8,3	3,5	5,1	4,3	90,8	47,3	49,3	36,1	30,8	59,4	18,8	5,7	4,0
SI - Slovenia	3,0	5,8	2,6	4,6	3,9	86,3	43,0	36,7	25,3	19,1	57,6	14,5	8,1	6,1
SK - Slovakia	3,0	4,0	4,6	3,6	3,1	86,4	33,4	23,0	15,7	13,9	59,2	31,6	9,2	5,7
UK - United Kingdom	s 2,8	s 7,2	s 5,0	s 6,6	s 3,6	0,00	47,2	48,1	39,2	34,3	56,4	8,3	5,4	2,9

Table B.1 – Coordinates and contributions of individuals of 2nd Step of DPCA - 2008

	Axis 1			Axis 2			Axis 3			Axis 4		
Countries	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
Austria	1.46	1.526	0.267	-0.060	0.006	0.000	-2.100	**7.877**	0.551	0.548	0.773	0.037
Belgium	-2.24	**3.564**	0.308	1.112	1.931	0.076	1.104	2.179	0.075	1.330	**4.556**	0.109
Bulgaria	0.71	0.354	0.101	0.333	0.174	0.023	-0.973	1.691	0.192	0.407	0.426	0.034
Croatia	2.99	**6.385**	0.748	-0.719	0.807	0.043	0.458	0.376	0.018	0.130	0.043	0.001
Cyprus	-2.04	**2.967**	0.295	-2.107	**6.938**	0.315	-0.176	0.055	0.002	0.578	0.859	0.024
Czech Republic	2.27	**3.685**	0.450	0.976	1.489	0.083	-1.855	**6.151**	0.300	0.448	0.518	0.017
Denmark	-2.46	**4.314**	0.494	0.487	0.371	0.019	-1.333	**3.177**	0.145	-0.523	0.703	0.022
Estonia	-2.49	**4.435**	0.397	1.618	**4.092**	0.167	1.319	**3.110**	0.111	-0.116	0.034	0.001
Finland	-2.77	**5.459**	0.512	0.228	0.081	0.003	1.297	3.004	0.112	-0.229	0.135	0.004
France	-0.11	0.008	0.002	-0.461	0.333	0.041	0.313	0.175	0.019	-0.781	1.572	0.117
Germany	1.121	0.896	0.105	-1.190	2.212	0.119	-0.560	0.560	0.026	-1.943	**9.723**	0.316
Greece	1.756	2.200	0.153	-2.410	**9.077**	0.289	2.411	**10.386**	0.289	0.907	2.117	0.041
Hungary	1.690	2.036	0.453	1.505	**3.540**	0.360	-0.238	0.101	0.009	-0.255	0.167	0.010
Iceland	-3.458	**8.528**	0.484	2.357	**8.679**	0.225	-1.338	**3.198**	0.073	0.799	1.644	0.026
Ireland	-1.522	1.653	0.129	-2.880	**12.957**	0.463	1.029	1.893	0.059	-0.189	0.092	0.002
Italy	2.453	**4.292**	0.544	0.068	0.007	0.000	-0.360	0.231	0.012	1.809	**8.427**	0.296
Latvia	-0.130	0.012	0.001	2.394	**8.951**	0.368	2.301	**9.457**	0.340	0.551	0.782	0.020
Lithuania	-0.055	0.002	0.000	0.727	0.826	0.023	0.459	0.376	0.009	-3.926	**39.699**	0.657
Malta	2.724	**5.293**	0.331	-0.972	1.476	0.042	-3.021	**16.305**	0.407	-0.576	0.855	0.015
Netherlands	-2.123	**3.214**	0.389	-1.790	**5.007**	0.276	-1.742	**5.419**	0.262	-0.356	0.326	0.011
Norway	-3.633	**9.415**	0.776	0.338	0.178	0.007	-1.303	**3.035**	0.100	-0.101	0.026	0.001
Poland	1.412	1.422	0.283	0.536	0.448	0.041	1.090	2.121	0.169	0.909	2.129	0.117
Portugal	2.982	**6.342**	0.512	-1.351	2.850	0.105	0.960	1.646	0.053	0.831	1.778	0.040
Romania	2.434	**4.226**	0.486	1.248	2.433	0.128	-0.638	0.727	0.033	1.125	**3.261**	0.104
Slovakia	3.595	**9.219**	0.425	2.612	**10.664**	0.225	0.475	0.404	0.007	-1.798	**8.327**	0.106
Slovenia	0.341	0.083	0.019	0.387	0.234	0.025	0.000	0.000	0.000	1.253	**4.043**	0.258
Spain	0.860	0.527	0.047	-1.463	**3.344**	0.137	2.609	**12.161**	0.436	-1.198	**3.698**	0.092
Sweden	-1.950	2.712	0.495	1.001	1.567	0.131	0.315	0.177	0.013	-0.621	0.992	0.050
Switzerland	-1.882	2.527	0.263	-2.442	**9.317**	0.443	-1.281	2.930	0.122	0.044	0.005	0.000
United Kingdom	-1.947	2.705	0.283	-0.082	0.010	0.001	0.777	1.080	0.045	0.943	2.290	0.066

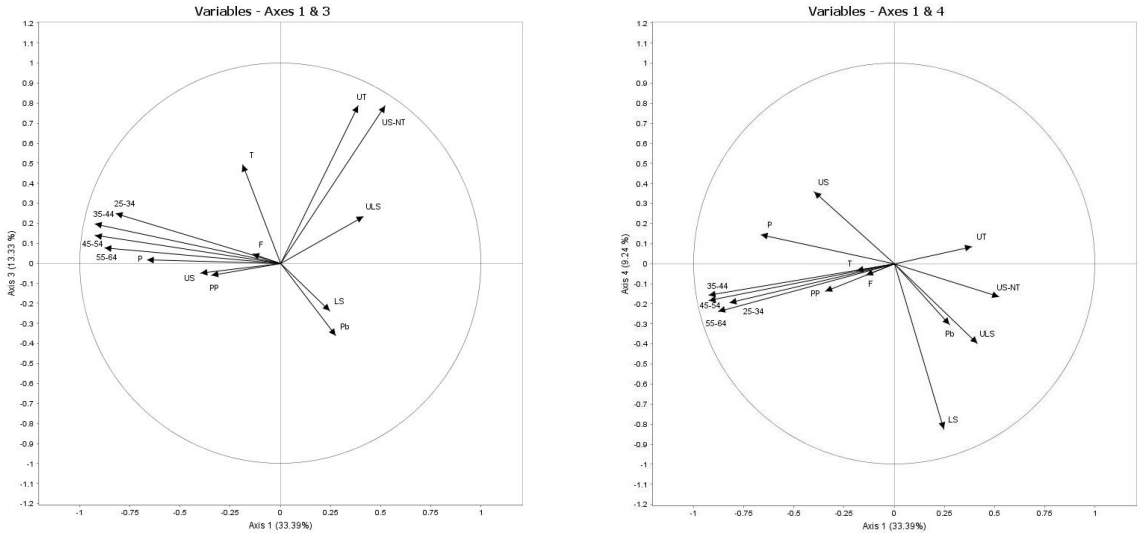

Figure B.1 - Correlation Circles of the variables in the principal plan [1,3] and in the principal plan [1,4] - 2nd Step of DPCA – 2008

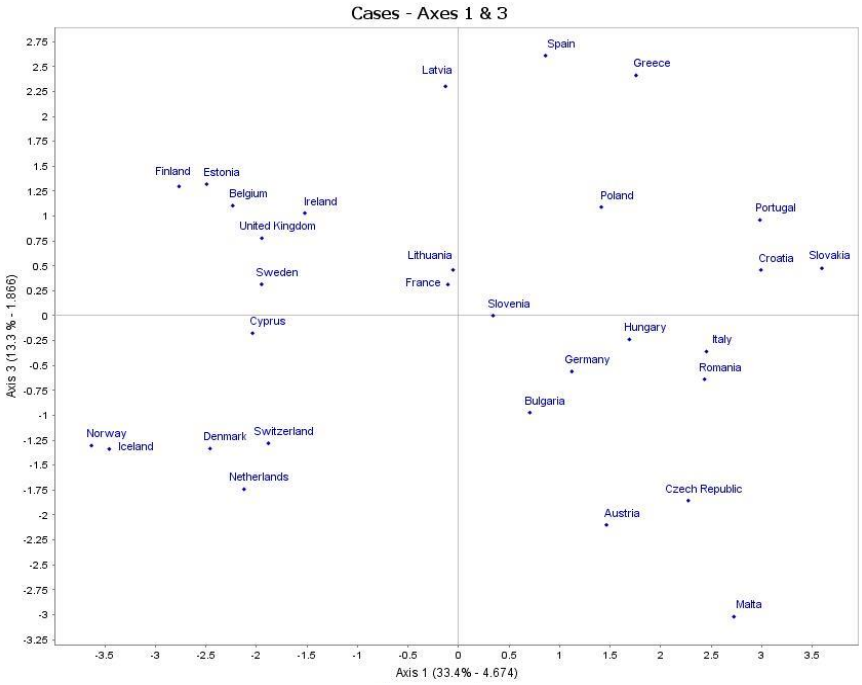

Figure B.2 - Individuals' representation in the principal plan [1,3] - 2nd Step of DPCA - 2008

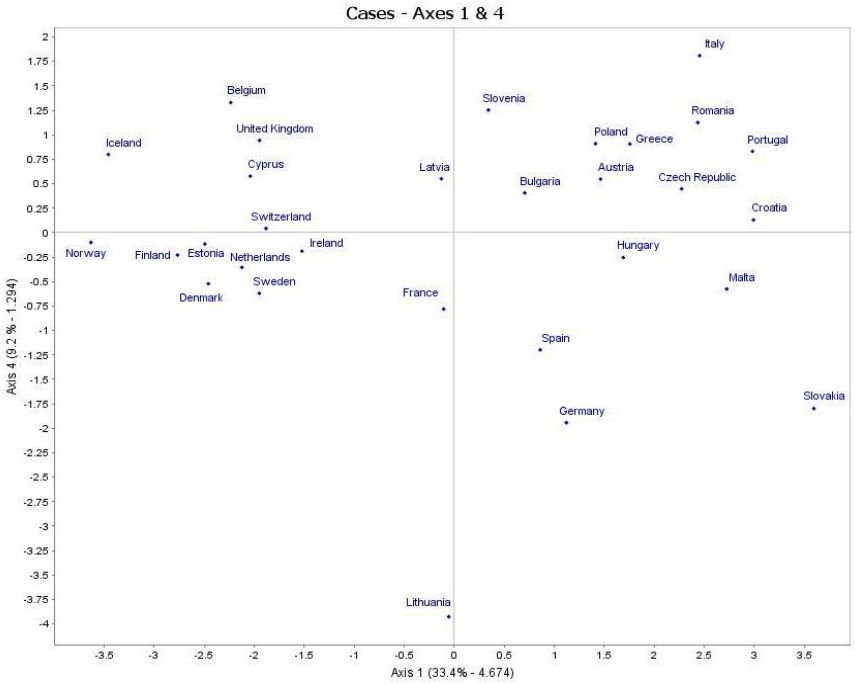

Figure B.3 - Individuals' representation in the principal plan [1,4] - 2nd Step of DPCA - 2008

Table B.2 - Coordinates and contributions of individuals of 2nd Step of DPCA - 2010

