FEATURES OF BUSINESS DEMOGRAPHY STATISTICS IN EUROPEAN COUNTRIES: RELATION OF ENTERPRISE DEATHS AND BIRTHS TO GDP PER CAPITA AND UNEMPLOYMENT

Berislav Žmuk[1, *], Ksenija Dumičić[1] and Anita Harmina[2]

[1]University of Zagreb – Faculty of Economics and Business
 Zagreb, Croatia

[2]VERN' University of Applied Sciences
 Zagreb, Croatia

ABSTRACT

The aim of this study is to analyse the variations and changes in business demography and macroeconomic features for 28 European countries in 2012 and 29 European countries in 2016 as well as the relations among these features based on the following variables: enterprises birth rates, enterprises death rates, the level of GDP per capita and the unemployment rate. Using the hierarchical cluster analysis approach, the second aim of this research is to form and compare the clusters of European countries with similar characteristics according to these variables for both years. The descriptive statistics results show that a relatively high heterogeneity of European countries relative to the observed variables was present in both years, but the change in variables averages suggests an overall improvement of economic and business demography features in 2016. The correlation analysis revealed that the death rate of enterprises had a statistically significant negative correlation with GDP per capita in both years and a significant positive correlation with the unemployment rate and with the birth rate of enterprises only in 2012. The hierarchical cluster analysis resulted in seven clusters in each observed year. The comparison of clustering solutions for 2012 and 2016 detected the European countries for which the movement between clusters could be seen as a deterioration and those for which it can be seen as an improvement in some or all of the observed variables. These results have important implications for public policies aimed at reducing death of enterprises and unemployment.

KEY WORDS

CLASSIFICATION

*Corresponding author, η: bzmuk@efzg.hr; +385 1 238 3372;
 Trg J.F. Kennedyja 6, HR – 10000 Zagreb, Croatia

INTRODUCTION

There is no doubt that new business creation is important for an economy. The business enterprise population for the 28 European Union countries in 2016 consisted of almost 27 million active enterprises, which employed about 150 million people. The share of newly born enterprises increased by 3,5 % in 2016 in comparison to 2015 and resulted in the total of 2,6 million enterprises births in 2016, which generated about 4,0 million jobs [1]. New business creation is considered as one of the main factors of innovation, economic productivity and job creation. Many empirical researches have confirmed the role that newly born enterprises have in stimulating the competitiveness and efficiency of other enterprises within a country.

Understanding the specific circumstances that affect the scale and the determinants of business demography is therefore very important for development of policies that enhance entrepreneurship and increase the level of employment – one of the main priorities of the European Union 2020 growth strategy [2]. The heterogeneity of birth and death of enterprises is context-dependent in nature. The business dynamics related policies and institutions may be a very important driving force for information and communication technology (ICT) related sectors especially since more significant cross-country differences in enterprise birth and death rates are present in younger than in mature ICT sectors [3]. According to the research conducted by [4], which examined the impact of business policies and laws on birth rates of enterprises across countries, the policies which restrain large enterprises competitiveness also encourage more new enterprises births.

Measurement of the differences in entrepreneurial activities across countries in this research is based on enterprise level indicators. The application of the enterprise approach in measuring enterprises dynamics such as the birth and death of enterprises provides a high consistency in methods and definitions of business events across European countries. Enterprise birth rates can therefore be used as a measure of entrepreneurial activity.

Hence, based on the data on the total number of active enterprises and enterprises birth and death rates, the first aim of this article is to analyse the variations and changes in business statistics indicators, as well as the accompanying effects that these events have on the unemployment rate and gross domestic product per capita for 28 (for the year 2012) and 29 (for the year 2016) European countries. For that purpose, the methods of descriptive statistics were used. By applying the methods of hierarchical cluster analysis, the second aim of the is article to investigate the changes that occurred in these two years regarding the grouping of European countries with similar characteristics based on the relative business demography variables and the chosen macroeconomic indicators. That way, the main stakeholders of this research, which are governments of the observed countries and leading macroeconomists, can, based on the results presented in this article and selected macroeconomics variables, form new regulations and give additional instructions which could lead to an increased level of wellbeing in their countries.

This article is organized as follows. After a brief introduction, the second section of this article presents the review of the relevant literature. The data and methods are described in the third section. The analysis, along with the discussions, is presented in section four. The article ends with final conclusions summarized in section five.

LITERATURE REVIEW

BUSINESS DEMOGRAPHY AND JOB CREATION

Previous research on the impact of business demography activities on job creation considers the birth of new enterprises as one of the key factors of employment creation and productivity of an economy [5-9], and the death of enterprises as a crucial factor in the process of "creative destruction" [10].

Every year in most OECD countries about 20 % of enterprises enter and exit the market and within the first two years about 20-40 % of newly born enterprises exit and only about 50 % of the newly born enterprises are still active after 5 years [11]. Nevertheless, the birth of new enterprises is still the most important generator of employment growth [7]. The number of jobs created by newly born enterprises outnumbers the employment lost due to the death of enterprises in all size classes [7]. According to [9] in the EU context, smaller enterprises contribute more to employment creation than larger enterprises do.

BUSINESS DEMOGRAPHY AND UNEMPLOYMENT

Business demography indicators are sensitive to changes in macroeconomic environment. Previous research regarding the relationship between unemployment and the birth of enterprises proves a significant relationship between these variables [12-15]. However, the empirical research often shows contradictory results regarding the direction of the relationship between unemployment and the birth of enterprises. A negative relationship would be expected: an increase in unemployment should be followed by a decrease in the number of new enterprises. According to [12] unemployment is mainly negatively related to birth and death rates in Italy, but this relationship is also sector dependent. An update of that study conducted by [13] revealed that a positive impact of unemployment on the net birth rate is due to a negative impact of unemployment on the firm death rate. For European countries, [16] found a significant positive correlation between the death rate of enterprises and the unemployment rate and a significant negative relationship for the unemployment rate and the difference between the birth and the death rate of enterprises (enterprise dynamics). The authors in [17] found a negative relationship between the number of newly born enterprises and unemployment for the EU countries.

The research performed by [14] empirically confirmed the existence of two different relationships between unemployment and self-employment for the OECD countries: the "entrepreneurial" effect whereby higher rates of self-employment may lead to a rise of entrepreneurial activity and a reduction of unemployment was found to be much stronger than the "refugee" effect of start-ups whereby high unemployment rates may be responsible for start-up activity of self-employed people. In a more recent study a positive and statistically significant relationship between self-employment and unemployment was found for more than 50 per cent of the European OECD countries, but the presence of a negative or statistically insignificant relationship was also reported [15].

The use of different indicators and/or application on different units of analysis (different industries, different countries, regions, different cross-country context or different time) may be the reason for these mixed research results on the relationship between unemployment and the birth and death of enterprises.

ENTREPRENEURSHIP AND ECONOMIC DEVELOPMENT

There is no doubt that entrepreneurial activities and economic development are closely related. The studies on the relationship between entrepreneurship and economic development (measured as GDP per capita) are generally also controversial. It has been found that the GDP growth rate has an impact on start-up rates [18, 19] and the authors [17] explored the relationship between GDP growth and the absolute number of newly-born enterprises for the EU countries by applying a random-effects negative binomial regression model and they found that the number of newly-born enterprises is positively related to GDP growth. If income is measured as GDP per capita, the impact on entrepreneurship could be negative [20] as well as positive [6, 8] and some authors suggest a U-shaped relationship for GDP per

capita and entrepreneurship [5, 21]. The level of economic development fosters new-firm formation [22] and it is after the countries reach a certain level of income that a higher GDP per capita is related to higher start-up rates [6].

CLUSTER ANALYSIS IN ENTREPRENEURSHIP

Cluster analysis is a widely used multivariate method for data analysis in which data are grouped into different clusters based on a set of variables, in the way that similar objects are placed in the same cluster. Many researchers in the area of business enterprises also use this method.

