Population Review

Volume 61, Number 2, 2022

Type: Article, pp. 1-25

How to Study Life Expectancy at Birth (e0) Differences between The Two Genders: A Methodological Proposition

Author: Konstantinos N. Zafeiris
Author affiliation: Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace. P. Tsaldari 1, 69132-Komotini, Greece
Corresponding author/address: Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace. P. Tsaldari 1, 69132-Komotini, Greece; e-mail: kzafiris@he.duth.gr

Abstract

This methodological paper proposes a combination of methods for evaluating the gender gap in life expectancy at birth between different populations. This approach consists of three steps. First, the abridged life tables per gender are calculated. Second, the gender differences in life expectancy at birth are decomposed by age by applying a relevant procedure. In the final step, an agglomerative cluster analysis method is applied. Several problems of this method are discussed and ways to overcome them are proposed. Results indicate the method's validity and sensitivity in portraying the segmentation or differences between countries in a cross-sectional study.

Keywords

Mortality, Life Table, Decomposition Methods, Cluster Analysis, Mahalanobis Distance

Introduction

It is common knowledge that women live longer than men. This phenomenon – a result of "male excess mortality" (see, for example, Wisser and Vaupel 2014) – can be illustrated by the differences in their life expectancy at birth. The earliest evidence for this comes from the 18^{th} and 19^{th} century (see Luy 2003; Barford et al. 2006). In the modern era, the width of the life expectancy at birth (e0) gender gap increased in the wealthier countries, at least up to the 1970s when it started to decrease (see Trovato and Heyen 2006). Gradually this trend started to spread worldwide. Overall, there is an inverse relationship between the longevity increase and the relevant gender gap (Fedotenkov and Derkachev 2020).

As Zafeiris (2020a) summarizes, the gender disparities in e0 bear several biological and social components. The first may include genetic, anatomical, and physiological elements (see, for example, Lindahl-Jacobsen et al. 2013; Brønnum-Hansen and Juel 2001). The second consists of behavioural and lifestyle agents (like smoking, alcohol consumption, violence) along with occupational hazards, medical and health developments, and other factors (see, for example, Alberts et al. 2014; Luy and Wegner-Siegmundt 2015; Hemstrom 2016; Sundberg et al. 2018). It is worth mentioning that by using panel data, in a fascinating publication, Clark and Perk (2012) found that the length of the life expectancy at birth differences between the two genders depends on gender equity and economic, political, and cultural elements.

Thus, the gender disparities in e0 is a multivariate phenomenon, encompassing all the aspects of human life. Such longitudinal studies are of great importance in identifying the modulating factors of longevity's gender gap. However, these factors may have a differential effect throughout human life, regulating human health and mortality accordingly. Thus, the mortality patterns produced under their action may differ significantly over time or between the two genders, a fact which is directly related to the study of e0 gender gap.

Of course, there is no doubt that life expectancy at birth (e0) is a straightforward and age-standardized measure of mortality (see Modig et al. 2020), i.e., an estimation of the life span. Surely then, any e0 differences between two population groups outline their overall mortality inequalities. However, e0 represents the quotient of the number of person-years lived after an exact age x (Tx) to the number of persons surviving that age (lx) (See Preston et al. 2001, pp. 38-69). It is then an average measure of longevity, affected by infant, child, juvenile and adult mortality, i.e., the mortality patterns throughout the whole life span. When comparing e0 in two groups of people, any differences result from the simple subtraction of two averages without considering the variability within these groups, while ignoring the fact that quite different mortality patterns may give similar values of life expectancy at birth (e0), not to mention that these patterns retain a dynamic character over time. It is what Arriaga (1984, 1989) notices when comparing a population's average longevity between two-time points: *"a* *change in life expectancy (at any age) does not necessarily mean that mortality rates change in the same magnitude, or even in the same direction at all ages".*

An example of this situation comes from Canada. Auger et al. (2014) compared the mean longevity in Quebec and the rest of Canada and found negligible differences (<0.1 years). However, after applying a decomposition procedure to check how the age-specific differences in mortality patterns affected longevity, they demonstrated that populations with equal life expectancy at birth might have excessive mortality inequality. Quebec's higher lung cancer mortality, affecting mainly the earlier ages, was offset by cardiovascular mortality in the rest of Canada, which afflicted the older ages. Thus, significant mortality differences can be masked even in the cases where there is no life expectancy gap.

Obviously, the same happens when studying the e0 differences between the two genders, which may be the result of a plethora of differences in their mortality patterns. An example of this fact is found in Zafeiris (2020a). After studying the evolution of the gender gap in Greece over time, he demonstrated that any developments come under the differential mortality and its differentiated evolution in each of the age groups of the human life span (for the whole problem, see also, for example, Zafeiris 2020b; Yang et al. 2012; Le et al. 2015). In conclusion, any gender differences in life expectancy at birth must be examined first in the light of age-specific and afterwards of the cause-specific contributions to these differentials.

If then, in a cross-sectional study, the research question is the comparative analysis of the gender gap among several countries, is the univariate analysis of the gender gap length capable of accurately portraying their classification/segmentation according to their similarities/dissimilarities? Even if the answer initially seems to be simple and positive, after considering the discussion in the previous paragraphs, it seems that this approach can only partly serve this purpose as countries with totally different mortality patterns and similar gender gaps may be grouped together. Therefore, this is a partial endeavor because it cannot accurately express the actual mortality differences between men and women in every one of the spatial units studied.

Thus, the research question in this paper is how to develop a procedure to classify/group countries according to the gender gap length by considering – at the same time – the gender differences in their age-specific mortality rates, i.e., their actual mortality patterns. Thus, this is rather a methodological paper than a substantive one: it illustrates how to measure and portray changes in life expectancy.

For that, a three-step procedure is proposed, consisting of three discrete methods. First, the abridged life tables per gender are constructed. In the second step, a decomposition method following Arriaga (1984, 1989), is applied. This discrete analysis aims to decompose the e0 gap between two groups of people according to their mortality pattern differences, i.e., by examining their differences in age-specific death rates and their contribution to the overall gender gap. The details of this method are described in the data and methods section of this paper.

The analysis will take place in seven age groups (dicussed later in the text), creating a multi-dimensional environment for each of the countries. Each dimension represents the effects of the age-specific mortality differences between females and males on the gender gap. The problem then is finding a way to summarise this information without any significant loss of mortality differential details and identifying any segmentation among the countries studied.

An obvious solution is to apply the method of Principal Component Analysis (see Jolliffe and Cadima, 2016) to reduce dimensionality, ideally into two parameters and construct a scatter diagram afterwards. Unfortunately, this method's application unmasked significant problems; the most important was the low proportion of the variance explained (about 60% for the year 2017). Thus, it was not used.

Instead, an agglomerative hierarchical cluster analysis (see Sokal and Rohlf, 1995) method is used in this paper, in which the countries are grouped according to the age-specific effects on the gender gap. According to Zafeiris and Tsoni (2021), this procedure starts with several sub-clusters of one point each. Subsequently, these are combined to form larger sub-clusters in an iterative process until the end of the procedure, based on a distance measure calculated from the data (see Hand 1981). In the end, the construction of a dendrogram (see later in the text) summarises the cluster solution. In that, one axis consists of the cases. The other describes the distance (dissimilarity) between the clusters. This dendrogram portrays the possible segmentation of the countries (for a discussion of several details, problems of this method, see the data and methods section).

Data from the European populations will be used as a case study to present this procedure and check its validity. This procedure can be applied to data from every country in the world. However, Eurostat's data were chosen because of their quality and consistency in estimating the yearly counts of deaths and mid-years populations, necessary for calculating life tables. For some European populations, for which no data were available in the Eurostat database, the estimates of the official Statistical Offices of each of them were used (see data and methods section). It should also be stressed that the analysis can easily be applied to lower spatial levels than the national populations, but this is a subject for another paper. Also, the causes of death have not been considered so far, as will happen in the future.

It must be stressed that the scope of this paper is not the examination of the temporal trends of genders' gap in life expectancy at birth in Europe, but the presentation of the procedure applied. Besides that, the analysis encompasses data from two subsequent years (2017 and 2018) to check the procedure's effectiveness, thus, to testify if any changes in the age-specific mortality schedule affect the generated interpretation of the cluster analysis. In that way, the sensitivity of the procedure in portraying the segmentation of countries when mortality changes will be affirmed.