	Axis 1			Axis 2			Axis 3			Axis 4		
Countries	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
Austria	-2.167	3.597	0.520	1.535	2.622	0.261	-0.626	0.662	0.043	0.157	0.064	0.003
Belgium	2.729	**5.705**	0.497	0.894	0.889	0.053	-1.635	**4.519**	0.178	0.851	1.892	0.048
Bulgaria	-1.188	1.081	0.301	-0.157	0.027	0.005	-0.313	0.166	0.021	0.656	1.123	0.092
Croatia	-2.478	**4.705**	0.642	-0.589	0.385	0.036	0.559	0.528	0.033	-0.018	0.001	0.000
Cyprus	1.531	1.796	0.150	1.519	2.567	0.147	2.218	**8.316**	0.314	1.500	**5.878**	0.144
Czech Republic	-2.487	**4.739**	0.657	0.148	0.024	0.002	-1.201	**2.436**	0.153	0.219	0.125	0.005
Denmark	1.834	2.576	0.322	1.013	1.142	0.098	-1.063	1.910	0.108	-1.467	**5.617**	0.206
Estonia	2.673	**5.477**	0.317	-3.257	**11.797**	0.470	-0.205	0.071	0.002	1.027	**2.753**	0.047
Finland	3.167	**7.687**	0.667	0.702	0.548	0.033	-0.438	0.323	0.013	0.084	0.019	0.000
France	0.219	0.037	0.010	0.610	0.413	0.077	0.372	0.234	0.029	-1.064	**2.955**	0.234
Germany	-1.160	1.032	0.114	1.325	1.954	0.149	1.106	2.066	0.104	-1.573	**6.458**	0.210
Greece	-0.925	0.656	0.053	-1.009	1.133	0.063	2.297	**8.919**	0.324	1.579	**6.513**	0.153
Hungary	-1.707	2.232	0.512	-0.436	0.212	0.033	-1.090	2.008	0.209	-0.196	0.100	0.007
Iceland	2.708	**5.619**	0.332	0.779	0.675	0.027	-3.048	**15.700**	0.420	-0.688	1.238	0.021
Ireland	2.207	**3.732**	0.208	-0.430	0.206	0.008	3.362	**19.105**	0.484	0.073	0.014	0.000
Italy	-2.882	**6.366**	0.678	0.622	0.430	0.032	-0.916	1.419	0.068	1.066	**2.964**	0.093
Latvia	1.036	0.823	0.045	-4.199	**19.609**	0.747	-0.842	1.198	0.030	1.609	**6.755**	0.110
Lithuania	0.767	0.451	0.018	-4.040	**18.156**	0.493	1.074	1.949	0.035	-3.311	**28.629**	0.331
Malta	-3.831	**11.246**	0.605	1.792	3.571	0.132	0.897	1.359	0.033	-1.518	**6.017**	0.095
Netherlands	0.867	0.576	0.078	2.516	**7.041**	0.655	1.245	**2.618**	0.160	-0.635	1.051	0.042
Norway	2.639	**5.335**	0.451	2.078	**4.805**	0.280	-0.944	1.507	0.058	-0.886	2.051	0.051
Poland	-0.714	0.391	0.087	-0.739	0.607	0.094	-0.815	1.122	0.114	0.758	1.498	0.098
Portugal	-2.542	**4.952**	0.592	0.131	0.019	0.002	0.491	0.408	0.022	0.465	0.565	0.020
Romania	-2.871	**6.317**	0.611	0.168	0.031	0.002	-1.399	**3.305**	0.145	0.784	1.606	0.046
Slovakia	-2.446	**4.586**	0.340	-2.433	**6.587**	0.337	-0.976	1.609	0.054	-0.994	2.579	0.056
Slovenia	-0.585	0.263	0.056	0.309	0.106	0.016	-0.799	1.079	0.104	1.102	**3.168**	0.198
Spain	0.377	0.109	0.009	-2.416	**6.494**	0.378	1.834	**5.686**	0.218	-0.453	0.536	0.013
Sweden	2.198	**3.703**	0.537	0.339	0.128	0.013	-1.414	**3.379**	0.222	-0.953	2.372	0.101
Switzerland	0.843	0.544	0.055	2.570	**7.347**	0.512	1.913	**6.183**	0.284	0.454	0.539	0.016
United Kingdom	2.188	3.669	0.322	0.656	0.478	0.029	0.355	0.213	0.009	1.373	**4.919**	0.127

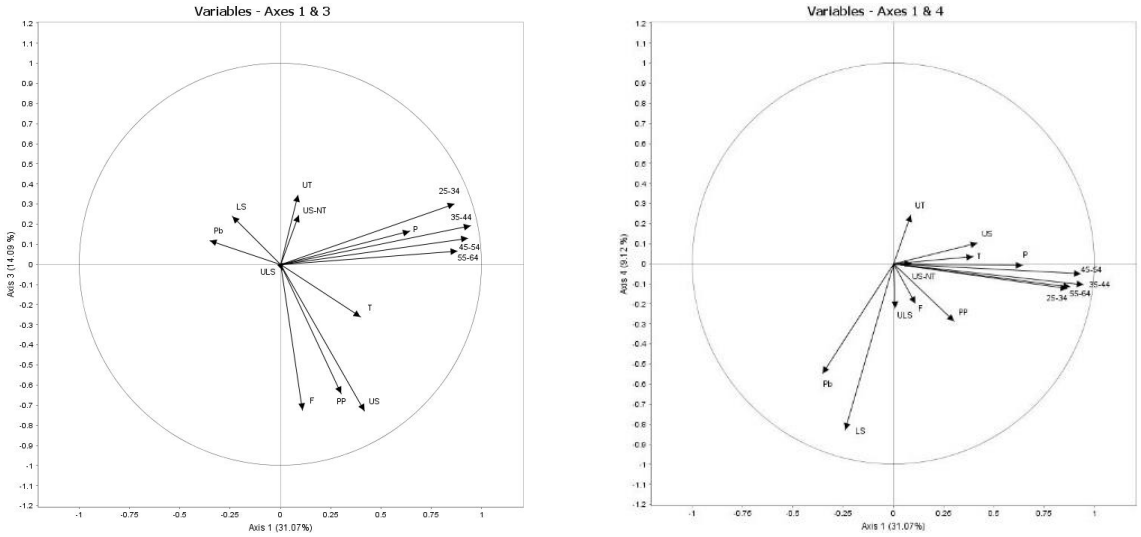

Figure B.4 - Correlation Circles of the variables in the principal plan [1,3] and in the principal plan [1,4] - 2nd Step of DPCA - 2010

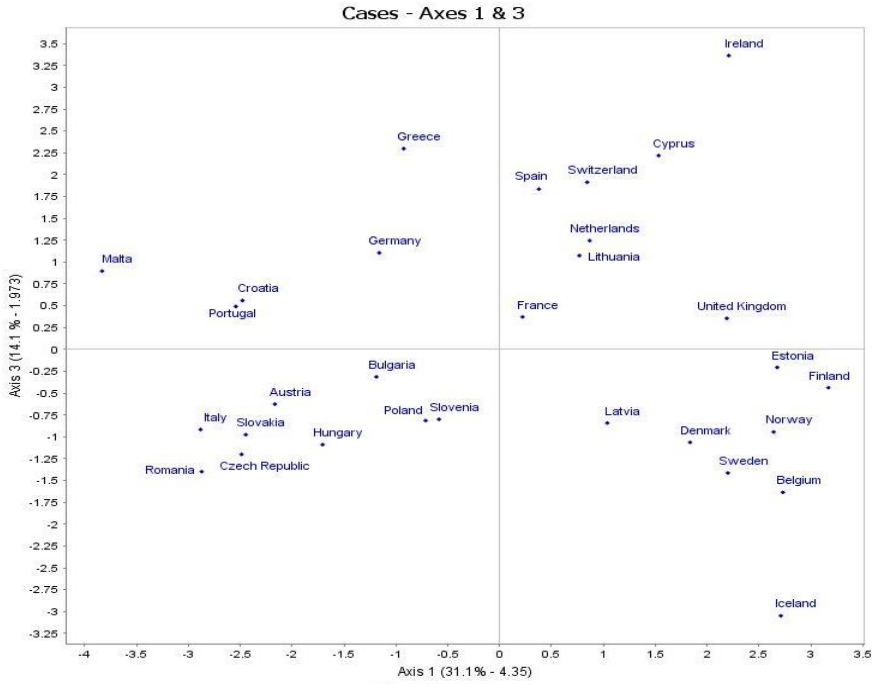

Figure B.5 - Individuals' representation in the principal plan [1,3] - 2nd Step of DPCA - 2010

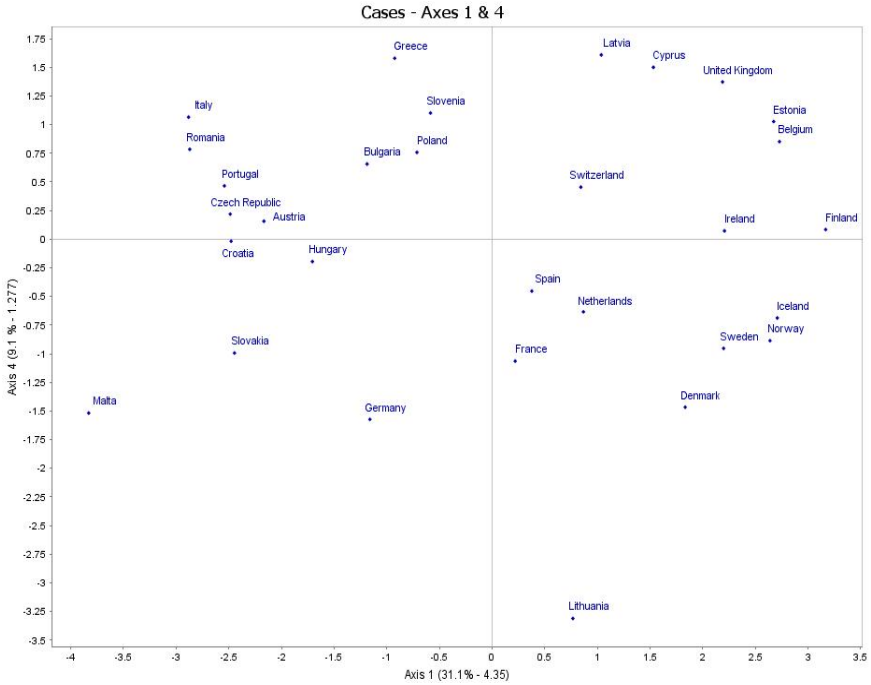

Figure B.6 - Individuals' representation in the principal plan [1,4] - 2nd Step of DPCA – 2010

Table B.3 - Coordinates and contributions of individuals of 2nd Step of DPCA - 2012