For instance, the author [23] explored the differences in ICT adoption among enterprises in different groups of European countries (innovation leaders and followers, moderate and modest innovators). Authors [24] used a cluster analysis in order to confirm their results of the discriminant analysis carried out for the purpose of examination of the differences between applied funding policies in wine industry enterprises in France and Hungary.

One of the examples of previous research that uses the methods of cluster analysis in entrepreneurship and business demography is the research conducted by [25] on the role that national culture plays as an explanatory variable alongside entrepreneurial variables for the level of economic development. According to this research, national culture reinforces the impact of entrepreneurship on economic development: the used variables explain over 60 % of the variance in the linear regression model for GDP per capita. By applying the cluster analysis for the European Union countries the same author [25] found four entrepreneurial clusters, each characterized by a different entrepreneurial dynamics explained to some extent by cultural value variables and income as measured by GDP per capita. By performing a cluster analysis on variables connected to economic development, institutions and entrepreneurship for the European Union countries the authors [26] examined the entrepreneurial performance of transition economies with a focus on productive entrepreneurship. In the study conducted by [27] five entrepreneurial types of European countries were distinguished by using principal component analysis and hierarchical cluster analysis on variables related to entrepreneurial activity and economic development. Their research showed that the diversity in entrepreneurial activities can be explained and discriminated by five explanatory themes on innovation, employment, formal institutions, entrepreneurship and governance.

Our research takes an approach in line with previous research [27] but explores exclusively the business demography activity indicators, the death and birth rates of enterprises as classifying variables for the cluster analysis along with the economic environment indicators – GDP per capita and the unemployment rate [28]. We believe that our research contributes to the existing empirical literature by using the most recent data set on business demography indicators for selected European countries from 2016 and by presenting a comparison of the results for this data set with older data from 2012.

DATA AND METHODS

In the article, a focus is given to three selected business demography variables, listed with their brief descriptions in Table 1. The values of business demography variables have been observed for the total number of enterprises excluding holding companies. In addition, to set those business demography variables into the context, two macroeconomic variables related to the business demography are included in the analysis as well.

Table 1. List of observed variables.

Variable group	Variable code	Variable description
Business demography	ActiveEnterprises	Population of active enterprises in t (number)
	BirthRate	Birth rate: number of enterprise births in the reference period (t) divided by the number of enterprises active in t (percentage)
	DeathRate	Death rate: number of enterprise deaths in the reference period (t) divided by the number of enterprises active in t (percentage)
Macroeconomics	GDPCapita	GDP per capita, PPP (constant 2011 international $)
	UnempRate	Unemployment, total (% of total labour force) (modelled ILO estimate)

In order to collect the data for the selected variables, Eurostat and the World Bank databases have been used [29-31]. The data from 2012 and 2016 are going to be observed in the analysis. The analysis is going to be conducted using the sample of European countries for which data were available. Consequently, the following 29 European countries have been included in the analysis: Austria. Belgium, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, North Macedonia, Norway, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom. Out of the 29 observed countries, 26 are the European Union member states. Unfortunately, the data for Greece are missing in 2012 and therefore the count of the observed countries in 2012 is equal to 28.

The observed variables are going to be first analysed by using the descriptive statistics approach. Within the descriptive statistics analysis, variables are going to be plotted, descriptive statistics measures are going to be calculated and relations between the variables are going to be observed. In addition, the outlier analysis will be conducted to check whether there is a country (or more of them) that has values of the observed variables significantly lower or higher in comparison to other observed countries.

A hierarchical cluster analysis will be performed to detect groups of countries with similar characteristics according to the observed variables. Overall two hierarchical cluster analyses will be conducted. The first one will include data from 2012 and the second one will use data from 2016. In order to get sensible results, standardized values of relative business demography measures, *BirthRate* and *DeathRate*, are going to be considered. In both hierarchical cluster analyses, both observed macroeconomic variables will be included as well.

ANALYSIS AND DISCUSSION

DESCRIPTIVE STATISTICS ANALYSIS

The observed European countries differ in both their land and their population sizes. Consequently, it is expected that the number of active enterprises is quite different among the observed countries. Therefore, the number of active enterprises in the observed countries in 2012 and in 2016 is shown in Figures 1 and 2, respectively.

Figures 1 and 2 confirmed that there are large differences in the number of active enterprises among the observed European countries. Five countries stand out: Italy, France, Spain, Germany and the United Kingdom. Those five countries have convincingly the largest number of active enterprises among all the observed countries. On average the observed countries
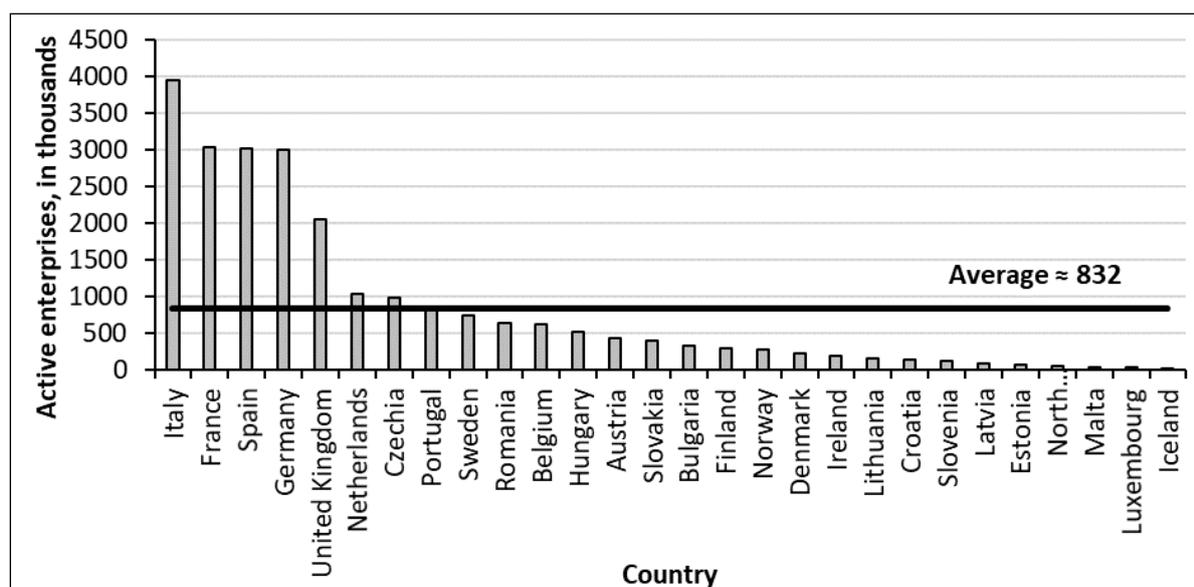
Figure 1. Number of active enterprises in the observed 28 European countries, in thousands, in 2012 [28].
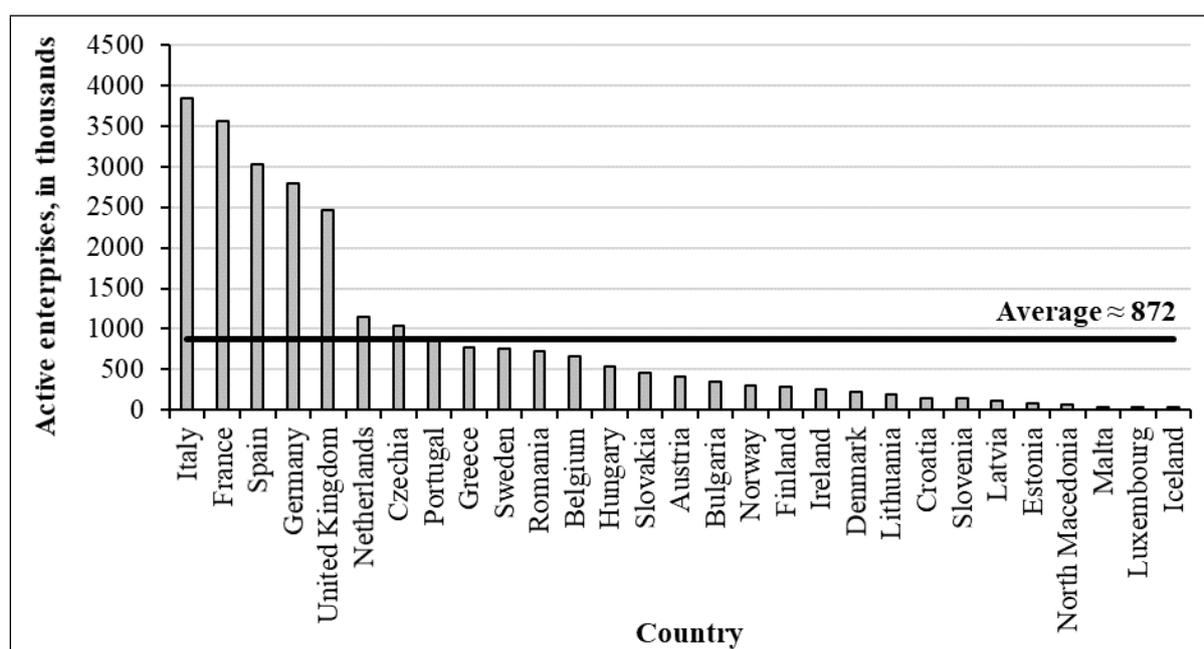