This paper is organized as follows. In the Data and Methods section, the details of the procedure used will be discussed. The Results section presents the life expectancy at birth by gender and the changes observed between 2018 and 2017 among the countries to outline the longevity situation in Europe before applying the clustering procedure. For the same reason, a univariate examination of the gender gap and the relevant differences between the two years will occur. Afterwards, the clustering procedure results, and the characteristics of the major clusters formed will be presented. In the Discussion and Conclusions section, the findings of the analysis will be discussed and summarised, along with a discussion of the validity of this methodology in portraying the actual situation when comparing multiple populations.

Data and Methods

Data come from the Eurostat database and refer to population estimates by age and gender, along with the relevant distribution of deaths (https://ec.europa.eu/eurostat/data/database). The microstates of Europe were omitted due to the lack of data (Andorra, Liechtenstein, San Marino, Monaco). In the analysis, the area of Kosovo is also included. Data for Russia come from the Federal State Statistical Service (https://eng.gks.ru/, https://showdata.gks.ru/finder/descriptors/278982). For Moldova from the National Bureau of Statistics of the Republic of Moldova (https://statistica.gov.md/index.php?l=en). For Bosnia and Herzegovina from the Statistical Yearbook of 2019 (Institute for Statistics of FBiH, 2019). Eurostat does not publish detailed data for these countries. The abridged life tables of the populations were constructed after Chiang (1984).

As Zafeiris (2020a) notes, numerous papers have been published for the decomposition of the differences in life expectancy at birth (e0) (see, for example, Andreev 1982; Pollard 1982, 1988; Pressat 1985; Andreev and Shkolnikov 2012; see also Andreev et al. 2002). However, in this paper, the decomposition method of Arriaga (1984; see also 1989) will be applied because it is straightforward and has broad applicability and consistency in its results (see, for example, Yang et al. 2012; Auger et al. 2014; Le et al. 2015; Sunberg et al. 2018; Chisumpa et al. 2018). According to this method, the effects of mortality change on life expectancies can be classified as "direct" and "indirect". The *"direct effect on life expectancy is due to the change in life years within a particular age group as a consequence of the mortality change in that age group"*. The indirect effect *"consists of the number of life years added to a given life expectancy because the mortality change within (and only within) a specific age group will produce a change in the number of survivors at the end of the age interval"*. Another effect springs from the interaction between the exclusive effect of each age group and the overall effect (for the relevant formulas for the calculations, see Arriaga 1984; see also Nusslelder and Looman 2004).

Therefore, any difference in life expectancy at birth is *"the sum of the number of years added to (positive contribution) or removed (negative contribution) from life expectancy at birth attributable to the decrease or increase, respectively, of mortality at each age for a given period",* as Yang et al. (2010) noted when they studied the rapid increase of e0 in South Korea.

In this paper, e0 differences correspond to the genders' gap length in a particular year. The analysis will be done in 15-year-long age groups except the two first (<1 and 1-14 years) and the last one (75+ years), in total seven age groups (called variables in the following paragraphs). The algebraic sum of these age groups' effects corresponds to the width of the gender gap, which will not be used as a separate variable in the analysis. The use of smaller length age groups (for example, five-year age groups) limited the procedure's analytic capability because of the minimal mortality differences between the two genders in these age groups in many European countries.

After estimating the effects of the mortality differences in the seven age groups, the cluster analysis procedure will be applied using the variables mentioned above (a detailed illustration of this procedure, its statistical requirements and problems can be seen in Zafeiris and Tsoni 2021).

Table 1. Multicolinearity statistics

Statistic	<1	1-14	15-29	30-44	45-59	60-74	75+
				2017			
R^2	0,054	0,300	0,864	0,920	0,947	0,873	0,467
Tolerance	0,946	0,700	0,136	0,080	0,053	0,127	0,533
VIF	1,057	1,429	7,352	12,538	18,854	7,857	1,878
				2018			
R^2	0,197	0,398	0,791	0,918	0,967	0,923	0,434
Tolerance	0,803	0,602	0,209	0,082	0,033	0,077	0,566
VIF	1,246	1,662	4,776	12,146	30,501	13,010	1,767

Source: own calculations

One of the solid prerequisites for carrying out a successful cluster analysis is avoiding multicollinearity among the variables or finding a way to overcome it (for its effects, see Ketchen and Shook, 1996). The term multicollinearity describes a situation of high intercorrelation or inter-association among the independent variables. The strategy to identify multicollinearity is to perform linear regressions of each of the variables as a function of the others. Afterwards, the coefficient of determination R^2 for each of the regressions is calculated. This coefficient defines the proportion of the variance in the dependent variable that is predictable from the independent variables. A value of 1 indicates a perfect fit of a regression model. The Tolerance is $1-R^2$ and the Variance Inflation Factor (VIF), is the inverse of the Tolerance. Commonly a VIF above 5 (or 10 in respect of the sample size) indicates multicollinearity (see, for example, Daoud, 2017; Allison 1999). Therefore, a multicollinearity test was applied in this analysis.

According to this test, the VIF is high for many variables (age groups; Table 1), Thus, the existing multicollinearity between the variables constitutes a significant problem, relating to the estimation of the proper distance measure of cluster analysis. Measures like the Euclidean or the squared Euclidean cannot be used, as they are affected by this problem. Instead, one can apply the Mahalanobis distance because it accounts for correlation among the variables (see Hair et al., 1998, p. 444; Masnan et al. 2015), as actually did happen.

Afterwards, several cluster procedures were applied (complete linkage, flexible linkage, single linkage, unweighted pair group average and weighted pair group average). According to the cophenetic correlation coefficient, the method of unweighted pair group (see Sokal and Michener, 1958) gave the best results (see Sokal and Rohlf, 1962). This coefficient measures the validity of the analysis and is 0.8 for the year 2017 and 0.9 for 2018; thus, the cluster solutions were of high quality (see Clarke et al. 2016). The cluster analysis was carried out with the XLSTAT software (Addinsoft 2020).

Results

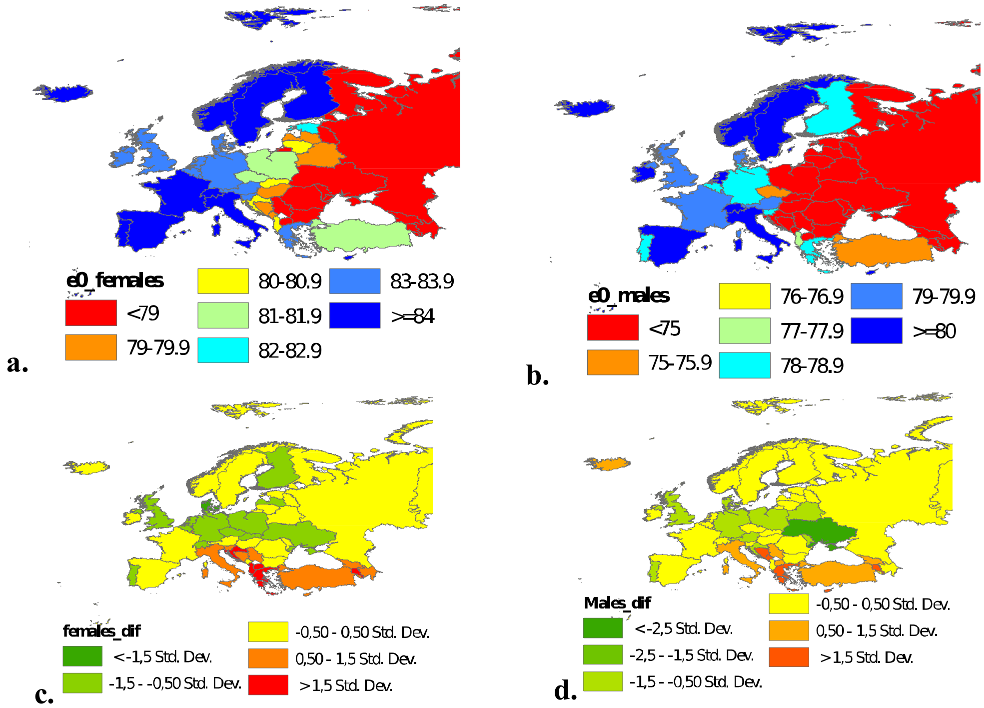
A. **Presenting life expectancy ta birth (e0) in Europe in the years 2017 and 2018**

1. **Life expectancy at birth**

In the year 2017, the mean duration of life is remarkably diverse in Europe, judging by the life expectancy at birth (e0) (Figure 1: a, b). A minimum e0 of 67.6 years for males in the Russian Federation (RU) and 77.1 years for females in Moldova (MD) indicates the lower bounds of longevity in the continent. On the other edge of the longevity distribution, Switzerland (CH) has the most extended life span in males (81.4 years) and Spain (ES) in females (85.7 years). Thus, the range (the difference between the maximum and minimum values) is 13.8 years in males and 8.6 in females. Overall, the European continent's average e0 is 75.9 years in males and 81.6 years in females (standard deviation of 4.1 and 2.6 years, respectively). In males living in the European Union (EU) and Eurozone aggregates, e0 is 78.1 and 79.2 years. In females 83.3 and 84.2 years; thus, e0 is more elongated in the economically more developed part of the European Union, i.e. in the Eurozone. Note that the UK is included in the EU estimations even though it exited the Union on January 1st, 2021.