	Axis 1			Axis 2			Axis 3			Axis 4			Axis 5		
Countries	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
Austria	1.755	2.219	0.323	-1.905	**4.384**	0.381	0.716	0.947	0.054	-0.400	0.407	0.017	0.157	0.076	0.003
Belgium	-2.641	**5.030**	0.489	-1.351	2.206	0.128	-0.578	0.617	0.023	1.242	**3.928**	0.108	0.080	0.020	0.000
Bulgaria	1.698	2.079	0.482	-0.044	0.002	0.000	-0.421	0.328	0.030	0.216	0.119	0.008	0.769	1.807	0.099
Croatia	3.089	**6.879**	0.763	0.354	0.152	0.010	-0.143	0.038	0.002	-0.165	0.069	0.002	-0.224	0.153	0.004
Cyprus	-1.364	1.341	0.156	1.720	3.574	0.249	1.161	2.493	0.113	1.415	**5.099**	0.168	1.274	**4.963**	0.136
Czech Republic	2.112	3.215	0.496	-1.271	1.950	0.180	-0.970	1.740	0.105	-0.202	0.104	0.005	0.161	0.079	0.003
Denmark	-2.201	**3.493**	0.612	-0.542	0.355	0.037	0.252	0.117	0.008	-0.880	1.973	0.098	-1.168	**4.174**	0.172
Estonia	-2.458	**4.355**	0.353	0.612	0.452	0.022	-1.801	**6.000**	0.189	0.994	2.517	0.058	1.997	**12.205**	0.233
Finland	-3.232	**7.531**	0.686	-0.049	0.003	0.000	0.188	0.065	0.002	0.000	0.000	0.000	-0.619	1.172	0.025
France	-0.277	0.055	0.015	-0.175	0.037	0.006	0.512	0.485	0.050	-0.992	2.508	0.189	0.554	0.940	0.059
Germany	0.741	0.396	0.039	-0.649	0.509	0.030	2.014	**7.504**	0.288	-2.250	**12.887**	0.360	1.641	**8.238**	0.191
Greece	1.991	2.857	0.114	4.628	**25.870**	0.617	0.415	0.319	0.005	1.227	**3.835**	0.043	-1.939	**11.502**	0.108
Hungary	1.650	1.962	0.508	-0.931	1.047	0.162	-0.648	0.777	0.078	-0.264	0.177	0.013	0.248	0.187	0.011
Iceland	-3.604	**9.362**	0.488	-2.052	**5.088**	0.158	-1.619	**4.849**	0.099	-0.333	0.283	0.004	-2.360	**17.037**	0.209
Ireland	-1.487	1.594	0.090	3.327	**13.369**	0.451	2.342	**10.142**	0.223	-0.040	0.004	0.000	-1.166	**4.161**	0.055
Italy	2.941	**6.234**	0.658	-1.591	3.056	0.192	0.099	0.018	0.001	1.052	**2.818**	0.084	-0.488	0.729	0.018
Latvia	-0.497	0.178	0.018	0.759	0.696	0.043	-2.431	**10.934**	0.442	1.771	**7.988**	0.234	1.581	**7.648**	0.187
Lithuania	-0.325	0.076	0.004	2.203	**5.865**	0.193	-1.590	**4.676**	0.101	-3.826	**37.275**	0.583	0.698	1.491	0.019
Malta	2.589	**4.831**	0.408	-2.255	**6.144**	0.310	1.519	**4.269**	0.141	0.630	1.010	0.024	-0.462	0.654	0.013
Netherlands	-1.480	1.578	0.305	-0.379	0.174	0.020	1.622	**4.865**	0.367	-1.043	2.772	0.152	-0.314	0.302	0.014
Norway	-3.126	**7.046**	0.679	-1.337	2.159	0.124	0.312	0.181	0.007	-0.454	0.524	0.014	-0.491	0.738	0.017
Poland	0.659	0.314	0.066	-0.315	0.120	0.015	-1.385	3.550	0.293	0.659	1.106	0.066	-0.744	1.694	0.084
Portugal	2.722	**5.341**	0.622	0.959	1.112	0.077	0.568	0.597	0.027	0.717	1.309	0.043	-0.729	1.626	0.045
Romania	2.902	**6.072**	0.566	-2.101	**5.330**	0.296	0.460	0.391	0.014	0.252	0.162	0.004	0.186	0.106	0.002
Slovakia	2.834	**5.792**	0.449	0.269	0.087	0.004	-2.179	**8.779**	0.266	-1.087	**3.010**	0.066	0.283	0.244	0.004
Slovenia	0.319	0.073	0.024	-0.463	0.259	0.051	-0.341	0.215	0.028	0.932	2.210	0.207	-0.931	2.654	0.207
Spain	0.640	0.296	0.021	3.526	**15.016**	0.637	-1.017	1.912	0.053	-0.401	0.409	0.008	0.079	0.019	0.000
Sweden	-2.419	**4.217**	0.562	-0.865	0.904	0.072	-1.116	2.305	0.120	-0.358	0.326	0.012	-0.829	2.101	0.066
Switzerland	-0.898	0.581	0.060	-0.220	0.059	0.004	3.265	**19.718**	0.798	0.172	0.075	0.002	0.861	2.266	0.055
United Kingdom	-2.634	**5.002**	0.392	0.136	0.022	0.001	0.794	1.167	0.036	1.415	**5.095**	0.113	1.897	**11.011**	0.203

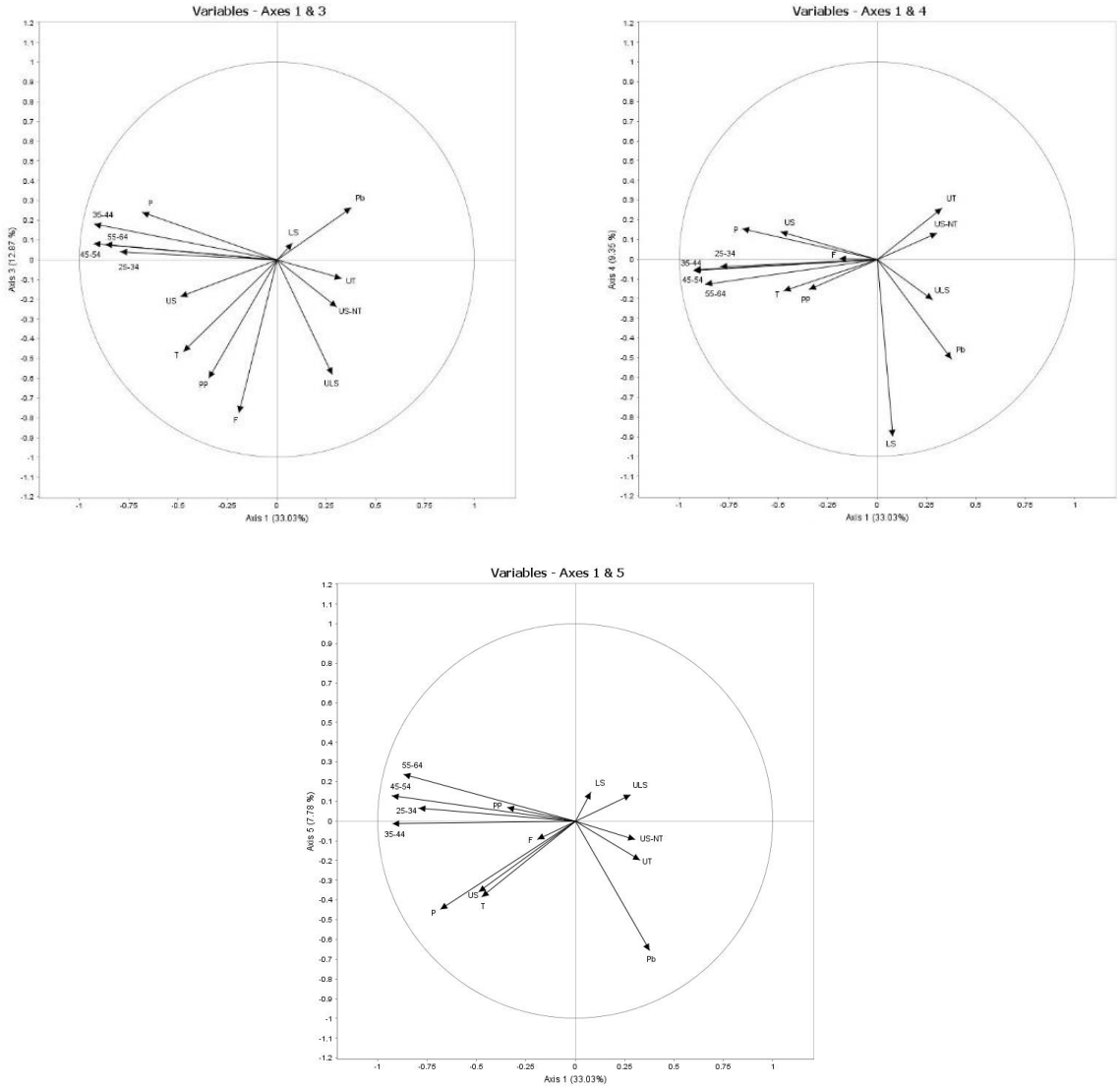

Figure B.7 - Correlation Circles of the variables in the principal plan [1,3], in the principal plan [1,4] and in the principal plan [1,5] - 2nd Step of DPCA – 2012

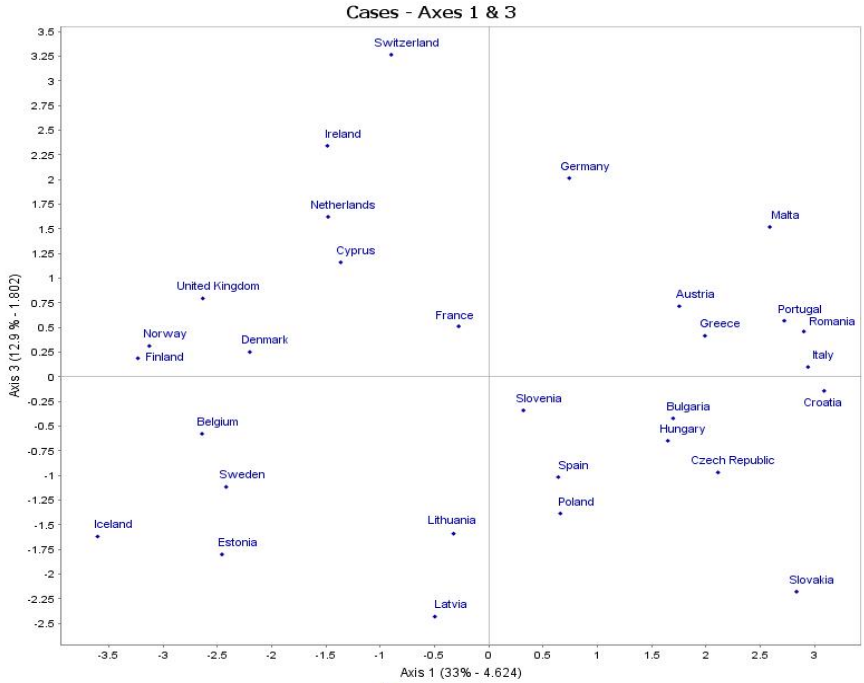

Figure B.8 - Individuals' representation in the principal plan [1,3] - 2nd Step of DPCA – 2012

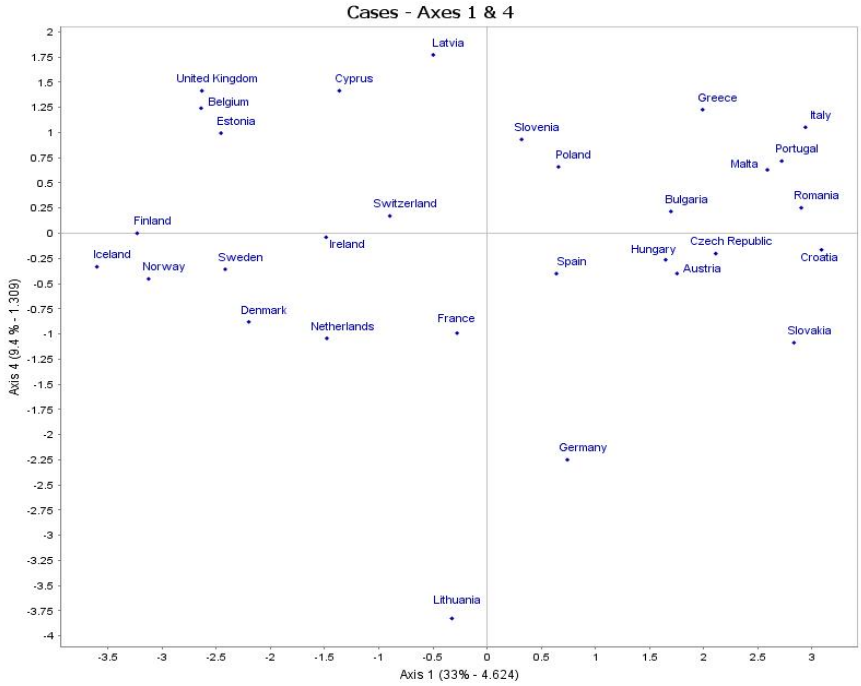

Figure B.9 - Individuals' representation in the principal plan [1,4] - 2nd Step of DPCA – 2012