Figure 2. Number of active enterprises in the observed 29 European countries, in thousands, in 2016 [28].

had about 832 thousands of active enterprises in 2012 and about 872 thousands in 2016. Along with the five mentioned countries, the Netherlands and Czechia are the only countries which have the number of active enterprises above the average value in both observed years. All other countries have the number of active enterprises below the average. Such relations suggest that a researcher should be careful when the business demography is observed by taking only absolute values into account. The absolute indicators of the total enterprises births (or deaths) are also less comparable across countries if they are based on different definitions of enterprises. The data on business demography variables used for analysis in this article are already highly consistent in methods and definitions of business events across European countries, but comparability can be additionally improved by using relative business

demography indicators (enterprise birth or death rates) since biases (if they exist in the absolute values) are then present in both the numerator and the denominator [32]. Therefore, in the article the focus will be given to relative variables *BirthRate* and *DeathRate* whereas, for the purpose of comparison of different countries only descriptive statistics results of the variable *ActiveEnterprises* will be shown.

Table 2. Descriptive statistics of the observed variables for all the observed European countries, in 2012 and in 2016 [29-31].

Variable code	Mean	St. Dev.	Coef. Var.	Skewness	1st Quar.	Median	3rd Quar.
2012							
ActiveEnterprises	832 055	1103 549	132 63	1,74	142 871	364 057	852 421
BirthRate	10,27	3,76	36,64	2,34	8,14	9,81	11,69
DeathRate	9,69	3,45	35,60	0,55	7,32	9,34	11,39
GDPCapita	34 932	15 711	44,98	1,58	25,281	33,168	42 956
UnempRate	10,58	6,14	58,03	1,81	6,67	8,41	13,51
2016							
ActiveEnterprises	872 107	1116,872	128,07	1,69	147 181	406 079	843 693
BirthRate	10,38	3,50	33,70	0,72	7,71	9,95	11,64
DeathRate	7,89	3,24	41,03	0,49	6,14	8,04	9,41
GDPCapita	37,268	16 442	44,12	1,67	27,124	34,655	44 464
UnempRate	8,78	5,29	60,28	1,86	5,90	7,57	9,67

Table 2 shows main descriptive statistics results of the observed variables in 2012 and in 2016. According to the coefficient of variation values it can be concluded that at the absolute business demography variable *ActiveEnterprises* a much larger variation level is present than at relative business demography variables *BirthRate* and *DeathRate*. The descriptive statistics results also show a moderately high variability of variables related to business demography activities (*BirthRate* and *DeathRate*) as well as to the indicator of economic development (*GDPCapita*) and a relatively high variability in UnempRate in both observed years. Hence, these results indicate a relatively high heterogeneity between the European countries, which did not change much in 2016 in comparison to 2012. However, the observation of the average values of variables reveals that the total number of active enterprises increased in 2016 in comparison to 2012. The birth rate of enterprises in 2016 had almost the same average value as in 2012, but the average death rate in 2016 was lower than in 2012 (with only a little higher dispersion of death rates in 2016 than in 2012). The average GDP per capita was higher in 2016 than in 2012 with about the same variation in both years and the average unemployment rate in 2016 was lower than the average unemployment rate in 2012 with only a slight increase in variation of unemployment rates among European countries in 2016. The range of variation in unemployment rates for the middle 50 % of European countries was much larger in 2012 (interquartile range 6,84) than in 2016 (interquartile range 3,77) which may indicate that the situation regarding unemployment in European countries was much better in 2016 than in 2012. Finally, these results may indicate an overall positive economic and business demography change that occurred among European countries in 2016 relative to 2012.

Table 3 presents the list of the top three and the last three countries according to the value of each observed variable. In both years Lithuania recorded the highest birth rates of enterprises, but its birth rate decreased in 2016 compared to 2012. It is interesting to notice that Malta, which was among the three countries with the lowest birth rate in 2012 moved to the list of the three countries with the highest birth rate in 2016, while being among the three countries with the lowest death rate in 2016. Among countries with the highest levels of GDP per capita,

Table 3. Top three and last three countries according to the values of the observed variables, in 2012 and in 2016 [29-31].

Variable code	Top three countries	Last three countries
2012		
ActiveEnterprises	Italy (3953 714); France (3 039 203); Spain (3012 443)	Malta (31 427); Luxembourg (29 122); Iceland (24 164)
BirthRate	Lithuania (24,88); Latvia (16,83): Bulgaria (13,02)	Ireland (6,76); Belgium (6,22); Malta (6,20)
DeathRate	Lithuania (18,65); Portugal (16,02); North Macedonia (14,83)	France (5,64), Norway (4,18); Belgium (3,12)
GDPCapita	Luxembourg (89 505); Norway (63 003); the Netherlands (45 949)	Romania (18 361); Bulgaria (15 772); North Macedonia (11 550)
UnempRate	North Macedonia (31,02); Spain (24,79); Croatia (15,94)	Luxembourg (5,14); Austria (4,86); Norway (3,12)
2016		
ActiveEnterprises	Italy (3849 594); France (3558 735); Spain (3026 237)	Malta (36 215); Luxembourg (32 391); Iceland (27 653)
BirthRate	Lithuania (18,82); Malta (16,36); Latvia (16,16)	Austria (6,59); Belgium (6,24); Greece (4,51)
DeathRate	Bulgaria (16,79); Portugal (14,84); Denmark (10,65)	Belgium (3,33); Malta (3,04); Ireland (0,96)
GDPCapita	Luxembourg (94 921); Norway (64 160); Ireland (63 227)	Romania (21 782); Bulgaria (17 793); North Macedonia (13 113)
UnempRate	North Macedonia (23,72); Greece (23,54); Spain (19,63)	Germany (4,12); Czechia (3,95); Iceland (2,98)

*values of the variables are shown in brackets

in 2016 Ireland replaced the Netherlands which was in the third place in 2012. The list of the last three countries according to the GDP per capita level remained the same but their GDP per capita values were higher in 2016 than in 2012. The list of countries with the highest levels of unemployment rates was led by North Macedonia in both years, but its unemployment rate decreased in 2016 compared to 2012. The range of variation in unemployment rates was 28,9 in 2012, but only 20,74 in 2016.