One could suggest that the existing heterogeneity develops itself in the emerging of a clear spatial pattern, dividing the European continent into two major groups: the western and eastern ones. Greece (located geographically in the south-eastern part of the continent) and Cyprus are components of the first group. In this, the average life span is more extended than in the eastern group. But the variability within the two groups remains as high as it is between them. This heterogeneity indeed reflects the recent demographic history of the distinct spatial units under consideration, connected to the diversification of the modern and relatively recent historical developments in terms of political, social, economic, cultural and health grounds, though such discussion does not fall within the scope of this article. Besides that, it is obvious that the univariate approach for identifying the segmentation of the countries bears a subjective element. It is up to the researcher to decide the best interpretation of the findings after focusing on different aspects of this variability.

Figure 1. Life expectancy at birth (e0) by gender in the year 2017 (a, b) and changes of e0 between the years 2018 and 2017 (c, d).



The next year, 2018, life expectancy at birth on average changes about +0.2 years in both genders (standard deviation: 0.26 in males and 0.2 in females). It decreases in 9 countries in males and 6 in females. The opposite happens in 36 and 39 countries, respectively; the diversity thus remains high. Overall, the mean longevity tends to increase mainly in large parts of the Mediterranean basin. Italy (IT), Greece (GR), Turkey (TR) and other countries are among them.

In the rest of Europe, a mixed picture prevails in both genders, and no spatial pattern is visible (see maps c and d of Figure 1, where the European countries cluster according to the per gender differences of e0 between the years 2018 and 2017 and the relevant standard deviation of these differences). On the continent level, the average life expectancy at birth is 76.1 (std: 4.1) in males and 81.8 (std: 2.6) years in females. Life expectancy at birth is higher in the European Union (males: 78.2; females: 83.4) and Eurozone (males: 79.3; females: 84.4).

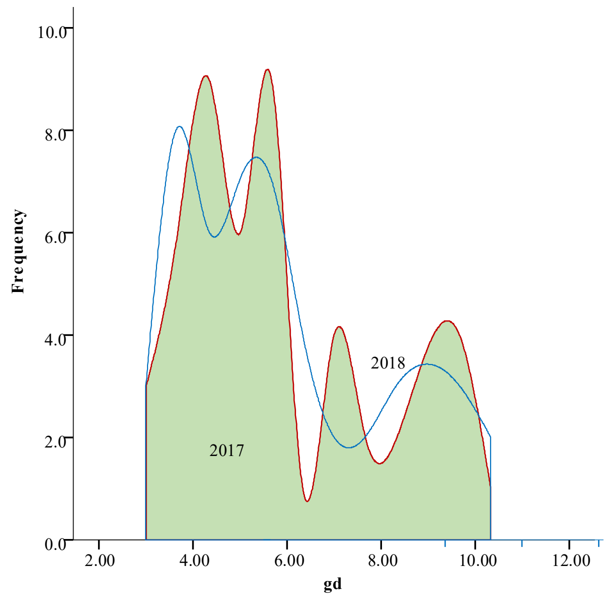
Even if these changes are not tremendous, they will almost surely affect the longevity gender gap. The question then is to what degree the segmentation/grouping of the countries will change, considering the univariate approach used so far.

2. The gender gap

The frequency distributions of the life expectancy at birth differences between the two genders (Figure 2), have a multi-modal shape and reveal that Europe is a highly diversified area. Females live longer than males (Figure 2; mean=5.7 years; std= 2.0

for 2017 and 2.1 for 2018). Somewhat smaller is the gender gap in the European Union and Eurozone (EUZone; 5.2 and 5 years, respectively). The gender gap range was 7.2 years in 2017 and 7.1 in 2018. Judging by these central tendency and dispersion measures, initially things do not seem to change a lot between the two years. However, this is not correct as revealed by the differences in their frequency curves. This diversification will be addressed in the forthcoming cluster analysis.

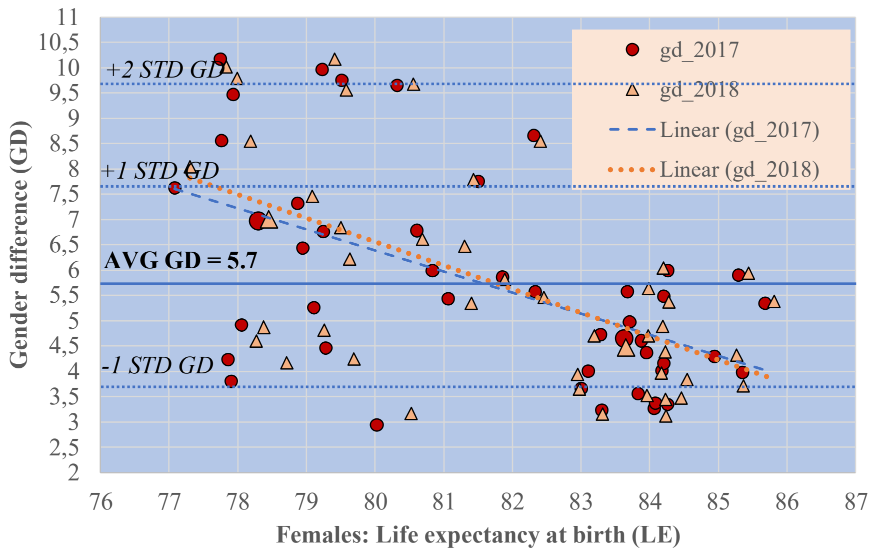
Figure 2. Frequency distribution of longevity differences between the two genders. Years: 2017 and 2018.



Moreover, the five most differentiated countries on the right side of the frequency curves are the Russian Federation (RU), Belarus (BY), Latvia (LV), Lithuania (LT) and Ukraine (UA), i.e., in some of the former Soviet Union countries, where the gender gap exceeds 9 years. In contrast, on the left side of these distributions, Albania (AL), The Netherlands (NL), Iceland (IS), Norway (NO) and Sweden (SW) have the five lowest values, below 3.5 years. Albania is probably an exception in this group, ranking in 68th position of the human development index (see United Nations, http://hdr.undp.org/en/content/human-development-index-hdi, UNDP 2018).

The Netherlands ranks in 10th position, Sweden in 7th, Iceland in 6th, Norway in 1st. According to this classification, Albania clusters in the "high human development group" and the others in the "very high development group". Therefore, countries with similar geographic, political, socio-economic, and other characteristics may have significant affinities in the longevity differences between women and men, but also other and very diverse countries may do so.

Figure 3. Female life expectancy at birth (x-axis) and gender differences. Years 2018 and 2017.



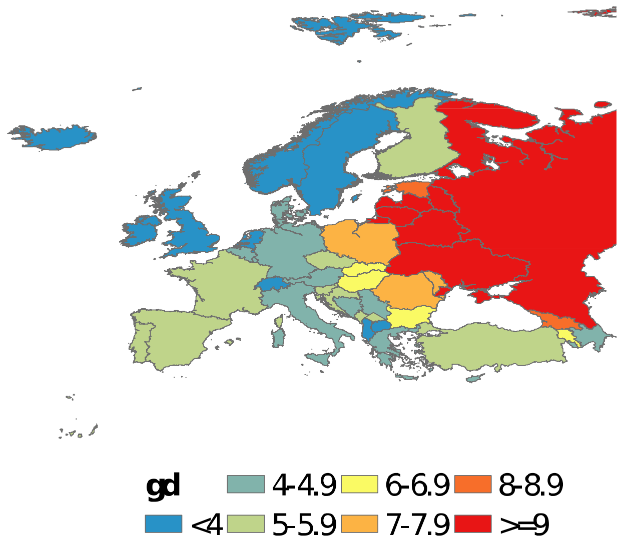
Another question emerges. Does the gender gap length have any relation to the life expectancy at birth? Or to put it in another way, is the heterogeneity in life expectancy at birth (e0) connected with the heterogeneity in gender gap? In the scatterplot of female life expectancy at birth against the relevant gender gap (Figure 3), it is apparent that the diversity in the mean duration of life discussed in the previous paragraphs only partially affects e0 gender differences. This relationship is hardly linear, as the coefficient of determination of the linear regression line between the two variables, R^2 is 0.29 for the year 2017 and 0.33 for 2018.