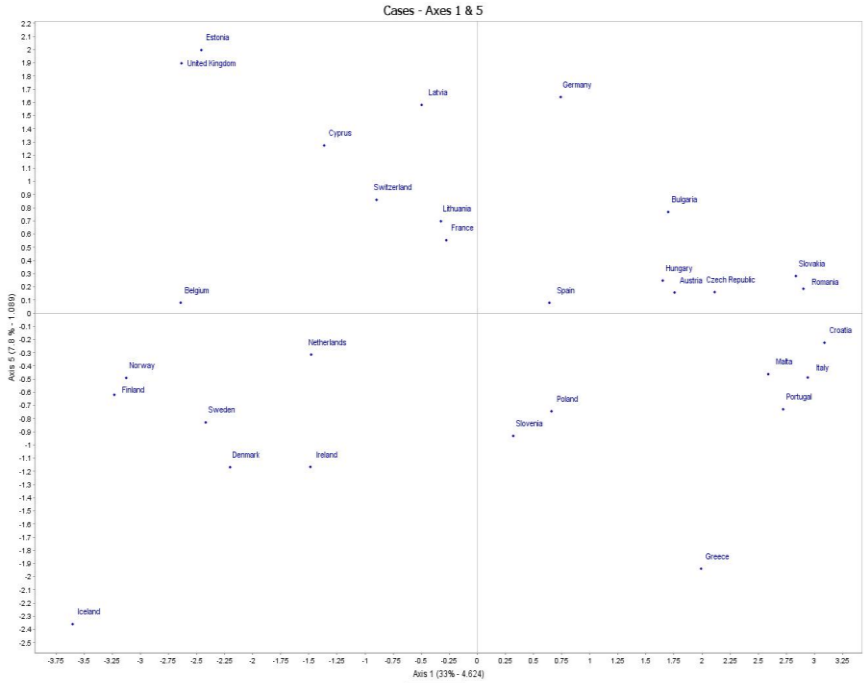

Figure B.10 - Individuals' representation in the principal plan [1,5] - 2nd Step of DPCA – 2012

Table B.4 - Coordinates and contributions of individuals of 2nd Step of DPCA – 2014

	Axis 1			Axis 2			Axis 3			Axis 4			Axis 5		
Countries	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
Austria	0.264	0.047	0.016	0.779	0.741	0.139	-0.901	1.630	0.185	-0.394	0.394	0.035	-0.102	0.033	0.002
Belgium	-2.465	**4.108**	0.583	0.688	0.577	0.045	0.674	0.912	0.044	0.291	0.215	0.008	-0.075	0.018	0.001
Bulgaria	1.495	1.512	0.391	0.169	0.035	0.005	0.656	0.864	0.075	-0.309	0.243	0.017	-0.473	0.707	0.039
Croatia	2.918	**5.758**	0.740	-0.710	0.615	0.044	0.688	0.951	0.041	-0.464	0.548	0.019	0.650	1.340	0.037
Cyprus	-0.946	0.606	0.069	-2.417	**7.129**	0.452	0.898	1.622	0.063	0.770	1.503	0.046	-1.319	**5.512**	0.135
Czech Republic	1.909	2.463	0.459	1.580	3.047	0.315	0.731	1.073	0.067	-0.243	0.150	0.007	0.429	0.582	0.023
Denmark	-2.279	3.511	0.589	0.448	0.245	0.023	-0.702	0.989	0.056	-0.327	0.272	0.012	1.569	**7.797**	0.279
Estonia	-2.383	**3.839**	0.281	0.710	0.616	0.025	2.387	**11.450**	0.283	0.260	0.172	0.003	-1.962	**12.191**	0.191
Finland	-3.268	**7.220**	0.675	-0.431	0.227	0.012	-0.201	0.081	0.003	0.256	0.167	0.004	0.400	0.506	0.010
France	-0.285	0.055	0.016	0.003	0.000	0.000	-0.310	0.193	0.019	-1.313	**4.375**	0.344	-0.425	0.573	0.036
Germany	1.572	1.671	0.129	1.199	1.754	0.075	-2.684	**14.469**	0.375	-1.978	**9.935**	0.204	-1.381	**6.045**	0.099
Greece	2.136	3.086	0.109	-5.576	**37.949**	0.745	-0.482	0.467	0.006	1.101	3.080	0.029	0.933	2.758	0.021
Hungary	1.773	2.125	0.467	1.685	3.466	0.422	0.048	0.005	0.000	-0.273	0.189	0.011	-0.093	0.028	0.001
Iceland	-3.894	**10.253**	0.564	1.507	2.774	0.084	0.844	1.432	0.026	0.327	0.271	0.004	2.722	**23.475**	0.276
Ireland	-1.551	1.627	0.110	-2.571	**8.070**	0.301	-2.619	**13.775**	0.313	0.448	0.509	0.009	0.223	0.157	0.002
Italy	3.306	**7.390**	0.733	1.084	1.435	0.079	0.407	0.333	0.011	1.326	**4.464**	0.118	0.567	1.017	0.022
Latvia	-0.274	0.051	0.007	-0.091	0.010	0.001	2.395	**11.522**	0.507	0.756	1.449	0.050	-1.779	**10.026**	0.280
Lithuania	-0.478	0.154	0.011	-1.174	1.682	0.066	0.447	0.401	0.010	-4.176	**44.279**	0.838	-0.204	0.132	0.002
Malta	2.761	**5.156**	0.435	2.200	**5.909**	0.276	-1.257	**3.176**	0.090	1.457	**5.387**	0.121	0.242	0.185	0.003
Netherlands	-1.314	1.167	0.307	0.004	0.000	0.000	-1.538	**4.751**	0.420	-0.812	1.674	0.117	0.329	0.343	0.019
Norway	-3.492	**8.246**	0.823	0.730	0.650	0.036	-0.631	0.799	0.027	-0.093	0.022	0.001	0.620	1.219	0.026
Poland	0.693	0.325	0.088	0.465	0.264	0.039	1.302	**3.404**	0.309	0.433	0.477	0.034	0.573	1.039	0.060
Portugal	2.471	**4.127**	0.696	-0.436	0.232	0.022	-0.476	0.456	0.026	0.951	2.297	0.103	0.160	0.081	0.003
Romania	3.392	**7.777**	0.616	2.099	**5.380**	0.236	-0.786	1.241	0.033	0.723	1.326	0.028	-0.019	0.001	0.000
Slovakia	3.212	**6.976**	0.520	-0.011	0.000	0.000	1.796	**6.480**	0.163	-1.765	**7.913**	0.157	0.375	0.444	0.007
Slovenia	0.449	0.136	0.058	0.226	0.062	0.015	0.545	0.597	0.085	1.055	2.825	0.320	0.698	1.545	0.140
Spain	0.636	0.273	0.025	-3.631	**16.091**	0.821	0.758	1.153	0.036	0.071	0.013	0.000	-0.076	0.018	0.000
Sweden	-2.450	**4.059**	0.576	0.494	0.298	0.023	0.942	1.781	0.085	-0.148	0.055	0.002	0.882	2.463	0.075
Switzerland	-1.042	0.735	0.090	0.231	0.065	0.004	-2.621	**13.797**	0.567	0.776	1.528	0.050	-1.382	**6.049**	0.158
United Kingdom	-2.864	**5.547**	0.390	0.744	0.676	0.026	-0.311	0.195	0.005	1.297	**4.268**	0.080	-2.081	**13.714**	0.206

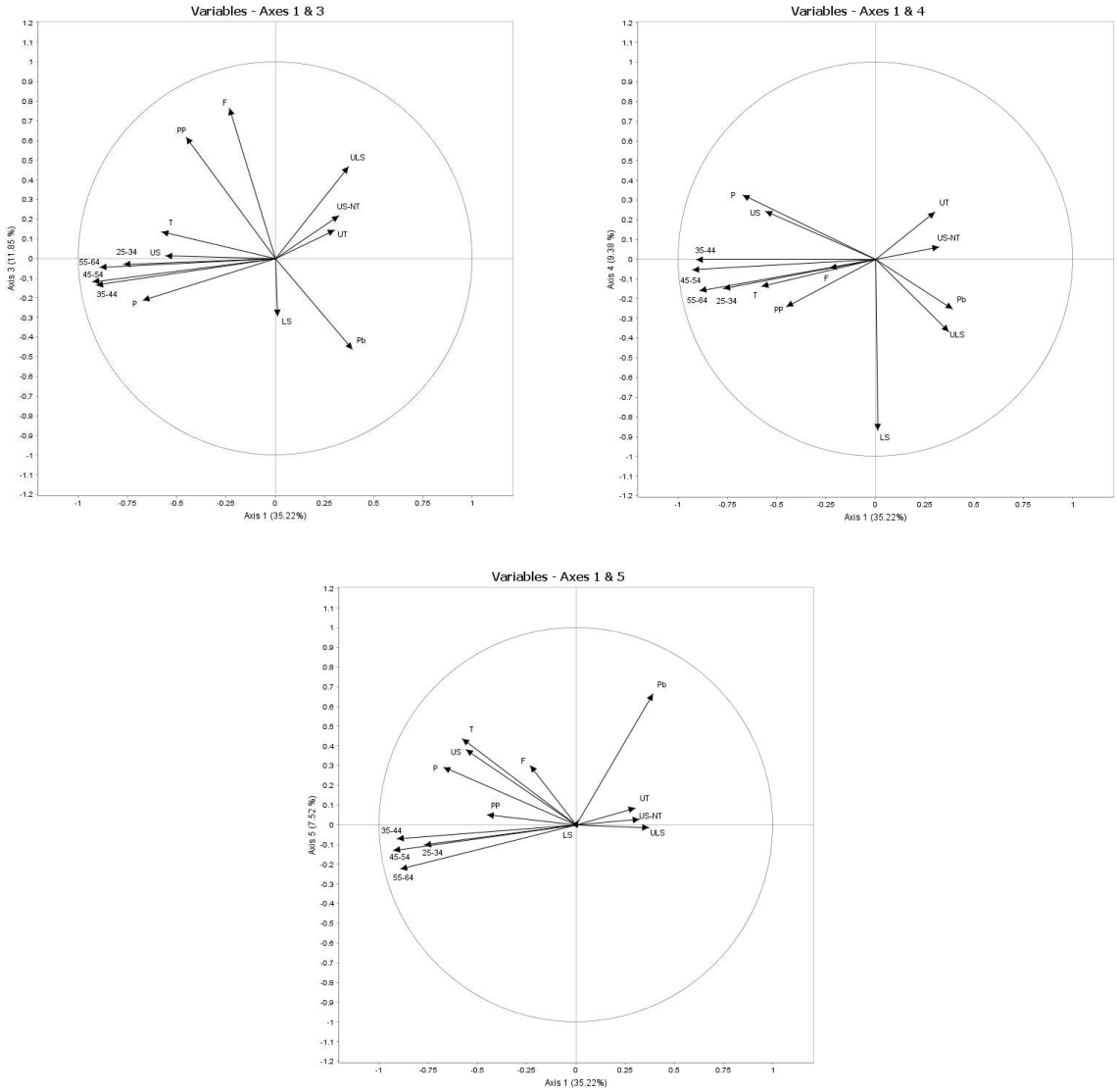

Figure B.11 - Correlation Circles of the variables in the principal plan [1,3], in the principal plan [1,4] and in the principal plan [1,5] - 2nd Step of DPCA - 2014