In order to check whether there are some countries that have significantly higher or lower variable values than the other countries, all variables are first standardized. After that those standardized variables are plotted by using box plots. In this case the width of the "box" in the box plots is defined by the values of the 1st and the 3rd quartiles whereas the whiskers range is defined by the last actual data point which can be found in the 1,5 interquartile range from those two quartiles. All other data points, which turned out to be outside that range, are considered to be potential outliers. The box plots of all the observed variables are shown in Figure 3.

According to Figure 3 there are many data points that could be considered as potential outliers. In addition, there are some countries (Lithuania, Luxembourg and North Macedonia) whose values for some variables deviate from the averages more than three standard deviations. However, the outlier problem seems to be improved in 2016 since only Luxembourg remained a potential outlier in that year. Hence, in order to be able to compare the situation in the observed periods it has been decided that no countries will be omitted Table 4 shows the matrix of Pearson's correlation coefficient values among the observed variables in 2012 and in 2016. It turned out that a statistically significant positive correlation between variables *BirthRate* and *DeathRate* was present in 2012, meaning that the European
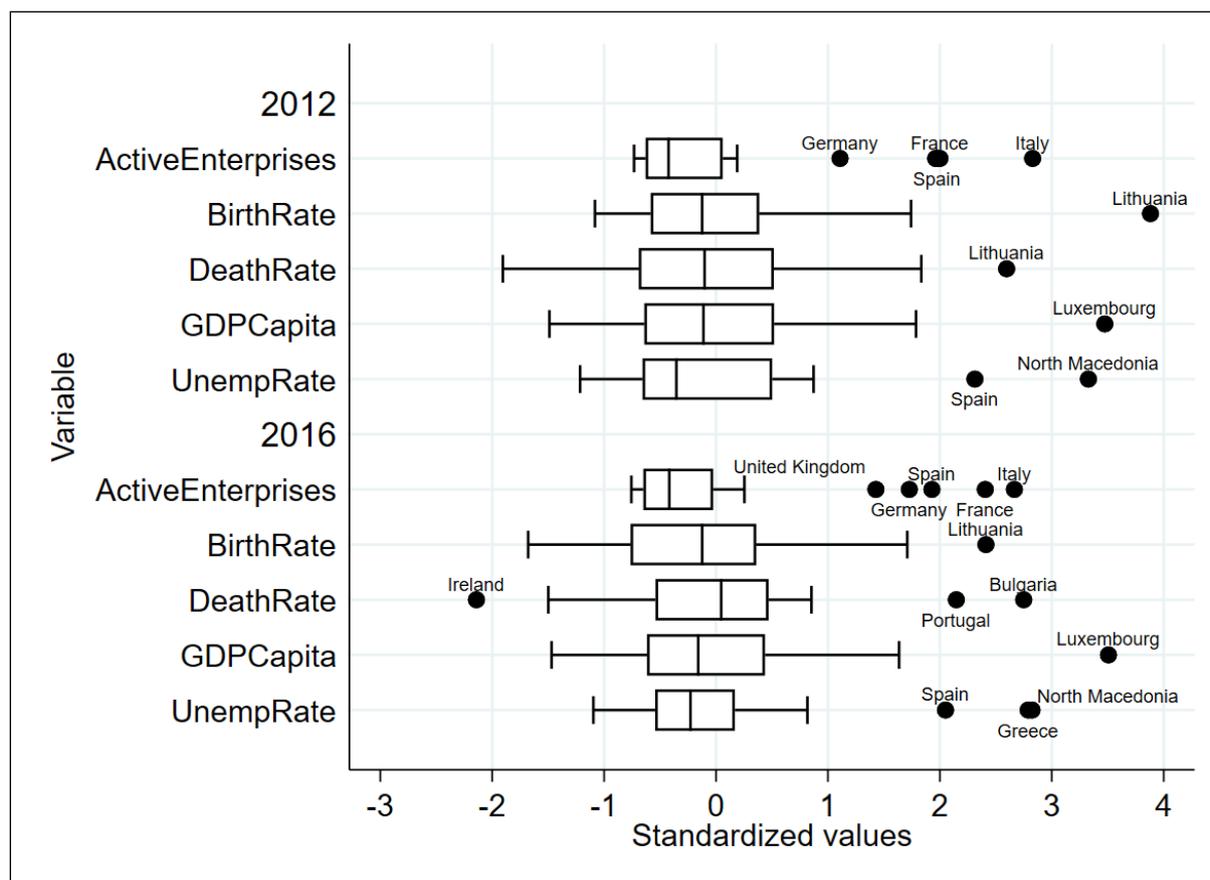
Figure 3. Box plots of the standardized values, in 2012 and in 2016 [29-31].

Table 4. Correlation matrix of the observed variables, Pearson's correlation coefficient, in 2012 ($n = 28$) and in 2016 ($n = 29$) [28-30].

Variable code	ActiveEnterprises	BirthRate	DeathRate	GDPCapita	UnempRate
2012					
ActiveEnterprises	1,0000				
BirthRate	−0,2396	1,0000			
DeathRate	−0,2347	0,7181*	1,0000		
GDPCapita	0,0359	−0,3259	−0,5788*	1,0000	
UnempRate	0,0367	0,1978	0,5035*	−0,5432*	1,0000
2016					
ActiveEnterprises	1,0000				
BirthRate	−0,1880	1,0000			
DeathRate	0,0224	0,3298	1,0000		
GDPCapita	−0,0226	−0,2674	−0,4370*	1,0000	
UnempRate	0,1353	−0,2120	0,1140	−0,4057*	1,0000

*statistically significant correlations at the significance level 0,05

countries that had higher birth rates of enterprises also had higher death rates at that time. However, in 2016 this relationship between birth and death rates for European countries is no longer evident, since the coefficient of correlation for these variables is statistically insignificant. The variable *DeathRate* had a statistically significant negative correlation with the variable *GDPCapita* and a positive correlation with the variable *UnempRate* in 2012, meaning that the European countries with higher death rates also had lower GDP per capita levels and higher unemployment rates. However, the variable *DeathRate* had a statistically significant negative correlation only with the variable *GDPCapita* in 2016, but not with the

variable *UnempRate*, which may also be seen as an indicator of change in determination of unemployment through the death rate of enterprises in 2016. In both observed periods a statistically significant negative correlation was present between variables *GDPCapita* and *UnempRate*.

HIERARCHICAL CLUSTER ANALYSIS

In the hierarchical cluster analysis Ward's clustering method will be applied. As a distance measure, the squared Euclidean distances will be used. Two hierarchical cluster analyses will be conducted. The first one by using data for 28 European countries (without Greece) from 2012, and the second one based on data for 29 European countries (with Greece) from 2016. In both cluster analyses standardized values of variables *BirthRate*, *DeathRate*, *GDPCapita* and *UnempRate* will be used to form clusters or groups of countries with similar characteristics according to the observed variables. In order to select the most appropriate or optimal number of clusters the Calinski and Harabasz pseudo-F index [33] and the Duda-Hart 01) Je(2)/Je(1) index F [34] have been consulted. High values of the Calinski and Harabasz pseudo-F index and the Duda-Hart Je(2)/Je(1) index F indicate distinct clustering. Along with the Duda-Hart Je(2)/Je(1) index F, the associated pseudo T-squared measures can be found as well. A lower pseudo T-squared indicates distinct clustering.

Table 5. The Calinski and Harabasz pseudo-F index and the Duda-Hart Je(2)/Je(1) index F values, data related to 2012 [29-31].