In other words, the increase in life expectancy is only weakly related to the decrease in the gender gap. However, it is evident that within the countries the increase in life expectancy is accompanied by a decrease in the gender gap (Fedotenkov and Derkachev 2020). But after using the European data, such findings are not confirmed due to the 'complex interaction of biological, environmental and social factors' (Zarulli et al. 2018), which regulate the length of e0 differences between women and men in each of the countries.

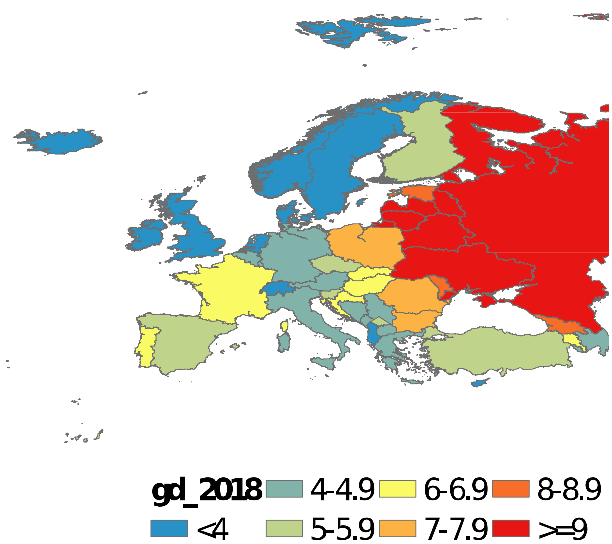
An additional problem has been identified. Countries with similar female life expectancies have quite diverse gender disparities in mortality. Others, with totally different mean duration of life (e0), have a similar gender gap. Thus, the problem of the segmentation/grouping of the countries becomes more complicated.

Figure 4. The gender gap in Life expectancy at birth. Years 2017 and 2018.

2017



2018



Meanwhile, if life expectancy at birth is omitted as a taxonomic criterion, the European countries – in a univariate analysis – can be classified in five categories if the classifying criterion is the standard deviation of the gender gap distribution (Figure 3). If the classifying criterion changes, the clustering of the countries alters too. Figure 4 shows a seven-group classification, emphasising at the same time on the subjectivity of this depiction, as is the case with the one mentioned above.

The gender gap is extensive in most former USSR democracies and other countries of eastern Europe, and small in Scandinavia, the British Isles and Iceland, and some scattered countries all over the continent. Among them, an intermediate but highly heterogeneous zone spread over much of Europe. In this pattern, several countries (for example, Iceland, Albania, and Cyprus) changed cluster between 2017 and 2018.

Between 2018 and 2017, the length of the gender gap decreases in 18 countries and increases in 27. On average, the changes are minimal, only -0.001 years. Nevertheless, the large standard deviation of 0.19 years indicates the high diversity of gender gap changes among the European countries. The minimum value is -0.44 in Montenegro and the maximum +0.48 in Croatia, representing a range of 0.92 years for the whole continent. Thus, the gender gap in life expectancy at birth is far from considered stable for many European countries in the two years studied.

Figure 5. Changes in the gender gap in life expectancy at birth between the years 2018 and 2017

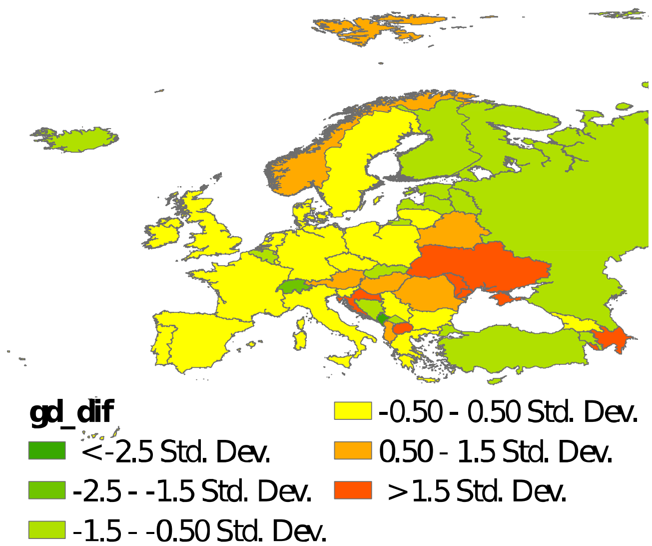


Figure 5 depicts this situation. In most of the countries, the gender gap changes lie in the range of ±0.5 standard deviations. In many others, mainly in the eastern part of the continent, the changes are more prominent, and the most extreme values are found – besides the already mentioned Croatia (HR) – in North Macedonia (MK), Moldova (MD), Ukraine (UA) and Azerbaijan (AZ). In these countries, the gender gap is longer in the year 2018 than in 2017. Somewhat smaller are the changes in Belarus (BY), Romania (RO), Austria (AT) and others. In contrast, in the countries marked with various shades of green in Figure 4, the gender gap decreases in 2018, though at different levels. Among them some countries like Serbia (RS), Turkey (TR), many ex-USSR democracies (i.e., most Baltic countries and the Russian Federation) and others rest.

In conclusion, the univariate analysis of the gender gap in Europe revealed a spatial segmentation, though the variability is high. The differences in the length of the gender gap between the two years studied, which in some countries are significant, affirm the dynamic nature of the phenomenon of male excess mortality, highlighting that on its part it comes under the influence of a variety of local agents, socioeconomic, cultural and health circumstances (see Beltrán-Sánchez 2015; Maiolo & Reid 2020). For this reason, minor changes occurred in the clustering of the countries between the two years studied.

B. The cluster analysis

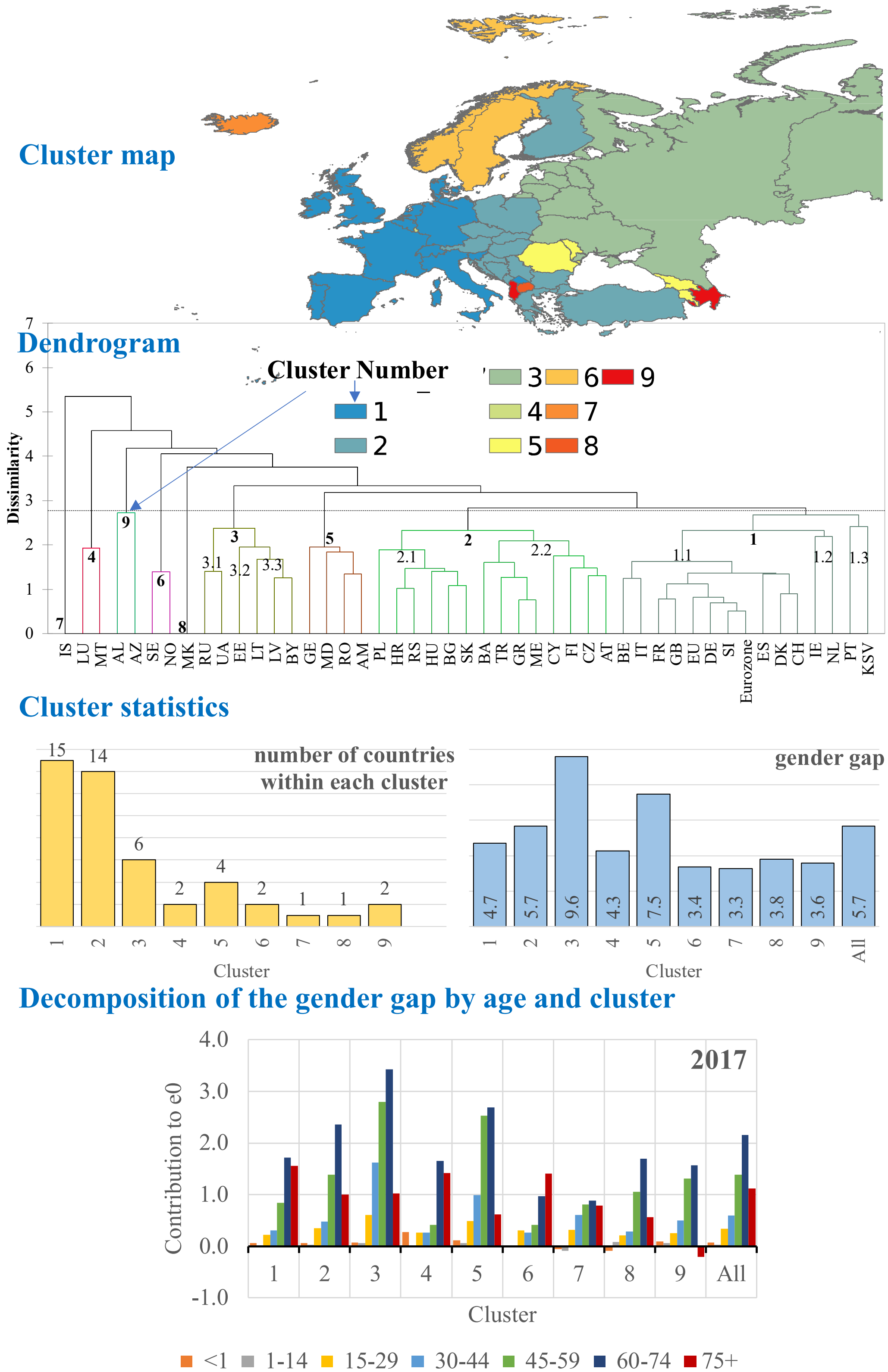
Any classification so far depends on the taxonomic criterion used; thus, it lacks objectivity, and it does not consider the problem of the comparability of the gender gap among the countries studied. Instead, the procedure presented in this section besides being objective, also takes into consideration the different mortality patterns occurring among the countries and between the two genders.