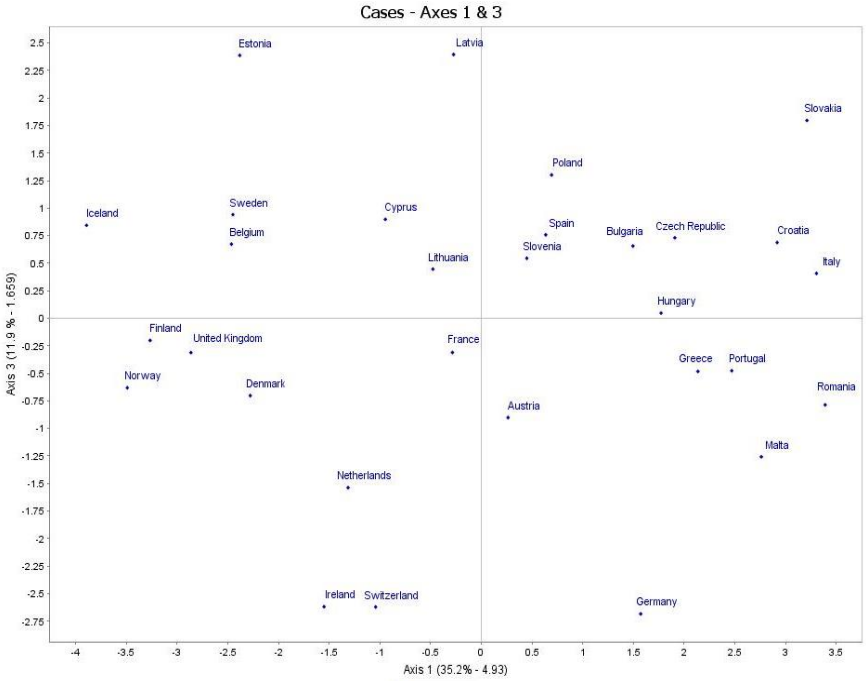

Figure B.12 - Individuals' representation in the principal plan [1,3] - 2nd Step of DPCA - 2014

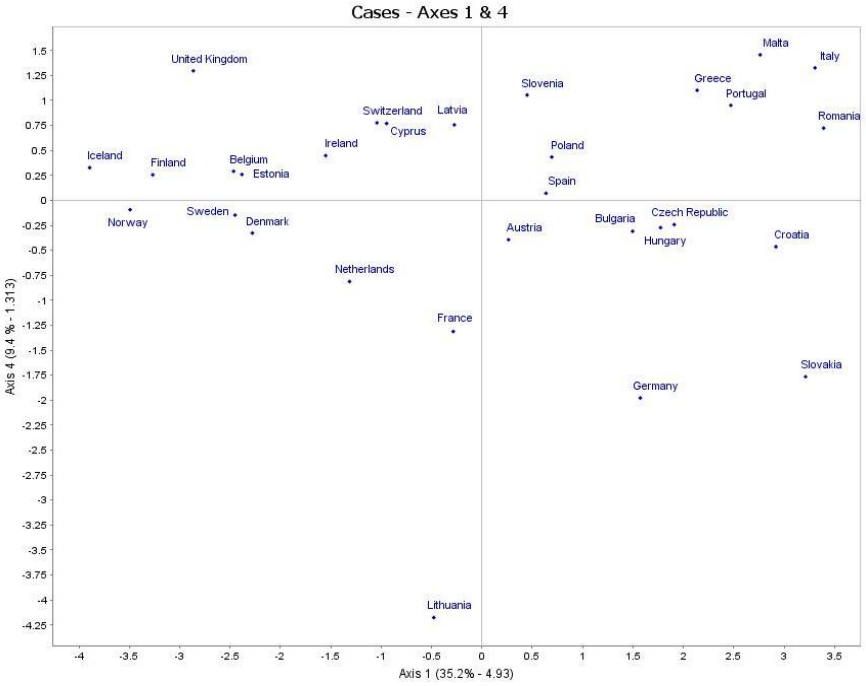

Figure B.13 - Individuals' representation in the principal plan [1,4] - 2nd Step of DPCA – 2014

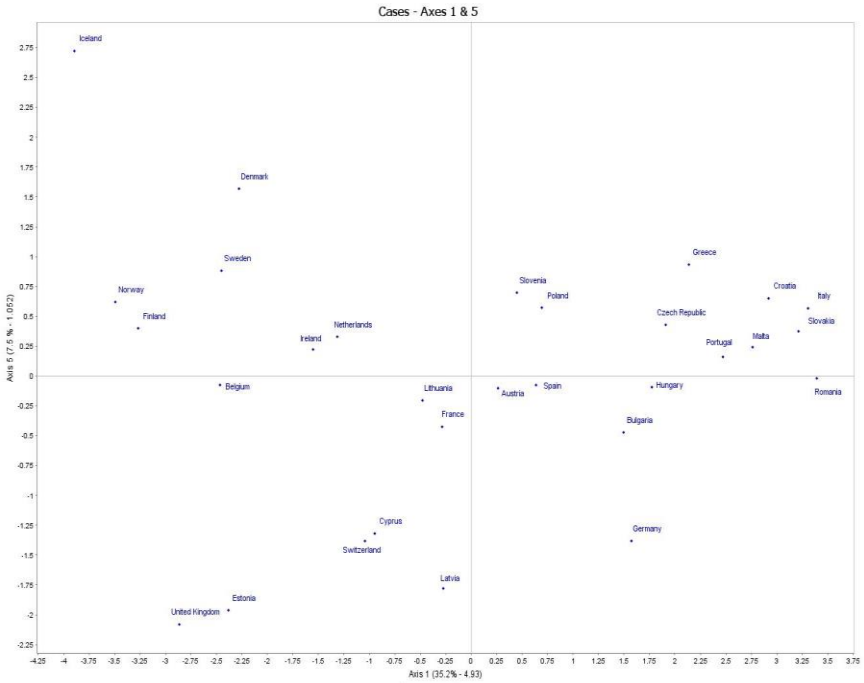

Figure B.14 - Individuals' representation in the principal plan [1,5] - 2nd Step of DPCA – 2014

Table B.5 - Coordinates and contributions of individuals of 2nd Step of DPCA - 2016

	Axis 1			Axis 2			Axis 3			Axis 4		
Countries	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**	**Coord**	**CTA**	**CTR**
Austria	0.488	0.159	0.064	0.200	0.045	0.011	-0.790	1.324	0.169	-0.570	0.824	0.088
Belgium	-2.239	**3.348**	0.514	0.711	0.568	0.052	0.804	1.372	0.066	0.322	0.263	0.011
Bulgaria	1.714	1.963	0.456	0.363	0.148	0.020	0.630	0.842	0.062	-0.491	0.611	0.037
Croatia	2.529	**4.270**	0.673	-0.672	0.507	0.048	0.473	0.475	0.024	-0.082	0.017	0.001
Cyprus	-1.485	1.473	0.168	-2.401	**6.476**	0.439	0.783	1.302	0.047	0.361	0.331	0.010
Czech Republic	2.192	**3.210**	0.510	1.550	2.699	0.255	0.717	1.090	0.054	-0.107	0.029	0.001
Denmark	-2.117	2.992	0.525	0.218	0.053	0.006	-1.266	**3.401**	0.188	0.425	0.458	0.021
Estonia	-2.054	2.817	0.222	1.230	1.700	0.080	2.664	**15.062**	0.374	-0.348	0.307	0.006
Finland	-3.195	**6.819**	0.618	-0.524	0.308	0.017	0.364	0.282	0.008	0.419	0.445	0.011
France	-0.340	0.077	0.025	-0.240	0.065	0.013	0.113	0.027	0.003	-1.416	**5.087**	0.442
Germany	1.688	1.902	0.166	0.903	0.916	0.048	-2.241	**10.662**	0.293	-2.386	**14.441**	0.332
Greece	1.364	1.242	0.039	-6.558	**48.306**	0.897	-0.264	0.147	0.001	1.027	2.678	0.022
Hungary	2.305	**3.549**	0.578	1.785	3.580	0.347	-0.309	0.202	0.010	-0.331	0.277	0.012
Iceland	-3.701	**9.145**	0.518	2.284	**5.859**	0.197	-0.001	0.000	0.000	1.704	**7.369**	0.110
Ireland	-2.126	3.019	0.218	-1.812	**3.686**	0.158	-2.573	**14.053**	0.319	0.444	0.500	0.009
Italy	3.498	**8.170**	0.773	0.377	0.160	0.009	0.512	0.557	0.017	1.464	**5.436**	0.135
Latvia	-0.594	0.236	0.030	-0.029	0.001	0.000	2.735	**15.871**	0.644	-0.012	0.000	0.000
Lithuania	-0.822	0.452	0.034	-0.796	0.712	0.032	0.257	0.141	0.003	-3.873	**38.054**	0.751
Malta	3.267	**7.128**	0.556	1.864	**3.901**	0.181	-1.078	2.467	0.061	1.525	**5.899**	0.121
Netherlands	-1.184	0.935	0.241	0.135	0.021	0.003	-1.804	**6.907**	0.561	-0.670	1.139	0.077
Norway	-3.078	**6.325**	0.776	0.620	0.432	0.032	-0.574	0.699	0.027	0.548	0.762	0.025
Poland	0.787	0.414	0.113	0.635	0.454	0.074	0.566	0.681	0.059	1.059	2.845	0.205
Portugal	2.265	**3.426**	0.631	-0.649	0.473	0.052	-0.496	0.522	0.030	0.871	1.924	0.093
Romania	3.660	**8.946**	0.717	1.782	3.565	0.170	-0.835	1.481	0.037	0.509	0.657	0.014
Slovakia	3.144	**6.603**	0.538	0.032	0.001	0.000	1.673	**5.941**	0.152	-1.343	**4.577**	0.098
Slovenia	0.455	0.138	0.062	0.129	0.019	0.005	0.435	0.402	0.056	1.326	4.459	0.524
Spain	0.178	0.021	0.002	-3.296	**12.200**	0.796	0.982	2.049	0.071	-0.123	0.038	0.001
Sweden	-2.514	**4.220**	0.536	0.755	0.639	0.048	0.970	1.996	0.080	0.079	0.016	0.001
Switzerland	-1.011	0.682	0.089	-0.087	0.009	0.001	-2.156	**9.862**	0.405	0.122	0.038	0.001
United Kingdom	-3.076	**6.319**	0.422	1.491	2.497	0.099	-0.292	0.181	0.004	-0.453	0.520	0.009

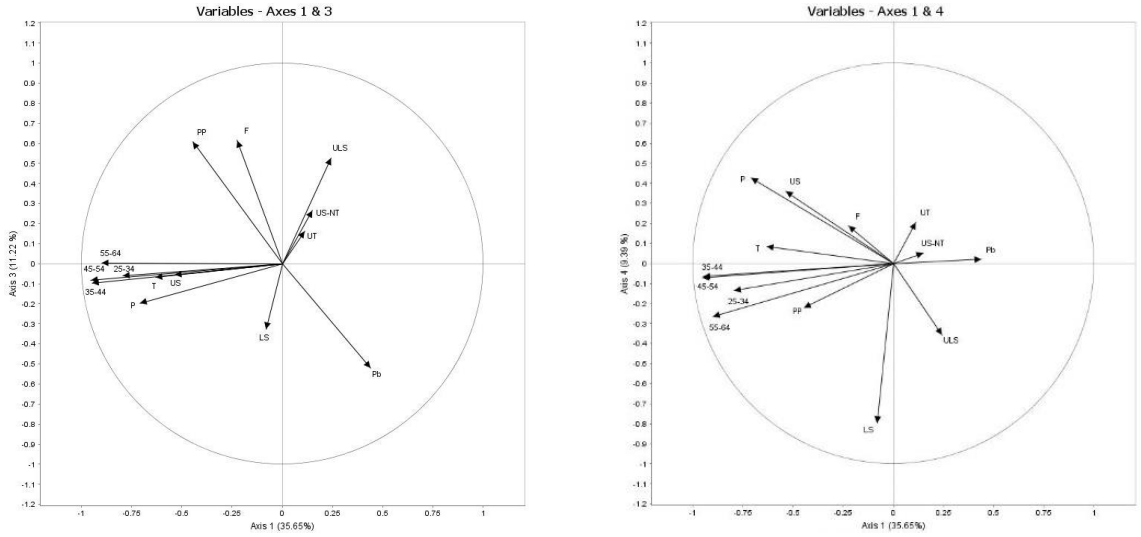

Figure B.15 - Correlation Circles of the variables in the principal plan [1,3] and in the principal plan [1,4] - 2nd Step of DPCA – 2016