Number of clusters	Calinski and Harabasz pseudo-F index	Duda-Hart	
		Je(2)/Je(1) index F	Pseudo T-squared
2	14,91	0,6180	6,18
3	13,40	0,5500	11,45
4	13,23	0,6580	4,68
5	13,93	0,2048	3,88
6	14,52	0,5666	9,18
7	**15,51**	**0,5908**	**4,16**
8	16,42	0,3602	8,88

Figure 4 shows the dendrogram for the cluster analysis based on data from 2012, whereas Table 5 presents the results of the Calinski and Harabasz pseudo-F index and the Duda-Hart Je(2)/Je(1) index F, with the associated pseudo T-squared values, related to that cluster analysis. The Calinski and Harabasz pseudo-F index values are steadily increasing from the solution with 4 clusters to the solution with 8 clusters, where the maximum value is achieved. Consequently, the Calinski and Harabasz pseudo-F index pointed out that the solution with 8 clusters should be taken as the optimal solution. However, the value of the Duda-Hart Je(2)/Je(1) index F is significantly lower when using the solution with 8 clusters than using the solution with 7 clusters. Similarly, the solution with 7 clusters has a considerably lower pseudo T-squared value than the solution with 8 clusters. Because of that it has been decided that the optimal number of clusters in the cluster analysis of data from 2012 is equal to 7.

Table 6 shows the list of countries according to their cluster membership in the first cluster analysis using data from 2012 whereas Table 7 presents the average values of variables included in the cluster analysis separately for each cluster. According to Table 7 the countries from the cluster A5 have the lowest average *BirthRate* whereas the countries from the cluster A3 have the lowest average *DeathRate*. On the other hand, Lithuania which is set to be in the cluster A7 has the highest *BirthRate* and *DeathRate* value. North Macedonia, which is the only one in the cluster A6, has among the highest values of variables *BirthRate* and *DeathRate* but it has the lowest *GDPCapita* and the highest *UnempRate*.
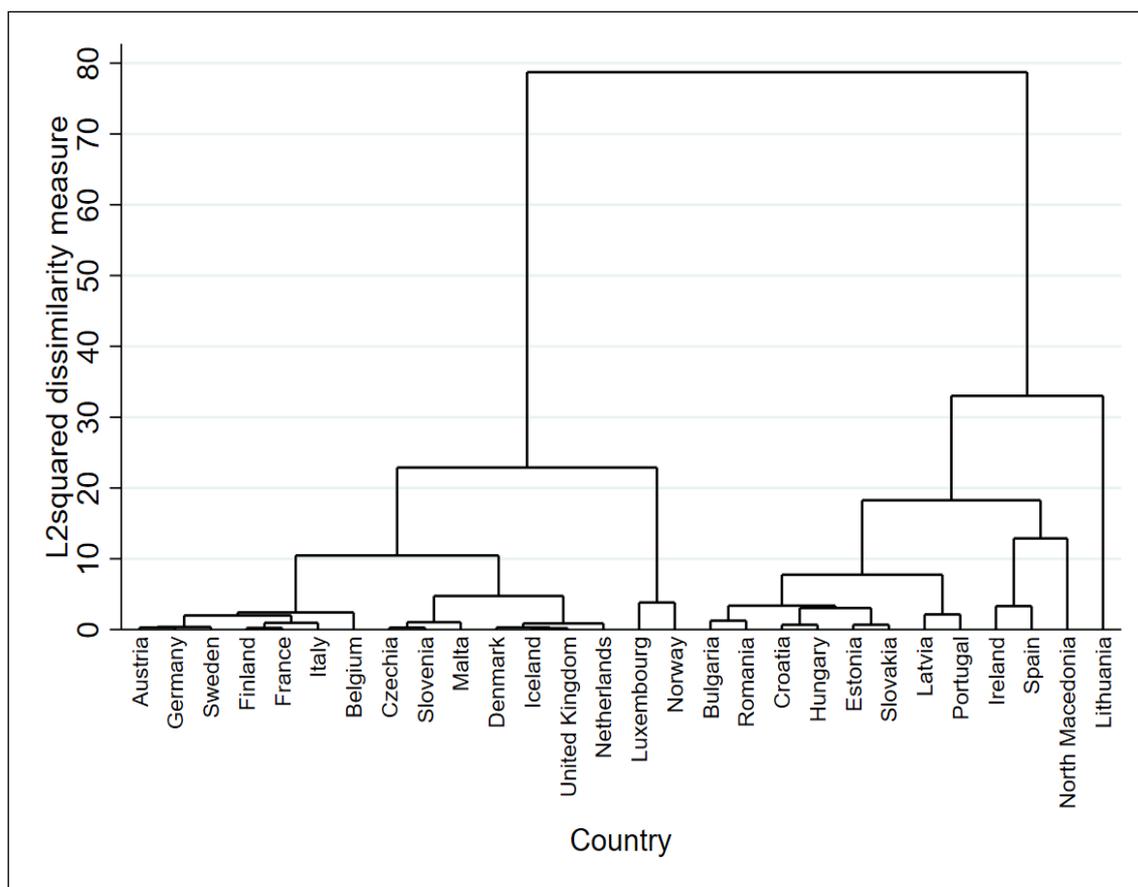
Figure 4. Dendrogram, included standardized values of variables *BirthRate*, *DeathRate*, *GDPCapita* and *UnempRate* for 28 European countries, Ward's clustering approach, the squared Euclidean distance, data related to 2012 [29-31].

Table 6. Members of clusters, data related to 2012 [29-31].

Cluster code	Countries
A1	Austria, Belgium, Finland, France, Germany, Italy, Sweden
A2	Czechia, Denmark, Iceland, Malta, the Netherlands, Slovenia, the United Kingdom
A3	Luxembourg, Norway
A4	Bulgaria, Croatia, Estonia, Hungary, Latvia, Portugal, Romania, Slovakia
A5	Ireland, Spain
A6	North Macedonia
A7	Lithuania

Table 7. Average values of variables included in the clustering process, data related to the 2012 analysis [29-31].

Cluster code	Number of countries	BirthRate	DeathRate	GDPCapita	UnempRate
A1	7	7,78	6,33	40,625	7,64
A2	7	9,93	9,92	36,298	7,05
A3	2	9,11	5,75	76,254	4,13
A4	8	11,77	11,77	21,951	12,57
A5	2	7,49	9,26	37,969	20,12
A6	1	11,40	14,83	11,550	31,02
A7	1	24,88	18,65	24,049	13,36