The results of the application of this procedure for the year 2017 are seen in Figure 6. The dendrogram represents the dissimilarities (y axis) among the countries (x axis). The closely related countries are linked tightly. Longer vertical and parallel lines connect countries will smaller affinities. The different clusters are presented with different colours, while they are numbered accordingly. These clusters have been mapped in the cluster map. However, a difference exists between the two representations. The map contains only the major clusters, while the dendrogram contains also the subclusters formed within each cluster, i.e., the groups of countries with greater affinities within each cluster. The cluster statistics represent the number of countries and the average length of the gender gap within each cluster (including the aggregates of the European Union and the Eurozone countries). The decomposition of the gender gap by age and cluster represents the results of the application of the Arriaga method (averages within each cluster), and in that way the difference in the mortality patterns the two genders have between the clusters.

In the year 2017, a tripartite division of Europe prevails, noting at the same time the existence of some smaller clusters or some exceptionally diversified countries. The first cluster (cluster 1) consists of the countries located west of the Adriatic Sea, Slovenia, Kosovo (KSV), the European Union (EU) and Eurozone aggregates, in total 15 countries/areas. The second cluster (cluster 2) consists of most of the countries of south-eastern Europe (14 countries), those included in the line up to Austria (AT), Czechia (CZ), Poland (PL), and Finland (FI). Romania (RO), Moldova (MD), Georgia (GE) and Armenia (AM) form the distinct cluster 5, located amid Cluster 2 and the one formed by most ex-USSR countries, including the Baltic ones (cluster 3). The other countries are more diverse, forming, however, some notable groups like that of the Scandinavian countries of Sweden and Norway.

As a result of the existing diversity of mortality patterns, these major clusters enclose smaller units (sub-clusters), which sometimes include neighbouring countries such as France (FR) and the United Kingdom (UK; sub-cluster 1.1) and Czechia (CZ) and Austria (AT; sub-cluster 2.2) or more distant ones like Ireland (IE) and The Netherlands (NL; sub-cluster 1.2). The European Union (EU) and Eurozone have greater affinities with Germany (DE), France (FR), Slovenia (SI) and the United Kingdom (UK). In these countries lives a large portion of the European population.

Figure 6. Clusters' details. European countries. Year 2017.



The first three clusters have significant differences in the average gender gap, which increases from 4.7 years in the first to 5.7 in the second and 9.6 in the third. Except for cluster 5 (7.5 years), the gender gap is shorter in the others. Also, cluster 4 is very close to cluster 1, judging by the total length of the gender gap (4.3 and 4.7 years, respectively). Indeed, in a univariate clustering process, both would have been merged into one, as happened in the mapping process of Figure 3. The same refers to clusters 6, 7, 8 and 9, all of them consisting of one country (keep in mind that these figures represent cluster averages or the unique values of one member clusters, and any mapping of the individual values per country might be slightly different). Therefore, it proves reasonable that the univariate approach may obscure the true nature of e0 differences between the two genders.

The nature of the classification discussed in the previous paragraphs is clarified after studying the decomposition of the gender differences by age. Overall, the most important fragment of the human life span is between 60 and 74 years, where the highest effect on the longevity differences between the two genders lies in most of the clusters. Except for some clusters, the second most crucial effect comes from the male excess mortality in the ages 45-59 years and secondarily in the ages 75+ years. A smaller effect springs from the age group 30-44, while the others contribute to a lesser degree.

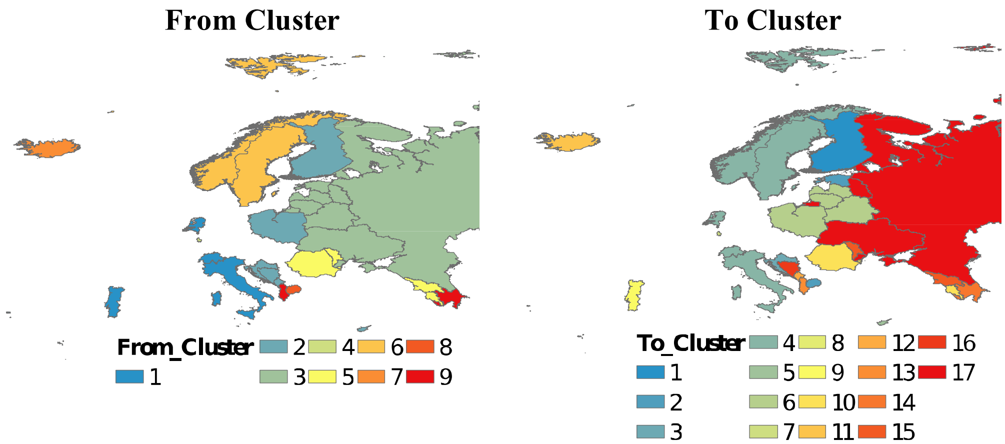
However, the effects of gender mortality disparities in all the age-groups of the human life span vary a lot between the clusters. Because of that, the differences among them are very clear. In cluster 1 (representing as already mentioned the European countries west of the Adriatic Sea along with Slovenia and Kosovo and the aggregates of the European Union and Eurozone) the gender gap is regulated by the male excess mortality in the older ages (60+) and to a lesser degree than that in the age group 45-59 years.

In contrast, in cluster 2 (south-eastern Europe countries and others) most of the e0 differences emanate from the male excess mortality in the ages 45-74 years and to a lesser degree in the 75+ years. In cluster number 3, corresponding to some ex-USSR countries, besides the other differences like the analogically small contribution of the older (75+ ages) to the overall gender gap, a strong effect of males' excess mortality in the ages 30-44 years is observed.

The same happens in cluster 5, but there the gender differences in the ages 45-59 and 60-74 have an almost equal effect on e0. Similar observations can be made for the other but smaller clusters, but it is worth noting that in cluster 9 (Albania and Azerbaijan) mortality in the older ages (75+ years) is higher in females than in males; a fact which needs further research and clarification.

So far, the cluster analysis revealed a clear and objective presentation of the countries' segmentation. However, as presented in the previous section of this paper, mortality patterns are unstable between the years 2017 and 2018. The question then is if the procedure presented in this paper bears the distinctive ability to record these changes. In other words, the question is about the sensitivity of the method in the examination of the gender differences when mortality changes even if in a small magnitude.

Figure 7. Changes in the clustering of countries between 2017 (From cluster) and 2018 (To cluster).