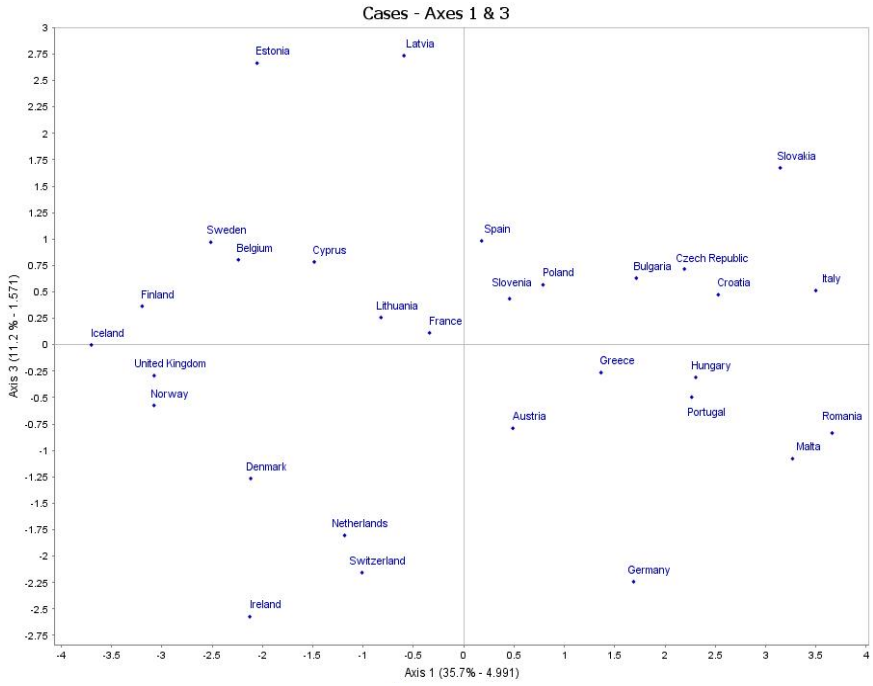

Figure B.16 - Individuals' representation in the principal plan [1,3] - 2nd Step of DPCA – 2016

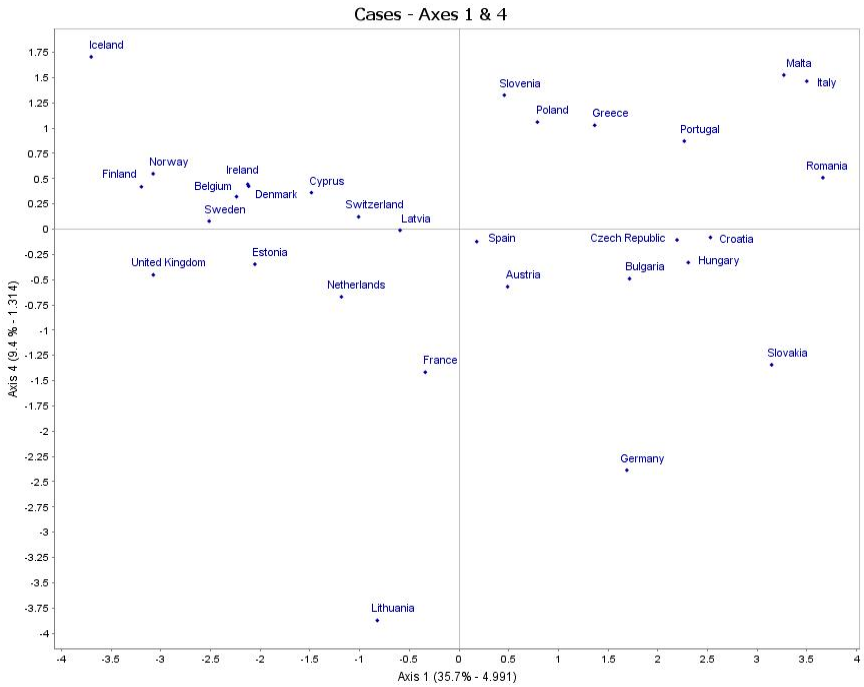

Figure B.17 - Individuals' representation in the principal plan [1,4] - 2nd Step of DPCA – 2016

3rd Step of DPCA – 2nd Criterion

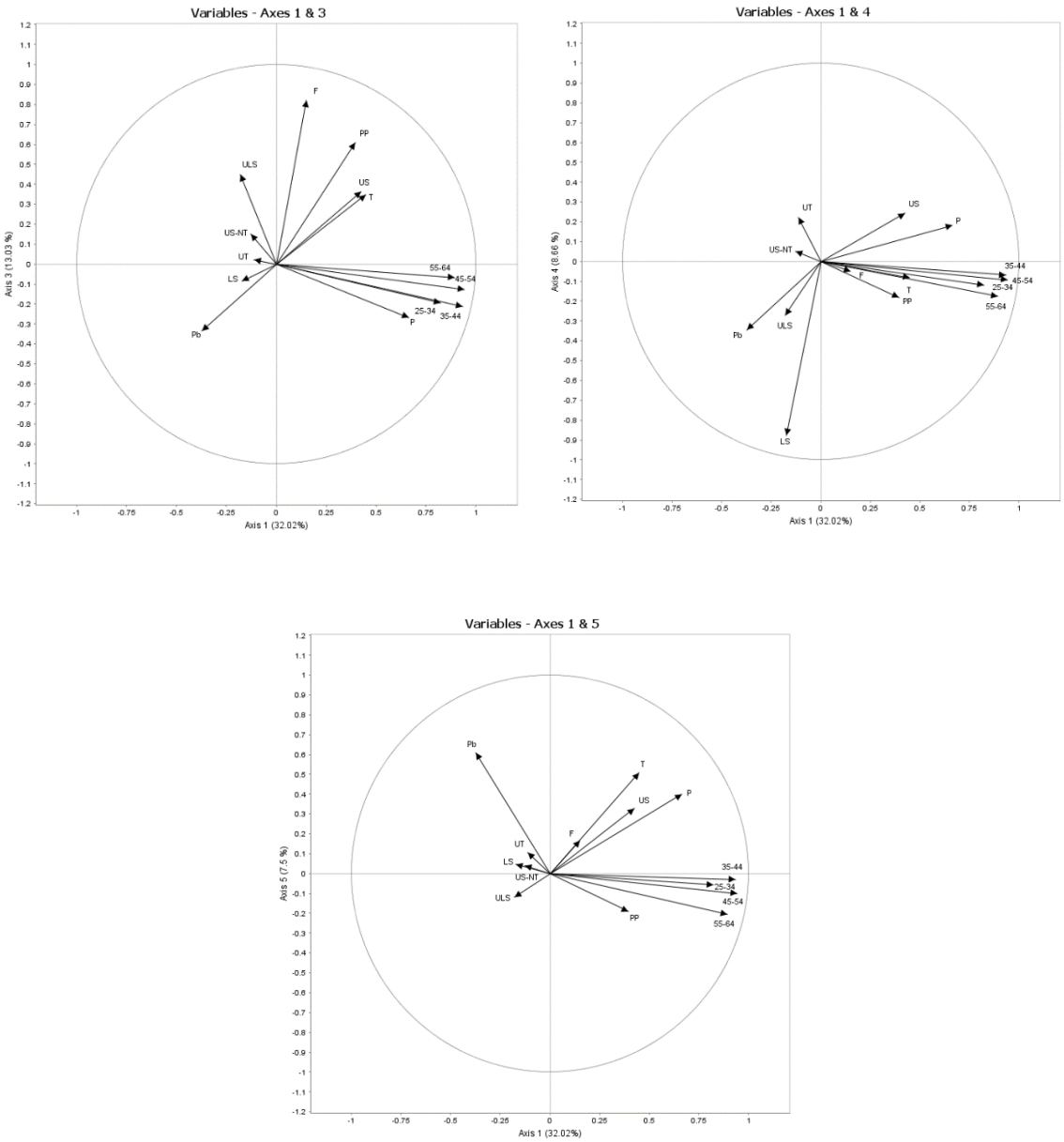

Figure C.1 - Correlation Circles of the variables in the principal plan [1,3], in the principal plan [1,4] and in the principal plan [1,5] – 3rd Step of DPCA – 2nd Criterion

Table D.1 – 3rd Step of DPCA – Coordinates of the Variables on the First Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Pre-Primary	-0.357	-0.306	-0.345	-0.442	-0.466
Primary	-0.640	-0.634	-0.683	-0.702	-0.735
Lower Secondary	0.221	0.215	0.079	0.021	-0.070
Upper Secondary	-0.419	-0.473	-0.489	-0.560	-0.564
Tertiary	-0.273	-0.308	-0.473	-0.586	-0.604
Public	0.309	0.274	0.376	0.406	0.448
25-34	-0.818	-0.773	-0.794	-0.807	-0.802
35-44	-0.918	-0.886	-0.927	-0.948	-0.965
45-54	-0.922	-0.863	-0.929	-0.962	-0.972
55-64	-0.876	-0.826	-0.870	-0.921	-0.917
Females	-0.198	-0.089	-0.193	-0.221	-0.246
Unemployment less than Secondary	0.354	0.148	0.281	0.374	0.309
Unemployment with Secondary to Non-Tertiary	0.451	0.119	0.304	0.319	0.237
Unemployment with Tertiary	0.341	0.119	0.328	0.299	0.194

Table D.2 – 3rd Step of DPCA – Coordinates of the Variables on the Second Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Pre-Primary	-0.271	-0.197	-0.349	-0.375	-0.365
Primary	0.006	0.115	0.161	0.130	0.136
Lower Secondary	0.058	0.109	0.138	0.031	0.001
Upper Secondary	-0.376	-0.398	-0.550	-0.429	-0.352
Tertiary	0.137	0.363	0.276	0.338	0.526
Public	-0.022	-0.219	-0.033	0.057	0.030
25-34	0.211	0.486	0.411	0.387	0.432
35-44	0.134	0.324	0.246	0.221	0.294
45-54	0.060	0.336	0.223	0.189	0.247
55-64	-0.006	0.259	0.175	0.120	0.201
Females	-0.271	-0.076	-0.384	-0.369	-0.447
Unemployment less than Secondary	0.214	0.576	0.564	0.575	0.584
Unemployment with Secondary to Non-Tertiary	0.648	0.836	0.881	0.877	0.890
Unemployment with Tertiary	0.626	0.795	0.812	0.822	0.870

Table D.3 – 3rd Step of DPCA – Coordinates of the Variables on the Third Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Pre-Primary	-0.521	-0.626	-0.601	-0.591	-0.567
Primary	0.376	0.335	0.240	0.168	0.079
Lower Secondary	-0.038	0.051	0.085	0.122	0.151
Upper Secondary	-0.406	-0.346	-0.187	-0.093	-0.103
Tertiary	-0.540	-0.659	-0.466	-0.315	-0.189
Public	0.317	0.371	0.265	0.281	0.315
25-34	0.292	0.078	0.041	-0.023	-0.042
35-44	0.281	0.175	0.180	0.083	-0.028
45-54	0.196	0.057	0.082	0.043	-0.041
55-64	0.107	0.030	0.078	-0.015	-0.091
Females	-0.810	-0.953	-0.776	-0.724	-0.606
Unemployment less than Secondary	-0.617	-0.801	-0.582	-0.492	-0.462
Unemployment with Secondary to Non-Tertiary	-0.341	-0.660	-0.239	-0.225	-0.215
Unemployment with Tertiary	-0.116	-0.429	-0.096	-0.140	-0.140