Figure 5 presents the dendrogram for the second cluster analysis, which is based on data from 2016. The values of the Calinski and Harabasz pseudo-F index, given in Table 8, suggest that it would be better to choose a higher number of clusters due to a steady increase of the index. ccording to that, the Duda-Hart Je(2)/Je(1) index F suggests choosing between the solution with 7 and 8 clusters (the highest, but very similar, values in the upper part of the table). Due to the fact that the solution with 8 clusters has a higher pseudo T-squared value, it has been decided that the optimal number of clusters is equal to 7.
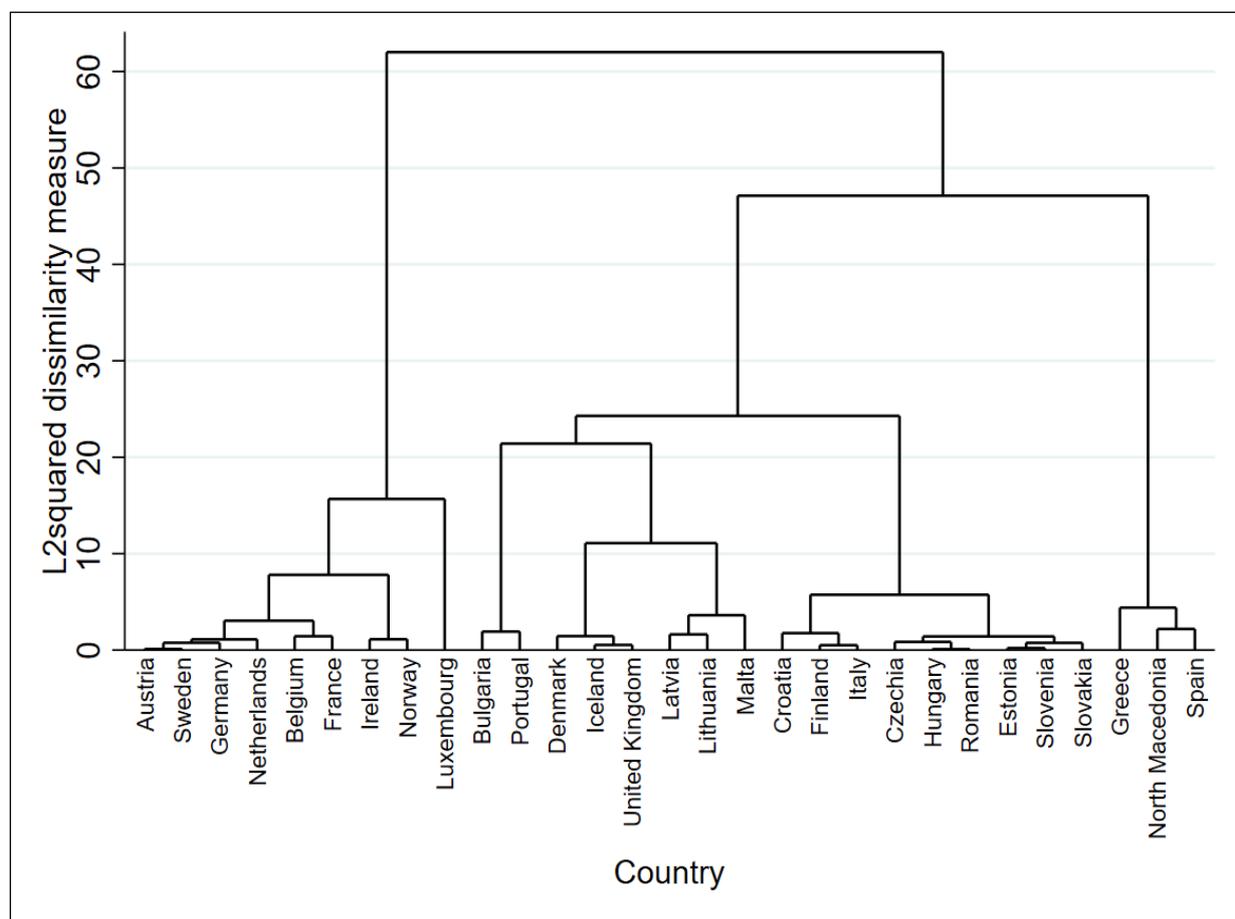

Figure 5. Dendrogram, included standardized values of variables *BirthRate*, *DeathRate*, *GDPCapita* and *UnempRate* for 29 European countries, Ward's clustering approach, the squared Euclidean distance, data related to 2016 [29-31].

Table 8. The Calinski and Harabasz pseudo-F index and the Duda-Hart Je(2)/Je(1) index F values, data related to 2016 [29-31].

Number of clusters	Calinski and Harabasz pseudo-F index	Duda-Hart	
		Je(2)/Je(1) index F	Pseudo T-squared
2	10,34	0,6401	10,12
3	12,35	0,6855	6,88
4	12,27	0,4858	6,35
5	13,43	0,4953	7,13
6	14,66	0,3943	6,15
7	**15,70**	**0,4928**	**6,17**
8	16,42	0,4922	7,22

Table 9 contains the list of countries according to their cluster membership in the cluster analysis based on data from 2016 whereas Table 10 presents average values of variables included in the cluster analysis for each cluster separately.

Table 9. Members of clusters, data related to 2016 [29-31].

Cluster code	Countries
B1	Austria, Belgium, France, Germany, Ireland, the Netherlands, Norway, Sweden
B2	Luxembourg
B3	Bulgaria, Portugal
B4	Denmark, Iceland, the United Kingdom
B5	Latvia, Lithuania, Malta
B6	Croatia, Czechia, Estonia, Finland, Hungary, Italy, Romania, Slovakia, Slovenia
B7	Greece, North Macedonia, Spain

Table 10. Average values of variables included in the clustering process, data related to 2016 [29-31].

Cluster code	Number of countries	BirthRate	DeathRate	GDPCapita	UnempRate
B1	8	7,73	4,92	48,783	6,70
B2	1	8,93	7,84	94,921	6,29
B3	2	14,04	15,82	22,458	9,32
B4	3	13,70	10,05	43,644	4,66
B5	3	17,11	5,86	28,943	7,40
B6	9	9,46	8,51	29,142	8,11
B7	3	8,25	8,56	23,548	22,30

According to Table 10, countries from the cluster B1 on average have the lowest birth and death rate of enterprises. On the other hand, the countries from the cluster B5 on average have the highest birth rates of enterprises whereas the countries from the cluster B3 on average have the highest death rates of enterprises.

DISCUSSION

Theoretical implications

The descriptive analysis results indicate that a very high variation level is present at the absolute indicator of business demography (*ActiveEnterprises*) and that the variation level in relative business demography indicators (*BirthRate* and *DeathRate*) is much lower but still moderately high. The results also show a relatively high variability of macroeconomic indicators (*GDPCapita* and *UnempRate*) in both observed years. High variability in economic and business demography indicators suggests the existence of different patterns in these variables among the European countries. However, even though a relatively high heterogeneity of European countries did not change much between these two years, in 2016 the average value of death rates was lower, the average GDP per capita was higher and the average unemployment rate was lower than in 2012, and only the enterprise birth rates had almost the same average value as in 2012. Thus, these results indicate an overall positive economic and business demography change that occurred among European countries in 2016 relative to 2012.

The analysis of the relationships between the variables revealed that the European countries on a lower level of development (measured as GDP per capita) also tend to have higher death rates of enterprises and higher unemployment rates, and that this statistically significant correlation was present in both years. Though, the correlation analysis also revealed that a

statistically significant and highly positive relationship between birth and death rates of enterprises found in 2012 was absent in 2016. The existence of a statistically significant and positive correlation between birth and death rates of enterprises would be expected, since this describes the process of *creative destruction*, through which dynamic and innovative new enterprises drive inefficient ones out of the market [10]. The same change in statistical significance of the correlation was found between death rates of enterprises and the unemployment rate, which was significantly positive in 2012, but insignificant in 2016. These results suggest that the underlying mechanisms of the correlation between birth and death rates of enterprises and between death rates of enterprises and unemployment rates have changed.

The expected significant relationship between unemployment and birth of enterprises [12-15, 17] was not confirmed in our results in either year. In accordance with the research conducted by [16], we found a significant positive correlation between the death rate of enterprises and the unemployment rate, but only in 2012. The insignificant linear correlation between the birth rate of enterprises and the level of economic development (as measured by the GDP per capita level) was expected, since the previous researches [5, 6, 21, 22] report on the more complex functional form of the relationship between these variables, which reflects the findings that it is only after a country reaches a certain level of development (income) that a positive relationship between income and start-up rates is found.

The hierarchical cluster analysis based on four classifying variables: business demography activity indicators – the birth and the death rate of enterprises, and macroeconomic indicators – GDP per capita and the unemployment rate resulted in 7 clusters of European countries each described by its own combination of characteristics related to the used variables in both observed year.

Practical implications

The comparison of cluster solutions reveals that clusters A1 and A3 in 2012 and clusters B1 and B2 in 2016 gather countries on a high or a relatively high level of development in terms of GDP per capita, which also have a low or a relatively low unemployment rate and whose main business demography features are a low or a relatively low birth and death rate of enterprises. Similar characteristics of macroeconomic indicators were found in clusters A2 and B4: these were also countries on a high or a relatively high level of development, but these countries had high or relatively high birth and death rates of enterprises. The cluster A5 gathered countries similar according to a relatively high level of development, a low or a relatively low birth and death rate on average, but a very high unemployment rate. This combination of business demography and macroeconomic features was not found in any of the clusters in 2016.