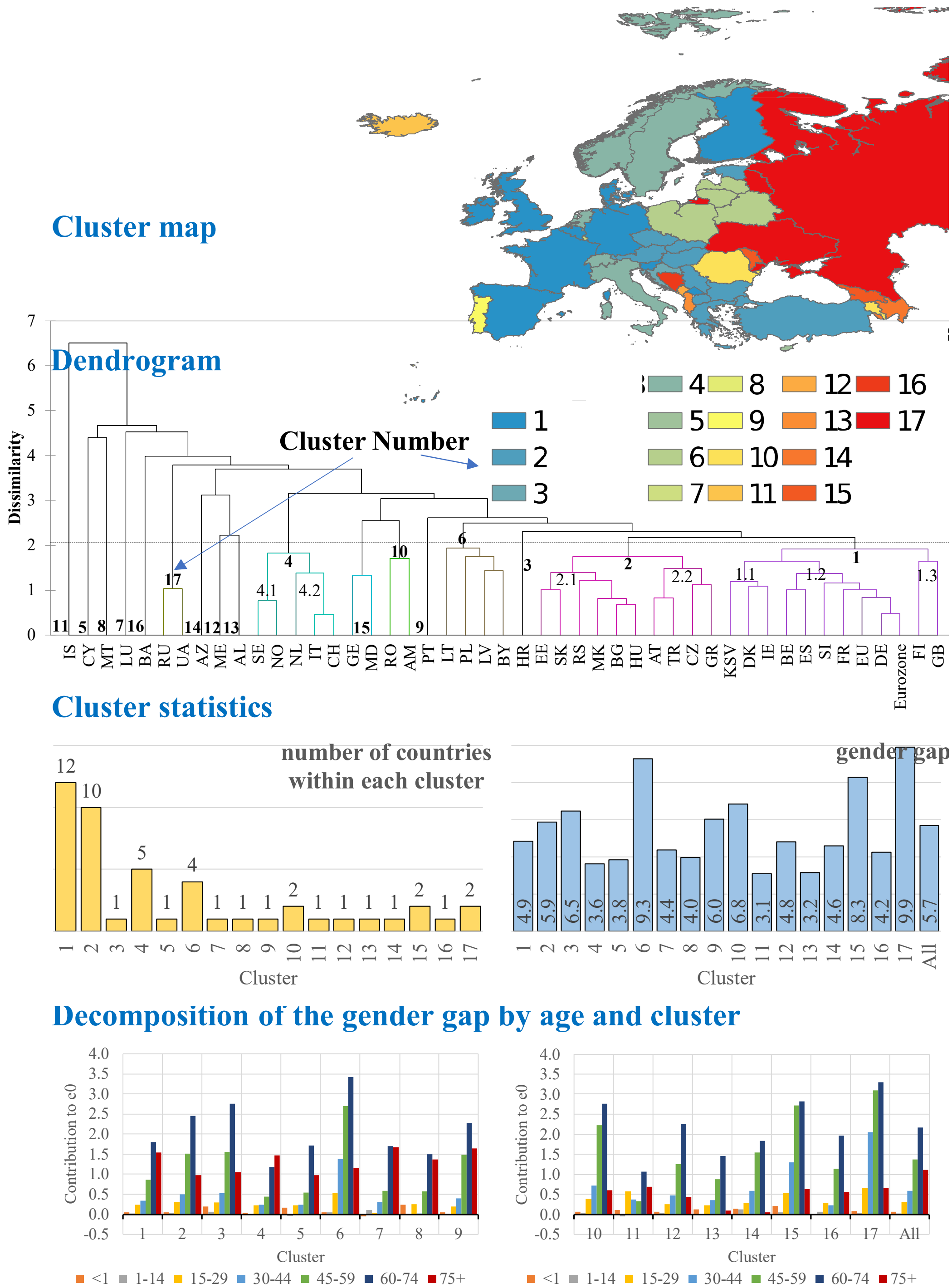
Thus, a year later, in 2018, several countries from eastern Europe, the ones around the Adriatic Sea basin and Scandinavia change cluster in the produced dendrograms (Figure 7).

Consequently, a total of 17 clusters were produced for 2018 (Figure 8) indicating a higher diversity for that year (compared with 9 for 2017; Figure 6). The differentiation of Eastern Europe from Western still holds. The Russian Federation (RU) and Georgia (GE) are the most distantly located countries. Azerbaijan (AZ), Montenegro (ME), Albania (AL), and Bosnia Herzegovina (BA) are 4 other highly diversified south-eastern and Eastern Europe countries positively contributing to the observed heterogeneity.

The dissimilarity decreases towards the other countries. Georgia (GE) and Moldova (MD) form Cluster 15 with high affinities with Romania (RO) and Armenia (AM; Cluster 10; Figure 8); both closer to the rest of the countries not discussed so far. One of the most striking groupings concerns the countries on the south-eastern shore of the Baltic Sea. Belarus (BY), Latvia (LV), Lithuania (LT), and Poland (PL) form a distinct cluster (Cluster 6), with more remarkable affinities with clusters 1 and 2. The latter bear many similarities with the homonymous clusters for 2017, described previously. Croatia (HR; Cluster 3) forms a unique cluster between clusters 6 and the group of clusters 1 and 2.

Italy (IT) clusters remarkably close to nearby Switzerland (CH), forming the first branch of cluster 4. The Netherlands (NL) and the Nordic countries of Sweden (SE) and Norway (NO) form the second branch of this cluster.The most diversified countries are Luxembourg (LU), Malta (MT), Cyprus (CY), and Iceland (IS). The characteristics of these clusters are discussed in the following paragraphs.

Figure 8. Clusters' details. European countries. Year 2018.



Judging by the number of countries participating in a cluster, the most populous ones are cluster 1 (12 countries), where the length of the gender gap is 4.9 years and cluster 2 (10 countries), where the gender gap is 5.9 years. This gap is smaller in cluster 4 (5 countries 3.6 years) and becomes larger in cluster 6 (9.3 years). The variability between the clusters is high. Cluster 11 (Iceland) has the minimum value (3.1 years) and cluster 17 (Russian Federation and Ukraine) the maximum (9.9 years).

In this classification – as happened also for the one of year 2017 but to a lesser degree- it is evident that countries with similar gender gaps belong to different clusters. Clusters 4 and 5, for example have identical length in their gender gap but many differences in the age specific mortality disparities, as seen in the decomposition of the gender gap by age and cluster. The same happens with clusters 11 and 13, and 7, 16 and 8. Similar observations can be made for the other clusters, indicating the sensibility of the procedure used.

The variability, then, between the clusters formed during the analysis elucidates the effects of the different age and gender-specific mortality disparities in regulating the gender gap along with the length of the gender gap. Thus, the gender gap is evaluated and ascribed to the mortality patterns' differences during the whole life span. This information must be considered in the ranking process or the proper clustering of the different human populations, addressing the gender gap estimations' comparability problem and thus giving reliable and consistent results.

However, while the sensitivity of the procedure used here is good enough to produce an accurate picture of the countries' segmentation, it is still subject to year-by-year mortality fluctuations. As a result, small mortality changes between the years 2017 and 2018 changed the clustering of the countries. If needed, the production of more robust results requires a way to avoid these fluctuations. A possible solution is to use two- or three-years data to smooth the death probabilities and their temporal fluctuations.

Discussion and Conclusions

If someone wants to discuss the longevity situation in Europe presented in the Results section of this paper, besides not being the scope of this paper, he/she will acknowledge that several determining factors are responsible for the development of the differential mortality patterns among its populations. Besides the well-known effects of economic development, the social structure, education, public medical developments, access to the health care system, differences in health-related behaviours, health interventions, vaccination, urbanization, nutrition, and others are also playing a crucial role (see, for example, Cutler et al. 2006; Jeong et al. 2020; Martín Cervantes et al. 2020).

The differential effect of these factors produced an extreme heterogeneity in the mean duration of life in the European countries, within a mortality transition regime (see, for example, Vallin and Meslé 2004). This heterogeneity and the segmentation that still exists and distinguishes many of the former Eastern Europe countries are also based on historical and political developments. As Meslé and Vallin (2011) state, by the time of the cardiovascular revolution, some of them started to converge with the rest of the European countries during the 1990s, while others (like Russia, for example) seem to retain a more diversified position today. Additionally, the economic crisis, which started in 2008 and continued to affect European populations for several years later, played an essential role in maintaining mortality levels (see, for example, Alicandroet al. 2019; Ballester et al. 2019). These developments occurred in a changing disease and morbidity environment (see GBD 2019 Diseases and Injuries Collaborators 2020).

Mortality is not stable between 2017 and 2018 in many European countries; an expected finding due to a possible differential action of the factors roughly described in the previous paragraph. Additionally, random (stochastic) e0 changes indeed occurred; a problem which is becoming increasingly important in the smaller populations of the continent.

These changes have a gender-specific element. In this multivariate scheme, the width of life expectancy at birth gender gap alters even if the changes are small in many populations. A brief discussion of the regulating factors of this phenomenon took place in the introductory section.

Apart from these, which are not the subject of this paper, the question of the classification of different populations was originally addressed as a univariate problem. In this display, the categorization of these populations comes solely from the length of the gender gap in life expectancy at birth. The mapping of such data may reveal the segmentation patterns occurring among the countries.