Table D.4 – 3rd Step of DPCA – Coordinates of the Variables on the Fourth Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Pre-Primary	-0.155	-0.096	-0.153	-0.140	-0.132
Primary	0.204	0.168	0.156	0.172	0.168
Lower Secondary	-0.872	-0.855	-0.900	-0.905	-0.853
Upper Secondary	0.214	0.209	0.140	0.099	0.113
Tertiary	0.005	0.065	-0.159	-0.153	-0.068
Public	-0.521	-0.588	-0.506	-0.473	-0.376
25-34	0.042	0.031	-0.036	-0.111	-0.121
35-44	0.051	0.046	-0.051	-0.039	-0.112
45-54	0.001	0.070	-0.058	-0.073	-0.084
55-64	-0.076	0.001	-0.126	-0.142	-0.180
Females	-0.114	-0.023	0.004	0.057	0.090
Unemployment less than Secondary	-0.254	-0.155	-0.204	-0.144	-0.114
Unemployment with Secondary to Non-Tertiary	0.166	0.070	0.134	0.218	0.263
Unemployment with Tertiary	0.382	0.300	0.261	0.346	0.345

Table D.5 – 3rd Step of DPCA – Coordinates of the Variables on the Fifth Axis for the trajectories according to the 1st Criterion

	2008	**2010**	**2012**	**2014**	**2016**
Pre-Primary	0.065	-0.015	0.069	0.219	0.375
Primary	-0.322	-0.393	-0.448	-0.504	-0.507
Lower Secondary	0.192	0.158	0.148	0.170	0.261
Upper Secondary	-0.212	-0.302	-0.360	-0.334	-0.305
Tertiary	-0.342	-0.409	-0.385	-0.440	-0.471
Public	-0.503	-0.514	-0.658	-0.654	-0.626
25-34	0.052	0.002	0.066	-0.007	-0.045
35-44	0.050	-0.031	-0.013	-0.075	-0.096
45-54	0.201	0.134	0.128	0.018	-0.030
55-64	0.270	0.215	0.238	0.176	0.202
Females	-0.122	-0.211	-0.092	-0.009	0.023
Unemployment less than Secondary	0.135	0.135	0.134	0.119	0.125
Unemployment with Secondary to Non-Tertiary	-0.090	-0.004	-0.093	-0.190	-0.206
Unemployment with Tertiary	-0.284	-0.056	-0.199	-0.305	-0.355

Table E.1 – 3rd Step of DPCA – Coordinates of the Individuals on the First Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Austria	3.341	3.769	3.647	0.625	0.999
Belgium	-4.831	-5.684	-5.490	-5.247	-4.891
Bulgaria	1.632	2.398	3.530	3.166	3.520
Croatia	6.140	5.112	6.421	6.174	5.493
Cyprus	-3.851	-3.003	-2.835	-2.071	-2.666
Czech Republic	4.800	4.677	4.390	4.066	4.330
Denmark	-4.888	-4.038	-4.575	-4.769	-4.492
Estonia	-5.359	-4.395	-5.109	-5.123	-4.649
Finland	-5.850	-6.375	-6.718	-6.880	-6.707
France	-0.254	-0.596	-0.576	-0.653	-0.651
Germany	2.444	1.929	1.541	3.315	3.365
Greece	3.681	2.600	4.138	4.537	4.118
Hungary	3.293	3.324	3.429	3.758	4.502
Iceland	-7.222	-5.959	-7.490	-8.164	-8.224
Ireland	-3.019	-3.727	-3.090	-3.303	-4.110
Italy	5.166	5.483	6.112	6.995	7.270
Latvia	-0.785	-0.999	-1.033	-0.652	-1.261
Lithuania	-0.463	-0.609	-0.675	-0.995	-1.449
Malta	5.949	6.930	5.381	5.889	6.537
Netherlands	-3.909	-2.234	-3.076	-2.757	-2.514
Norway	-7.304	-5.944	-6.498	-7.340	-6.599
Poland	2.695	1.501	1.371	1.505	1.576
Portugal	6.086	5.053	5.657	5.192	4.881
Romania	5.002	5.427	6.032	7.166	7.349
Slovakia	6.755	5.194	5.892	6.782	6.697
Slovenia	0.779	1.054	0.664	1.002	0.938
Spain	1.569	0.216	1.331	1.281	0.999
Sweden	-4.189	-4.594	-5.027	-5.175	-5.446
Switzerland	-3.339	-2.027	-1.867	-2.194	-2.144
United Kingdom	-4.071	-4.485	-5.475	-6.131	-6.769

Table E.2 – 3rd Step of DPCA – Coordinates of the Individuals on the Second Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Austria	-2.070	-2.807	-3.059	-1.251	-0.598
Belgium	-0.641	-1.207	-2.170	-1.179	-0.677
Bulgaria	-1.251	-0.673	-0.070	-0.096	-0.879
Croatia	0.737	0.077	0.569	1.090	0.313
Cyprus	1.023	1.521	2.762	3.767	4.017
Czech Republic	-2.380	-2.287	-2.040	-2.258	-2.662
Denmark	-1.573	-1.265	-0.870	-1.024	-0.221
Estonia	0.191	3.404	0.983	-0.695	-0.782
Finland	1.034	0.457	-0.078	0.267	1.094
France	0.559	-0.248	-0.281	0.210	0.813
Germany	0.573	-1.053	-1.042	-1.465	-1.499
Greece	3.945	3.492	7.431	8.009	8.677
Hungary	-1.260	-1.448	-1.495	-2.439	-3.146
Iceland	-3.136	-2.378	-3.296	-2.911	-3.027
Ireland	3.181	4.895	5.342	3.810	3.218
Italy	-0.925	-2.331	-2.554	-1.712	-1.683
Latvia	0.548	3.102	1.219	0.549	0.728
Lithuania	1.314	4.319	3.538	2.206	2.194
Malta	-1.803	-2.596	-3.622	-3.366	-3.727
Netherlands	-0.270	-0.606	-0.609	-0.100	-0.049
Norway	-1.335	-1.778	-2.147	-1.323	-0.410
Poland	0.619	-0.218	-0.506	-0.709	-1.142
Portugal	1.516	-0.429	1.541	0.657	0.306
Romania	-1.664	-2.512	-3.373	-3.099	-3.552
Slovakia	-0.751	-0.082	0.432	0.409	-0.272
Slovenia	-0.323	-0.903	-0.744	-0.550	-0.569
Spain	3.614	4.120	5.662	5.491	4.917
Sweden	-0.399	-0.770	-1.389	-0.896	-0.473
Switzerland	0.417	-0.050	-0.354	-0.382	0.246
United Kingdom	0.509	0.254	0.219	-1.012	-1.155

Table E.3 – 3rd Step of DPCA – Coordinates of the Individuals on the Third Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Austria	0.942	1.645	0.929	0.830	0.739
Belgium	-0.967	-0.568	-0.750	-0.765	-0.851
Bulgaria	0.107	-0.302	-0.546	-0.800	-0.691
Croatia	0.052	0.072	-0.185	-0.865	-0.613
Cyprus	3.114	3.474	1.507	-0.188	0.268
Czech Republic	-0.587	-0.549	-1.259	-1.193	-1.384
Denmark	0.422	0.126	0.326	0.377	0.870
Estonia	-1.932	-4.320	-2.337	-2.457	-2.494
Finland	-0.226	0.142	0.244	0.172	0.008
France	0.252	0.948	0.665	0.170	-0.279
Germany	1.517	2.768	2.614	2.590	2.150
Greece	1.115	0.685	0.539	1.174	1.711
Hungary	-1.899	-1.223	-0.841	-0.273	-0.042
Iceland	-1.357	-1.861	-2.101	-1.559	-1.000
Ireland	2.987	1.890	3.039	3.296	3.035
Italy	-0.353	0.328	0.129	-0.317	-0.492
Latvia	-3.839	-5.835	-3.155	-2.302	-2.449
Lithuania	-1.578	-4.140	-2.064	-1.288	-0.882
Malta	1.881	3.385	1.972	1.542	0.856
Netherlands	3.356	3.969	2.105	1.438	1.700
Norway	0.856	1.424	0.406	0.521	0.239
Poland	-1.510	-1.563	-1.798	-1.492	-0.832
Portugal	0.394	0.845	0.737	0.806	0.882
Romania	-1.608	-0.595	0.597	0.857	0.603
Slovakia	-3.999	-3.479	-2.827	-2.515	-2.325
Slovenia	-0.590	-0.271	-0.442	-0.471	-0.396
Spain	-0.034	-1.582	-1.320	-0.411	-0.259
Sweden	-1.026	-1.062	-1.449	-1.274	-1.538
Switzerland	4.027	4.640	4.237	3.399	2.901
United Kingdom	0.482	1.009	1.031	0.997	0.564

Table E.4 – 3rd Step of DPCA – Coordinates of the Individuals on the Fourth Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Austria	-0.254	0.082	-0.442	-0.735	-0.775
Belgium	1.521	1.214	1.374	0.495	0.656
Bulgaria	0.116	0.407	0.239	0.008	-0.206
Croatia	0.067	-0.243	-0.182	-0.193	-0.048
Cyprus	1.002	1.492	1.565	1.533	1.565
Czech Republic	-0.246	0.153	-0.223	-0.197	-0.380
Denmark	-1.131	-1.060	-0.974	-0.971	-0.745
Estonia	0.738	0.703	1.100	1.283	1.176
Finland	-0.088	-0.025	0.000	0.122	0.748
France	-0.674	-0.778	-1.098	-1.385	-1.237
Germany	-1.975	-1.774	-2.488	-2.736	-2.567
Greece	1.390	0.934	1.357	1.409	1.530
Hungary	-0.490	-0.334	-0.292	-0.353	-0.540
Iceland	0.016	-0.009	-0.369	-0.091	0.257
Ireland	0.492	0.125	-0.044	-0.328	-0.435
Italy	1.343	1.214	1.164	1.329	1.388
Latvia	1.582	1.308	1.959	1.953	1.651
Lithuania	-3.655	-4.075	-4.232	-3.914	-3.571
Malta	-1.304	-1.195	0.697	0.800	0.667
Netherlands	-0.825	-0.410	-1.154	-1.410	-1.310
Norway	-0.422	-0.227	-0.502	-0.624	-0.317
Poland	1.105	0.757	0.729	0.740	0.778
Portugal	0.943	0.600	0.793	0.853	0.859
Romania	0.605	0.792	0.279	0.337	0.098
Slovakia	-1.638	-1.373	-1.203	-1.141	-0.941
Slovenia	0.927	1.022	1.031	1.080	1.079
Spain	-0.251	-0.733	-0.444	0.686	0.675
Sweden	-0.657	-0.635	-0.396	-0.161	-0.390
Switzerland	-0.061	0.476	0.190	0.064	-0.077
United Kingdom	1.827	1.592	1.565	1.548	0.413

Table E.5 – 3rd Step of DPCA – Coordinates of the Individuals on the Fifth Axis for the trajectories according to the 1st Criterion

	2008	2010	2012	2014	2016
Austria	0.139	-0.395	0.159	0.263	0.290
Belgium	0.337	0.043	0.081	0.216	0.132
Bulgaria	0.701	0.709	0.775	0.782	0.888
Croatia	-0.061	0.186	-0.226	-0.338	-0.288
Cyprus	0.986	1.155	1.285	0.420	-0.027
Czech Republic	0.088	-0.274	0.162	0.349	0.572
Denmark	-0.708	-0.882	-1.178	-1.505	-1.560
Estonia	1.844	2.704	2.015	2.163	2.385
Finland	-0.455	-0.587	-0.625	-0.756	-0.432
France	0.035	0.018	0.559	0.635	0.585
Germany	1.469	1.066	1.656	1.761	1.821
Greece	-1.215	-0.652	-1.956	-2.728	-3.195
Hungary	0.423	0.252	0.250	0.799	1.060
Iceland	-1.624	-2.384	-2.381	-2.219	-1.778
Ireland	-1.238	-0.680	-1.177	-1.696	-1.930
Italy	-0.958	-0.886	-0.492	-0.200	-0.380
Latvia	1.407	2.358	1.595	1.774	1.670
Lithuania	0.753	1.524	0.704	0.879	1.446
Malta	-0.471	-0.679	-0.467	0.009	-0.186
Netherlands	-0.543	-0.969	-0.317	-0.476	-0.319
Norway	-0.863	-1.396	-0.495	-0.710	-1.059
Poland	-0.647	-0.675	-0.751	-0.292	-0.441
Portugal	-0.940	-0.710	-0.736	-0.430	-0.464
Romania	-0.151	-0.582	0.187	0.552	0.641
Slovakia	0.863	0.794	0.285	0.630	0.875
Slovenia	-1.053	-1.145	-0.940	-0.728	-0.767
Spain	0.278	1.111	0.080	-0.831	-0.844
Sweden	-0.439	-0.898	-0.836	-0.638	-0.559
Switzerland	0.398	0.195	0.868	0.639	0.192
United Kingdom	1.648	1.679	1.914	1.675	1.672