The clusters that gathered countries on a low or a very low level of development with also a relatively high or a very high unemployment rate are A4, A6 and A7 in 2012 and clusters B3 and B7 in 2016. Among these clusters, A4 and B3 have high birth and death rates, the cluster B7 has a relatively low birth rate but a relatively high death rate, the cluster A6 has the lowest GDP per capita and the highest unemployment rate among all the clusters, the cluster A7 has the highest birth and death rates of enterprises and the cluster B7 has a relatively low birth rate but a relatively high death rate with the highest unemployment rate among all the clusters in 2016.

Countries that are joined in the clusters B5 and B6 share a moderately low level of development. However, the cluster B5 has a relatively low death rate and unemployment rate and the highest birth rate among all the clusters in 2016, while the cluster B6 has a relatively low birth rate of enterprises but a relatively high death rate and unemployment rate.

By comparing the resulting clusters for 2012 and 2016 with similar main characteristics we observe some interesting country-specific movements between clusters. Since the main

features of clusters A1 and B1 are a low birth and death rate of enterprises, high GDP per capita and a relatively low unemployment rate, it is interesting to notice that some countries with these characteristics that were present in the cluster A1 are no longer present in the cluster B1 in 2016. Namely, Finland and Italy, formerly in A1 in 2012 were found in the cluster B6 in 2016, which is characterized by a still relatively low birth rate but a moderately high death rate of enterprises, a relatively low level of economic development (as measured by GDP per capita) on average and a moderately high unemployment rate on average that is still lower than the overall sample unemployment rate average for 2016. For Finland and Italy this shift from a cluster with more desirable characteristics in 2012 to a cluster with significantly less desirable characteristics in 2016 could mean a stagnation or even deterioration of business demography activities and macroeconomic performance.

A similar conclusion could be derived for Czechia and Slovenia, both members of the cluster B6 in 2016, since these countries were members of the cluster A2 in 2012, which is characterized by a moderately high birth rate, a moderately high death rate, a relatively high GDP per capita but a relatively low unemployment rate.

For the Netherlands, joining the cluster B1 in 2016 could be seen as an improvement in the observed variables since in 2012 this country was the member of the cluster A2. An improvement is also evident for Ireland which is found in the cluster B1 in 2016 since in 2012 it was a member of the cluster A5 characterized by a low birth rate, a relatively low death rate, a relatively high GDP per capita but a high unemployment rate.

For Lithuania, which was a potential outlier in 2012 since it grouped alone that year in the cluster A7 with the highest birth and death rate of enterprises, a relatively high unemployment rate and a low GDP per capita, joining the cluster B5 in 2016 could be seen as a positive movement. Countries in the cluster B5 still have the highest birth rate and on average a relatively low GDP per capita in 2016, but also have a low death rate of enterprises and a relatively low unemployment rate. Grouping in the cluster B5 can be seen as a positive movement between clusters for Malta (even though this country groups with significantly less developed countries in 2016) since Malta was in the cluster A2 in 2012.

Bulgaria and Portugal were found in cluster A4 in 2012, characterised by a high birth and death rate of enterprises, a relatively high unemployment rate and a low GDP per capita. In 2016 these two countries formed the cluster B3 whose main features are the highest death rate of enterprises and the lowest GDP per capita on average among all the clusters in 2016 and also a high birth rate of enterprises and a relatively high unemployment rate. For Bulgaria and Portugal this separation from other countries could mean a lack of adequate mechanisms for correction of undesirable characteristics related to the observed variables.

North Macedonia formed a 1-unit cluster A6 in 2012 and was a potential outlier that year since it had a relatively high birth and death rate of enterprises followed by the lowest GDP per capita and the highest unemployment rate. However, in 2016 North Macedonia clustered with Spain and Greece in the cluster B7, characterized by a relatively low birth rate but still a relatively high death rate, low GDP per capita and the highest unemployment rate on average among all the clusters in 2016. Hence, the fact that North Macedonia lost the status of an outlier in 2016 could be seen as an improvement in the observed indicators of business demography activities and economic development. On the other side, for Spain, a shift from the cluster A5 to the cluster B7 could point to some issues regarding the death of enterprises. Finally, even though Luxembourg was a potential outlier in 2012, this country was grouped with Norway in the cluster A3 that year, but in 2016 Luxemburg confirmed its outlier status by excluding itself from all other countries with far the highest GDP per capita level.

CONCLUSIONS

In this study the descriptive statistics methods and the hierarchical cluster analysis approach were used to explore the features of European countries regarding the business demography activities based on the data on the total number of active enterprises, enterprises birth rates and enterprises death rates and regarding the chosen macroeconomic indicators based on the level of economic development (measured as GDP per capita) and the unemployment rate. This research uses the data for 28 European countries in 2012 and for 29 European countries (the same 28 countries as in 2012 plus Greece) in 2016. Our results indicate a presence of a relatively high heterogeneity of countries in both observed periods relative to business demography indicators and macroeconomic indicators. However, by observing a lower average value of death rates and a higher average value of birth rates of enterprises alongside with a higher average value of GDP per capita and a lower value of unemployment rate, it is concluded that a positive economic and demographic change occurred in 2016 relative to 2012. The correlation analysis shows that the significance of the relationship between business demography activities and economic development indicators in 2012 changed in 2016: the death rate of enterprises had a statistically significant negative correlation with GDP per capita in both years and a significant positive correlation with the unemployment rate and with the birth rate of enterprises only in 2012. The hierarchical cluster analysis based on Ward's clustering approach and the squared Euclidean distances resulted in 7 distinct groups of countries in both years with respect to their enterprise birth rate, enterprise death rate, GDP per capita level and unemployment rate.

Previous research applying cluster analysis in this area mainly focused on the classification of groups of countries due to one or more measures of entrepreneurship such as the birth of new enterprises (number or rate), Global Entrepreneurship Monitor's TEA Index, Global Entrepreneurship and Development Index [25, 26] and others. In line with a recent research conducted by [27], our research excludes other measures of entrepreneurship but includes the death rate of enterprises in the cluster analysis in order to explore its capability to differentiate the groups of countries alongside the birth rate of enterprises and macroeconomic indicators.

All resulting clusters in 2012 and in 2016 are described in terms of low, relatively low or high, and high average values of demography statistics indicators and macroeconomic indicators. The country-specific situation regarding the movement between clusters the countries participated in 2012 to the ones they formed in 2016 is presented. For some countries this movement between clusters could be seen as a stagnation or deterioration of some or all of the observed business demography activities and/or macroeconomic performance. This is possibly the case for Finland and Italy (movement from A1 to B6), Czechia and Slovenia (from A2 to B6), Bulgaria and Portugal (from A4 to B3) and Spain (from A5 to B7). However, for some countries movement between clusters could be seen as an improvement in the observed indicators: for the Netherlands (from A2 to B1), Ireland (from A5 to B1), Malta (from A2 to B5), Lithuania (from being an outlier in A7 to B5) and North Macedonia (from being an outlier in A6 to B7). These results may have important practical implications for country-specific public policies aimed at (further) reduction of death of enterprises and unemployment.

Main limitations of this research are the lack of data on business demography indicators for Greece in 2012 (the data on these activities for Greece have been available only since 2015) and the outliers issue which was not treated deliberately in order to enable comparison of results for all the observed countries in 2012 and 2016. In future research the outlier problem should be resolved and the obtained results should be verified in accordance with the exclusion of outliers. Also, in future research the formal analysis of the differences across clusters in terms of the classifying variables should be performed and the main determinants that influence the movement of countries between clusters over time should be explored.