Thus, in all the European countries, women live longer than men, as expected. Note that e0's variability discussed in the previous paragraphs affects the gender gap only partially; an increase in life expectancy at birth is only weakly related to the decrease in the gender gap. Additionally, the countries' classification may alter when the taxonomic criterion changes; a long-known problem for statistics referring to subjectivity. By defining the groups' characteristics used, for example, in the mapping process, the researcher affects the countries' segmentation solely and thus their clustering without applying an objective method (see Blyth 1972 for the definition of the terms).

If the taxonomic criterion is the standard deviation of the gender gap distribution, the European countries form five groups, for both years studied. In a seven-group classification, the former USSR republics and other eastern Europe countries form a very diverse group. On the other edge of the observed variability lie the Scandinavia countries, the British Islands and Iceland, and some scattered countries. An intermediate but highly heterogeneous group includes the rest of the European countries. E0's gender gap alters in many European countries between 2017 and 2018 and several countries (for example Iceland, Albania, and Cyprus) change their position in this univariate classification. Thus, any depiction should consider the understudied phenomenon's dynamic nature for developing objective portraiture of the countries' segmentation.

However, the most critical issue is that the analysis so far does not consider the real gender gap's nature: any disparities in longevity spring from the age-specific differences in men and women's mortality patterns. This information is essential to record and evaluate the real dimensions of the phenomenon being studied. An example comes from the countries with an almost equal gender gap, but different mortality patterns (see results section). Then, the problem is to apply a method that considers all the information and is capable of a more accurate grouping of countries based on their similarities, this being the scope of this paper.

Such a method is the agglomerative cluster analysis. This method allows an objective analysis of the phenomenon because the solution is not left exclusively to the researcher, for example, when mapping the countries in a univariate analysis of the gender gap. However, the role of the researcher remains crucial. First, all the method's requirements discussed in the Data and Methods section must be fulfilled. Especially the multicollinearity problem to which this method is significantly sensitive must be addressed. Thus, the choice of the correct distance measure is crucial for the compelling depiction of the data. Therefore, the Mahalanobis distance was used in this paper as it takes this problem into account.

A second complication is that the cluster solution depends on the clustering procedure applied. Every time the cluster analysis method is applied, numerous solutions are given. Though, on the one hand, the principle of parsimony must not be forgotten – i.e. the principle that the most acceptable explanation of a phenomenon is the simplest one (see Oxford Reference, available at https://www.oxfodreference.com/view/10.1093/oi/authority.20110803100346221). On the other hand, a procedure's choice must be based on an objective criterion as much as possible. Thus the cophenetic correlation coefficient was used in this paper, allowing the choice of the most effective cluster solutions.

The results of the analysis indicate the validity of the procedure used. The clusters formed are discussed in detail in the Results section of this paper. For the year 2017, a tripartite division of Europe prevails. Surprisingly enough, this division partly recalls the old Hajnal's line (1965) that characterized Europe by different nuptiality levels. Southeastern Europe is differentiated along with some other countries at the boundaries of this line. However, in this part of Europe, the diversity remains high. Also, the Scandinavian countries retain a more diversified position.

In 2018, several countries from eastern Europe, the ones around the Adriatic Sea basin and Scandinavia change cluster in the produced dendrograms. Diversity becomes even higher, and a total of 17 clusters are formed. However, the differentiation between Eastern and Western Europe still exists.

Therefore, besides the small changes in the total length of the gender gap between the two years studied, a differentiated picture emerges if this gap's components are taken into consideration. This picture springs from the gender differences in the age-specific mortality rates throughout the entire life span, which are not stable over time. Any analysis of the gender differences in any spatial level, in order to be more accurate, must focus on the decomposition of the gender differences of mortality by age before any analytic procedure is applied. That is why Arriaga's decomposition method was used in this paper, though many others can be applied too.

In conclusion, the procedure used here is very effective in producing and portraying the segmentation of the countries in a period. The application of this procedure in two subsequent years indicated its sensitivity in recording the mortality developments, even if these are small. However, this sensitivity may not give a robust picture of the clustering of the countries, as mortality changes. The number of the clusters produced and the differences in the countries' segmentation for the two years studied indicate this problem. One solution could be the application of this procedure for longer time periods (maybe two or three years) for "smoothing" the mortality curves and as a result the production of more robust results.

Acknowledgements: This paper is part of the Research Project *Demographic Imperatives in Research and Practices in Greece*, supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "1st Call for H.F.R.I. Research Projects to support Faculty Members & Researchers and the Procurement of high-cost research equipment grant" (project number 2988).

ORCID: 0000-0002-2587-6565

Ethical approval: This article does not contain any studies with human participants performed by the author.

Funding: None

Conflicts of interest/Competing interests: no conflicts of interest.

Availability of data and material: Eurostat's and countries' Databases and publications.

References

Addinsoft (2020). XLSTAT statistical and data analysis solution. New York, USA. https://www.xlstat.com.

Alberts, S. C., Archie, E. A., Gesquiere, L. R., Altmann, J., Vaupel, J. W., & Christensen, K. (2014). The male–female health-survival paradox: A comparative perspective on sex differences in aging and mortality. In Committee on Population, Division of Behavioral and Social Sciences and Education, National Research Council, M. Weinstein, & M. A. Lane (Eds.), *Sociality hierarchy health comparative biodemography a collection of papers* (pp. 339–364). Washington (DC): National Academies Press.

Alicandro, G., Malvezzi, M., Gallus, S. et al. (2019). Worldwide trends in suicide mortality from 1990 to 2015 with a focus on the global recession time frame. *International Journal of Public Health* 64, 785–795. https://doi.org/10.1007/s00038-019-01219-y.

Allison, P. (1999). Multiple Regression: A Primer. (p. 142). Thousand Oaks, CA: Pine Forge Press.

Andreev, E. M. & Shkolnikov, V. M. (2012). An Excel spreadsheet for the decomposition of a difference between two values of an aggregate demographic measure by stepwise replacement running from young to old ages. *MPIDR* TECHNICAL REPORT 2012-002. APRIL 2012

Andreev, E. M., Shkolnikov, V. M. & Begun, A. Z. (2002). Algorithm for decomposition of differences between aggregate demographic measures and its application healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research* 7, 499–522.

Andreev, E. M. (1982). Metodkomponent v analize prodoljitelnosty zhizni. [The method of components in the analysis of length of life]. *Vestnik Statistiki* 9, 42-47.

Arriaga, E. E. (1984). Measuring and explaining the change in life expectancies. *Demography* 21, 83-96.

Arriaga, E. E. (1989). Changing trends in mortality declines during the last decades. In: Ruzicka, L., Wunsch, G. & Kane, P. (eds). Differential mortality: methodological issues and biosocial factors (pp. 105-129). Oxford, England: Clarendon, Press.

Auger, N., Feuillet, P., Martel, S., Lo, E., Barry, A. D. & Harper S. (2014). Mortality inequality in populations with equal life expectancy: Arriaga's decomposition method in SAS, Stata, and Excel. *Annals of Epidemiology* 24, 575-580.

Ballester, J., Robine, J-M., Herrmann, F. R. & Rodó, X. (2019).. Effect of the Great Recession on regional mortality trends in Europe. *Nature Communications* **10,** 679. https://doi.org/10.1038/s41467-019-08539-w.

Barford, A., Dorling, D., & Smith, G. D. (2006). Life expectancy: women now on top everywhere during 2006, even in the poorest countries, women can expect to outlive men. *BMJ*, 332(7545), 808.

Brønnum-Hansen, H., & Juel, K. (2001). Abstention from smoking extends life and compresses morbidity: A population based study of health expectancy among smokers and never smokers in Denmark. *Tobacco Control*, 10(3), 273–278.

Beltrán-Sánchez, H., Finch, C. E., & Crimmins, E. M. (2015). Twentieth century surge of excess adult male mortality. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 8993–8998.

https://doi.org/10.1073/pnas.1421942112.

Blyth, C. (1972). Subjective vs. Objective Methods in Statistics. *The American Statistician 26*(3), 20-22. doi:10.2307/2682860.

Chiang, C. L. (1984). The life table and its applications. Malabar: Robert E. Krieger Publishing Company.

Chisumpa, V. H. & Odimegwua, C, O. (2018). Decomposition of age- and cause-specific adult mortality contributions to the gender gap in life expectancy from census and survey data in Zambia. *SSM - Population Health* 5, 218-226.

Clark, R., & Peck, B. M. (2012). Examining the gender gap in life expectancy: A Cross-national analysis, 1980–2005. *Social Science Quarterly 93*(3), 820–837.