Table E.6 – 3rd Step of DPCA – Coordinates of the Individuals on the First Axis for the trajectories according to the 2nd Criterion

	2008	2010	2012	2014	2016
Austria	-2.603	-2.162	-1.983	-0.132	0.151
Belgium	1.835	2.330	2.473	2.589	2.782
Bulgaria	-1.612	-1.597	-1.614	-0.972	-0.798
Croatia	-3.237	-2.885	-2.938	-2.395	-1.897
Cyprus	0.956	1.079	1.472	1.622	2.251
Czech Republic	-3.127	-2.649	-2.196	-1.639	-1.426
Denmark	1.296	1.584	1.988	2.305	2.606
Estonia	2.040	1.768	2.462	2.713	2.905
Finland	2.362	2.658	3.024	3.481	3.890
France	-0.391	-0.114	0.196	0.598	0.942
Germany	-1.795	-1.409	-0.966	-1.368	-1.030
Greece	-1.713	-1.582	-1.455	-1.085	-0.518
Hungary	-2.041	-2.038	-1.685	-1.520	-1.571
Iceland	2.344	2.568	3.208	3.648	4.056
Ireland	0.892	1.247	1.633	2.142	2.892
Italy	-3.053	-3.025	-2.925	-2.856	-2.681
Latvia	0.283	0.175	0.614	0.759	1.402
Lithuania	-0.109	-0.114	0.221	0.729	1.345
Malta	-4.005	-3.785	-2.764	-2.515	-2.470
Netherlands	0.687	0.619	1.240	1.436	1.698
Norway	2.304	2.495	2.786	3.501	3.694
Poland	-1.482	-1.012	-0.731	-0.408	-0.090
Portugal	-3.081	-2.802	-2.503	-1.929	-1.520
Romania	-2.890	-2.813	-3.004	-3.158	-2.958
Slovakia	-3.293	-3.009	-2.761	-2.699	-2.429
Slovenia	-0.978	-0.811	-0.386	-0.106	0.329
Spain	-0.686	-0.524	-0.187	0.308	0.718
Sweden	1.388	1.841	2.262	2.690	3.210
Switzerland	0.445	0.458	0.723	1.395	1.949
United Kingdom	1.257	1.735	2.451	3.029	3.531

Table E.7 – 3rd Step of DPCA – Coordinates of the Individuals on the Second Axis for the trajectories according to the 2nd Criterion

	2008	2010	2012	2014	2016
Austria	-2.110	-1.660	-1.499	-0.450	-0.271
Belgium	-1.494	-0.859	-0.994	-0.454	-0.554
Bulgaria	-1.414	0.115	0.850	0.732	-0.198
Croatia	-0.889	0.285	1.426	1.968	0.909
Cyprus	-0.998	0.211	1.846	3.126	2.544
Czech Republic	-1.732	-0.730	-0.409	-0.758	-1.250
Denmark	-2.128	-1.002	-0.601	-0.603	-0.432
Estonia	-0.893	2.740	1.055	-0.096	-0.418
Finland	-0.612	0.045	-0.046	0.344	0.408
France	-0.964	-0.281	-0.047	0.477	0.466
Germany	-0.770	-0.742	-0.890	-0.828	-0.993
Greece	0.661	2.143	5.736	6.593	5.917
Hungary	-1.272	-0.217	-0.072	-0.880	-1.544
Iceland	-3.017	-1.656	-1.860	-1.713	-2.143
Ireland	0.066	2.606	3.027	2.179	1.190
Italy	-1.629	-1.147	-0.552	0.123	-0.131
Latvia	-0.583	2.899	1.710	0.989	0.634
Lithuania	-0.375	3.497	2.420	1.657	1.010
Malta	-2.195	-1.705	-1.762	-1.598	-1.780
Netherlands	-1.608	-1.075	-0.734	-0.095	-0.397
Norway	-2.032	-1.675	-1.626	-1.128	-0.688
Poland	-0.645	0.219	0.497	0.245	-0.417
Portugal	-0.744	-0.196	1.820	1.377	0.730
Romania	-1.824	-1.272	-1.349	-1.107	-1.544
Slovakia	-0.265	1.238	1.514	1.354	0.433
Slovenia	-1.277	-0.407	0.224	0.388	0.031
Spain	0.661	2.963	4.469	4.611	3.300
Sweden	-1.318	-0.475	-0.550	-0.343	-0.511
Switzerland	-1.218	-0.691	-0.648	-0.392	-0.230
United Kingdom	-1.074	-0.235	-0.033	-0.729	-1.226

Table E.8 – 3rd Step of DPCA – Coordinates of the Individuals on the Third Axis for the trajectories according to the 2nd Criterion

	2008	2010	2012	2014	2016
Austria	-0.543	-0.233	-0.146	-0.593	-0.630
Belgium	0.959	1.063	1.120	0.991	0.898
Bulgaria	0.039	0.594	0.689	0.564	-0.016
Croatia	-0.550	-0.069	0.518	0.812	0.334
Cyprus	-2.264	-2.479	-1.146	-0.048	-0.574
Czech Republic	0.536	1.166	1.263	0.848	0.436
Denmark	0.229	0.646	-0.068	-0.164	-0.587
Estonia	1.740	2.516	1.754	1.623	1.333
Finland	-0.232	-0.020	0.031	0.251	0.377
France	-0.230	-0.252	-0.315	-0.370	-0.378
Germany	-1.585	-1.293	-1.572	-2.166	-2.103
Greece	-1.959	-1.643	-0.881	-0.896	-1.205
Hungary	1.309	1.366	1.080	0.304	-0.391
Iceland	1.813	2.215	1.985	1.765	1.209
Ireland	-2.770	-2.457	-2.652	-2.981	-3.172
Italy	0.355	0.467	0.525	0.419	0.132
Latvia	2.848	3.492	2.430	1.616	1.373
Lithuania	0.406	1.785	1.129	0.631	-0.052
Malta	-1.256	-1.371	-0.834	-1.237	-1.266
Netherlands	-2.441	-2.279	-1.366	-1.116	-1.394
Norway	-0.042	0.031	-0.119	-0.483	-0.354
Poland	0.715	1.242	1.413	1.053	0.144
Portugal	-0.785	-0.453	-0.332	-0.681	-1.134
Romania	1.208	1.124	0.221	-0.613	-0.905
Slovakia	2.362	2.536	2.359	1.872	1.060
Slovenia	0.322	0.507	0.499	0.411	-0.023
Spain	-0.658	0.133	0.565	0.121	-0.393
Sweden	0.944	1.279	1.239	0.901	0.685
Switzerland	-3.006	-2.851	-2.832	-2.744	-2.840
United Kingdom	-0.253	-0.198	-0.332	-0.369	-0.286

Table E.9 – 3rd Step of DPCA – Coordinates of the Individuals on the Fourth Axis for the trajectories according to the 2nd Criterion

	2008	2010	2012	2014	2016
Austria	-0.063	-0.011	0.034	-0.251	-0.422
Belgium	1.570	1.282	1.289	0.364	0.360
Bulgaria	0.335	0.462	0.150	-0.073	-0.041
Croatia	-0.054	-0.206	0.020	-0.130	-0.076
Cyprus	0.671	0.955	1.029	0.908	0.938
Czech Republic	0.010	0.201	-0.097	-0.061	-0.129
Denmark	-0.885	-0.948	-0.565	-0.529	-0.523
Estonia	0.298	0.508	0.402	0.435	0.390
Finland	0.033	0.182	0.235	0.167	0.436
France	-0.830	-0.943	-0.997	-1.201	-1.223
Germany	-1.948	-2.090	-2.044	-1.822	-1.744
Greece	1.118	1.107	1.360	1.332	1.090
Hungary	-0.406	-0.218	-0.140	-0.102	-0.106
Iceland	0.186	0.309	0.190	0.112	0.132
Ireland	0.368	0.255	-0.098	-0.004	-0.214
Italy	1.470	1.405	1.334	1.382	1.447
Latvia	0.813	1.143	1.162	1.107	0.812
Lithuania	-4.409	-4.230	-3.988	-3.632	-3.479
Malta	-1.325	-1.505	1.030	1.276	1.233
Netherlands	-0.600	-0.500	-0.705	-0.844	-0.909
Norway	-0.257	-0.255	-0.304	-0.379	-0.299
Poland	0.762	0.754	0.722	0.590	0.749
Portugal	0.476	0.551	0.888	0.968	0.955
Romania	0.402	0.657	0.622	0.792	0.596
Slovakia	-1.699	-1.285	-1.160	-1.226	-0.959
Slovenia	0.871	1.174	1.140	1.028	1.031
Spain	-0.603	-0.692	-0.764	0.240	0.201
Sweden	-0.594	-0.443	-0.256	-0.299	-0.557
Switzerland	0.168	0.320	0.310	0.487	0.346
United Kingdom	1.290	1.032	1.050	1.244	0.103

Table E.10 – 3rd Step of DPCA – Coordinates of the Individuals on the Fifth Axis for the trajectories according to the 2nd Criterion

	2008	2010	2012	2014	2016
Austria	-0.288	-0.056	0.018	-0.142	-0.356
Belgium	-0.225	-0.332	-0.319	-0.430	-0.451
Bulgaria	-0.777	-0.741	-0.898	-0.998	-1.170
Croatia	0.160	0.213	0.299	0.241	0.139
Cyprus	-1.592	-1.594	-1.384	-1.048	-1.072
Czech Republic	0.115	0.131	-0.146	-0.234	-0.482
Denmark	0.441	0.455	1.295	1.480	1.445
Estonia	-1.373	-1.417	-1.934	-2.499	-2.877
Finland	1.452	1.074	0.694	0.303	-0.275
France	-0.371	-0.463	-0.551	-0.649	-0.641
Germany	-1.429	-1.280	-1.192	-1.238	-1.272
Greece	1.520	1.563	1.680	1.647	1.758
Hungary	0.163	-0.094	-0.288	-0.655	-0.922
Iceland	2.269	2.601	2.516	2.177	1.713
Ireland	1.267	1.422	1.136	1.221	1.249
Italy	0.500	0.398	0.252	0.103	-0.023
Latvia	-0.395	-0.840	-1.641	-2.097	-2.237
Lithuania	1.127	0.939	0.217	-0.559	-1.046
Malta	0.225	0.075	0.373	0.157	0.026
Netherlands	0.363	0.525	0.522	0.549	0.384
Norway	0.896	0.852	0.642	0.611	0.701
Poland	1.413	1.129	0.797	0.132	0.058
Portugal	0.667	0.687	0.555	0.214	-0.015
Romania	0.527	0.453	-0.261	-0.276	-0.483
Slovakia	0.706	0.364	-0.001	-0.441	-0.725
Slovenia	1.544	1.273	0.848	0.431	0.087
Spain	-0.529	-0.561	-0.437	-0.071	-0.203
Sweden	0.735	0.849	0.708	0.270	0.176
Switzerland	-0.913	-0.799	-0.805	-0.815	-0.923
United Kingdom	-1.277	-1.419	-1.555	-1.681	-1.732