REFERENCES

[1] Eurostat: *Business demography statistics. Statistics explained*.
https://ec.europa.eu/eurostat/statistics-explained/index.php/Business_demography_statistics, accessed 28th April 2019,

[2] European Commission: *Europe 2020: A strategy for smart, sustainable and inclusive growth*.
Brussels, European Commission, 2010,
http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf, accessed 25th April 2019,

[3] Brandt, N.: *Business Dynamics in Europe*.
OECD Science, Technology and Industry Working Papers, No. 2004/01. OECD Publishing, Paris, 2004,
http://dx.doi.org/10.1787/250652270238,

[4] Nofsinger, J. and Zykaj, B.B.: *Business Policies and New Firm Birth Rates Internationally*.
Accounting and Finance Research **3**(4), 1-14, 2014,
http://dx.doi.org/10.5430/afr.v3n4p1,

[5] Wennekers, A.R.M.; van Stel, A.; Thurik, A.R. and Reynolds, P.D.: *Nascent Entrepreneurship and the Level of Economic Development*.
Small Business Economics **24**(3), 293-309, 2005,
http://dx.doi.org/10.1007/s11187-005-1994-8,

[6] Pinillos, M.J. and Reyes, L.: *Relationship between individualist–collectivist culture and entrepreneurial activity: evidence from Global Entrepreneurship Monitor data*.
Small Business Economics **37**(1), 23-37, 2011,
http://dx.doi.org/10.1007/s11187-009-9230-6,

[7] de Kok, J., et al.: *Do SMEs create more and better jobs*.
Brussels, European Commission, 2011,
http://ec.europa.eu/growth/sites/growth/files/docs/body/do-smes-create-more-and-better-jobs_en.pdf,
accessed 25th April 2019,

[8] Audretsch, D.B.: *Entrepreneurship research*.
Management Decision **50**(5), 755-764, 2012,
http://dx.doi.org/10.1108/00251741211227384,

[9] de Wit, G. and de Kok, J.: *Do small businesses create more jobs? New evidence for Europe*.
Small Business Economics **42**(2), 283-295, 2014,
http://dx.doi.org/10.1007/s11187-013-9480-1,

[10] OECD: *Entreprenurship at a Glance 2017*.
OECD Publishing, Paris, 2017,
http://dx.doi.org/10.1787/entrepreneur_aag-2017-en,

[11] Bartelsman, E.; Scarpetta, S. and Schivardi, F.: *Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries*.
Industrial and Corporate Change **14**(3), 365-391, 2005,
http://dx.doi.org/10.1093/icc/dth057,

[12] Caree, M.; Santarelli, E. and Verheul, I.: *Firm entry and exit in Italian provinces and the relationship with unemployment*.
International Entrepreneurship and Management Journal **4**(2), 171-186, 2008,
http://dx.doi.org/10.1007/s11365-007-0060-1,

[13] Santarelli, E.; Carree, M. and Verheul, I.: *Unemployment and firm entry and exit: an update on a controversial relationship*.
Regional Studies **43**(8), 1061-1073, 2009,
http://dx.doi.org/10.1080/00343400801968361,

[14] Thurik, A.R.; Carree, M.A.; Van Stel, A. and Audretsch, D.B.: *Does self-employment reduce unemployment?*
Journal of Business Venturing **23**(6), 673-686, 2008,
http://dx.doi.org/10.1016/j.jbusvet.2008.01.007,

[15] Saridakis, G.; Mendoza, M.A.; Muñoz Torres, R.I. and Glover, J.: *The relationship between self-employment and unemployment in the long-run: A panel cointegration approach allowing for breaks.*
Journal of Economic Studies **43**(3), 358-379, 2016,
http://dx.doi.org/10.1108/JES-11-2013-0169,

[16] Ptak-Chmielewska, A.: *The Relation between Enterprise Population Dynamics and Economic Cycle.*
International Journal of Business, Humanities and Technology **2**(2), 17-25, 2012,

[17] Salman, A.K.; Zampatti, D. and Shukur, G.: *Macroeconomic Determinants, Innovation and the Birth of New Firms: Negative Binomial Regression Approach.*
International Journal of Economics and Finance **5**(11), 72-81, 2013,
http://dx.doi.org/10.5539/ijef.v5n11p72,

[18] Armington, C. and Acs, Z.J.: *The Determinants of Regional Variation in New Firm Formation.*
Regional Studies **36**(1), 33-45, 2002,
http://dx.doi.org/10.1080/00343400120099843,

[19] Lee, S.Y.; Florida, R. and Acs, Z.J.: *Creativity and entrepreneurship: A regional analysis of new firm formation.*
Regional Studies **38**(8), 879-891, 2004,
http://dx.doi.org/10.1080/0034340042000280910,

[20] Noorderhaven, N.; Thurik, R.; Wennekers, A.R.M. and van Stel, A.: *The role of dissatisfaction and per capita income in explaining self-employment across 15 European countries.*
Entrepreneurship-Theory and Practice **28**(5), 447-466, 2004,
http://dx.doi.org/10.1111/j.1540-6520.2004.00057.x,

[21] van Stel, A.; Wennekers, A.R.M.; Thurik, A.R. and Reynolds, P.: *Explaining nascent entrepreneurship across countries.*
EIM Business and Policy Research, Zoetermeer, SCALES-paper N2003-01, 2003,

[22] Fritsch, M. and Schroeter, A.: *Why does the effect of new business formation differ across regions?*
Small Business Economics **36**(4), 383-400, 2011,
http://dx.doi.org/10.1007/s11187-009-9256-9,

[23] Pejić Bach, M.: *Exploring information and communications technology adoption in enterprises and its impact on innovation performance of European countries.*
Ekonomický časopis **62**(4), 335-362, 2014,

[24] Luptak, M.; Boda, D. and Szucs, G.: *Profitability and Capital Structure: An Empirical Study of French and Hungarian Wine Producers in 2004-2013.*
Business Systems Research Journal **7**(1), 89-103, 2016,
http://dx.doi.org/10.1515/bsrj-2016-0007,

[25] Linan, F. and Fernandez-Serrano, J.: *National culture, entrepreneurship and economic development: different patterns across the European Union.*
Small Business Economics **42**(4), 685-701, 2014,
http://dx.doi.org/10.1007/s11187-013-9520-x,

[26] Szabo, Z.K. and Herman, E.: *Productive entrepreneurship in the EU and its barriers in transition economies: A cluster analysis.*
Acta Polytechnica Hungarica **11**(6), 73-94, 2014,
http://dx.doi.org/10.12700/APH.11.06.2014.06.5,

[27] Abdesselam, R.; Bonnet, J. and Renou-Maissant, P.: *The Demography of Enterprises and Employment in the European Union Countries*.
Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS No. 2017-10, 2017,

[28] Dumičić, K.; Žmuk, B. and Harmina, A.: *Clustering of European countries regarding recent changes in business demography statistics.*
The 15th International Symposium on Operational Research SOR'19. Slovenian Society Informatika, Section for Operational Research, Ljubljana, 2019,

[29] Eurostat: *Business demography by legal form (from 2004 onwards, NACE Rev. 2).*
http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=bd_9ac_1_form_r2&lang=en,

[30] World Bank: *GDP per capita, PPP (constant 2011 international $).*
http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD,

[31] World Bank: *Unemployment, total (% of total labor force) (modeled ILO estimate).*
http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS,

[32] OECD/Eurostat: *Eurostat-OECD Manual on Business Demography Statistics.*
OECD Publishing, Paris, 2007,
http://dx.doi.org/10.1787/9789264041882-en,

[33] Calinski, T. and Harabasz. J.: *A dendrite method for cluster analysis.*
Communications in Statistics **3**(1), 1-27, 1974,

[34] Duda, R.O.; Hart, P.E. and Stork, D.G.: *Pattern Classification.*
Wiley, New York, 2000.