Clarke, K. R., Somerfield, P. J. & Gorley, R. N. (2016). Clustering in non-parametric multivariate analyses. *Journal of Experimental Marine Biology and Ecology* 483: 147-155 doi: 10.1016/j.jembe.2016.07.010.

Cutler, D., Deaton, A. & Lleras-Muney, A. (2006). The Determinants of Mortality. *Journal of Economic Perspectives* 20(3), 97–120.

Daoud, J. (2017). Multicollinearity and Regression Analysis. *Journal of Physics Conference Series* 949(1):012009. DOI: 10.1088/1742-6596/949/1/012009.

Fedotenkov, I. and Derkachev, P. (2020). Gender longevity gap and socioeconomic indicators in developed countries. *International Journal of Social Economics* 47(1). 127-144. DOI 10.1108/IJSE-02-2019-0082

GBD 2019 Diseases and Injuries Collaborators. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396: 1204–22.

Hair, J., Anderson, R., Tatham, R., & Black, W. C. (1998). Multivariate data analysis (5th, ed.). Upper Saddle River, N. J.: Prentice Hall.

Hajnal, J. (1965). European marriage pattern in historical perspective. In Glass, D.V.; Eversley, D.E.C. (eds.). Population in History. London: Arnold.

Hand, D. 1981. Discrimination and Classification. New York: John Wiley and Sons.

Hemstrom, O. (2016). Changes in the gender gap in life expectancy in Sweden: A cohort analysis with the most recent trends. In M. Dinges& A. Weigl (Eds.), *Gender-specific life expectancy in Europe 1850–2010* (pp. 149–167). Stuttgart: Franz Steiner Verlag.

Institute for Statistics of FBiH. (2019). Statističkigodišnjak/ljetopis Federacije Bosnei Hercegovine. Statistical Yearbook. Sarajevo. Available at: http://fzs.ba/wp-content/uploads/2020/01/Godisnjak2019.pdf. Retrieved at: 1/10/2020.

Jeong W., Jang, S-I., Park, E-C. & Nam, J. Y. (2020). The Effect of Socioeconomic Status on All-Cause Maternal Mortality: A Nationwide Population-Based Cohort Study. *International Journal of Environmental Research and Public Health* 17, 4606; doi:10.3390/ijerph17124606 www.mdpi.com/journal/ijerph.

Jolliffe, I. & Cadima, J. (2016). Principal component analysis: a review and recent developments, *Philosophical Transactions of the Royal Society* A 374: 20150202. http://dx.doi.org/10.1098/rsta.2015.0202.

Ketchen, D. & Shook C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. *Strategic Management Journal* 17: 441-458.

Le, Y., Ren, J., Shen, J., Li, T. & Zhang, C-F. (2015). The Changing Gender Differences in Life Expectancy in Chinese Cities 2005-2010. *PLoS ONE* 10(4), e0123320. doi:10.1371/journal.pone.0123320.

Lindahl-Jacobsen, R., Hanson, H. A., Oksuzyan, A., Mineau, G. P., Christensen, K., & Smith, K. R. (2013). The male–female health-survival paradox and sex differences in cohort life expectancy in Utah, Denmark, and Sweden 1850–1910. *Annals of Epidemiology 23*(4), 161–166.

Luy, M. (2003). Causes of male excess mortality: insights from cloistered populations. *Population and Development Review* 29(4), 647-676.

Luy, M., & Wegner-Siegmundt, C. (2015). The impact of smoking on gender differences in life expectancy: more heterogeneous than often stated. *European Journal of Public Health, 25*(4), 706–710.

Maiolo, V. & Reid, A. M. (2020). Looking for an explanation for the excessive male mortality in England and Wales since the end of the 19th century. *SSM - Population Health* 100584. https://doi.org/10.1016/j.ssmph.2020.100584.

Martín Cervantes, P.A., Rueda López, N., Cruz Rambaud, S. (2020). Life Expectancy at Birth in Europe: An Econometric Approach Based on Random Forests Methodology. *Sustainability 12*, 413.

Masnan, M. J., Mahat, N. I., Md Shakaff, A. Y., Abdullah, A. H., Ishqi Zakaria, N. Z., Yusuf, N., Subari, N., Zakaria, A. & Abdul Aziz, A. H. (2015). Understanding Mahalanobis distance criterion for feature selection. *AIP Conference Proceedings* 1660, 050075; doi: 10.1063/1.4915708.

Meslé, F., & Vallin, J. (2011). Historical trends in mortality. In R. G. Rogers & E. M. Crimmins (Eds.), International handbook of adult mortality (2nd ed., pp. 9–48). International handbooks of populations Springer: Dordrecht.

Modig K, Rau, R. & Ahlbom A. (2020). Life expectancy: what does it measure? *BMJ Open* 10: e035932. doi:10.1136/ bmjopen-2019-035932.

Nusslelder, W. J. & Looman, C. W. N. (2004). Decomposition of differences in health expectancy by cause. *Demography* 41(2), 315-334.

Pollard, J. H. (1988). On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography* 25, 265-276.

Pollard, J. H. (1982). The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries* 109, 225-240.

Pressat, R. (1985) Contribution des écarts de mortalité par âge à la différence des vies moyennes. *Population* 4-5, 766-770.

Preston, S. H., Heuveline, P. and Guillot, M. (2001). Demography. Measuring and modelling population processes. Maiden: Blackwell Publishers.

Sokal, R. R. & Michener, C. D. (1958) A Statistical Methods for Evaluating Relationships. *University of Kansas Science Bulletin* 38, 1409-1448.

Sokal, R. & Rohlf, J. (1962). The comparison of dendrograms by objective methods. *Taxon* 11, 33-40.

Sokal, R. & Rohlf, J. (1995). Biometry. 3rd edition. New York: W. H. Freeman and Company.

Sundberg, L., Agahi, N., Fritzell, J., & Fors, S. (2018). Why is the gender gap in life expectancy decreasing? The impact of age-and cause-specific mortality in Sweden 1997–2014. *International Journal of Public Health, 63*(6), 673–681.

Trovato, F. & Heyen, N. B. (2006). A varied pattern of change of the sex differential in survival in the G7 countries. *Journal of Biosocial Science* 38, 391-401.

UNDP. 2018. Human Development Indices and Indicators. 2018 Statistical Update. United Nations: New Yoro. Available at: http://hdr.undp.org/sites/default/files/2018_human_development_statistical_update.pdf . Retrieved at: 18/3/2019.

Vallin, J. & Meslé, F. (2004). Convergences and divergences in mortality. A new approach to health transition. *Demographic Research* special collection 2(2), 11-44.

Wisser, O. & Vaupel, J. W. (2014). The sex differential in mortality: A historical comparison of the adult-age pattern of the ratio and the difference. *MPIDR WORKING PAPER WP* 2014-005. JUNE 2014. Available at: https://www.demogr.mpg.de/pa-pers/working/wp-2014-005.pdf. Retrieved 8/12/2020.

Yang, S., Khang, Y. H., Harper, S., Davey Smith, G., Leon, D. A., & Lynch, J. (2010). Understanding the rapid increase in life expectancy in South Korea. *American journal of public health 100*(5), 896–903. https://doi.org/10.2105/AJPH.2009.160341

Yang. S., Khang, Y-H., Chun, H., Harper, S. & Lynch, J. (2012). The changing gender differences in life expectancy in Korea 1970-2005. *Social Science & Medicine* 75, 1280-1287.

Zafeiris, K. N. & Tsoni, A. (2021). Clustering of the Demographic Regimes of Europe at the National Level: A Methodological Assessment. *Population Review* 60(1), 26-51.

Zafeiris, K. N. (2020a). Gender differences in life expectancy at birth in Greece 1994–2017. *Journal of Population Research* 37(1), 73-89. DOI: 10.1007/s12546-019-09239-4.

Zafeiris, K. N. (2020b). Tobacco Related Mortality in Greece: The Effect of Malignant Neoplasms, Circulatory and Respiratory Diseases, 1994–2016. In: Makrides, A., Karagrigoriou, A. & Skiadas Ch. (eds)., *Data Analysis and Applications 4: Financial Data Analysis and Methods, Volume 6*. (pp. 223-247). London: iSTE WILEY.

Zarulli, V., Barthold Jones, J. A., Oksuzyan, A., Lindahl-Jacobsen, R., Christensen, K., & Vaupel, J. W. (2018). Women live longer than men even during severe famines and epidemics. *Proceedings of the National Academy of Science* 115(4), E832–E840. https ://doi.org/10.1073/pnas.17015 35115 .