Disparities in premature mortality: Evidence for the OECD countries

Ana Ledesma-Cuenca[1], *Antonio Montañés*[2] and *María Blanca Simón-Fernández*[3]

[1] University of Zaragoza, Spain. aledesma@unizar.es
[2] University of Zaragoza, Spain. Corresponding author. amontane@unizar.es
[3] University of Zaragoza, Spain. bsimon@unizar.es

Abstract

This paper studies the existence of international health outcome disparities. We focus on the use of the potential years of life lost for a database that includes information from 33 OECD countries and covers the period 1990-2017. The methodology proposed by Phillips and Sul (2007) allows us to reject the existence of a single pattern of behaviour between countries for both males and females, suggesting the existence of severe health outcome inequalities. This methodology estimates the existence of four convergence clubs whose composition slightly varies when comparing the male and female cases. Some socioeconomic factors are found to be very important in explaining the forces that may drive the creation of these convergence clubs. In particular, the evolution of the economy and health policies are pivotal to understanding the creation of these estimated convergence clubs. Additionally, our results offer evidence in favour of the importance of environmental policies to explain these health outcome differences.

Key words: Potential years of life lost, Club convergence analysis, OECD countries, Phillips-Sul, premature mortality.

Funding Sources

This work was supported by the "Fundación Caja Rural de Aragón" Chair of the University of Zaragoza, Spain, the Gobierno de Aragón, grant to the research group Cassetem, and by the Ministerio de Ciencia e Innovación, grant PID2020-113229RB-C44.

Acknowledgments

This version has largely benefited from the comments and suggestions of two anonymous referees. The usual disclaimer applies.

1.- Introduction

The improvement of the health levels of its citizens is one of the main objectives of any society. Governments have continuously been allocating increasing resources to achieve the objectives set by various supranational organizations since the end of the last century. These include the resolution of September 25, 2015 of the United Nations Assembly, commonly known as the 2030 Agenda, which explicitly includes Health as one of its main objectives. Specifically, Sustainable Development Goal 3 aims to ensure healthy lives and promote well-being for all people, regardless of age. To achieve this goal, health policymakers face challenges such as economic and social crises, changes in population structure, evolving disease patterns, increasing health inequalities and environmental transformations that challenge health systems (Rajmil and Fernández de Sanmamed, 2019; McCartney *et al.*, 2019; Gavurova *et al.*, 2020).

The search for improvements in health levels has led to a significant increase in health spending since the second half of the last century, especially in the Organization for Economic Co-operation and Development (OECD) countries, an international group of countries that works to build better policies for better lives. This effort has been fruitful and has involved a significant increment in the health indicators of these countries, reducing mortality rates and increasing the life expectancy of their inhabitants. However, despite these generic improvements, it does not appear that the evolution of these health indicators has followed a similar pattern of behavior, as can be deduced from the results recently obtained by Liou *et al.* (2020) and Xu *et al.* (2021). Rather, it seems that the differences in the health levels between countries have grown. These disparities have generated some strange situations, for example where countries with similar levels of health spending do not exhibit similar levels of health, as noted in Barthold *et al.* (2014). Following Costa *et al.* (2021), this disparity in health levels could be interpreted as the result of an inefficient allocation of resources. This is a serious problem in economics, and it is especially so in a situation in which the COVID 19 crisis has raised public spending to very high levels, seriously aggravating the problems of unsustainability noted in Christopoulos and Eleftheriou (2020b) for the period previous to the outbreak of the pandemic. Health policymakers are therefore faced with the task of maintaining or improving the level of health of their fellow citizens under restrictive economic conditions, as noted by Reibling *et al.* (2019), which makes it extremely necessary to evaluate the evolution of health indicators.

In this regard, we should note that several indicators have been previously used in the literature to assess health system outcomes. The most frequently used are life expectancy at birth, adjusted mortality, infant mortality ratio and premature mortality, as seen in the papers by Martín Cervantes *et al.* (2019), Park and Nam (2019) and Varbanova and Beutels (2020), amongst many others. In our view, a very interesting alternative is the use of the Potential Years of Life Lost, hereafter PYLL. This indicator allows us to assess the loss suffered by society as a result of the death of young people or premature deaths. In addition, the assessment of PYLL rates allows us to advise and help guide state health units, as well as to identify new health problems that require the formation of new community policies (McDonnell *et al.*, 1998). Dickinson and Welker (1948) and Haenszel (1950) pioneered the use of this measure, observing that counting deaths did not give a complete picture of mortality, and that this can be complemented by PYLL. This health indicator is able to capture better the burden of loss and lifetime productivity.

Despite the above-mentioned improvement in health indicators, this does not seem to have behaved similarly in all the countries and it is possible to find disparities in the evolution of this. Quite an interesting method to analyze the existence of these disparities is convergence analysis, as Liou *et al.* (2020) note. In this regard, we should note the existence of a group of very interesting papers that have studied convergence in health indicators using the standard concepts of β- and σ-convergence. We can cite the papers of Jaworska (2014), Richardson *et al.* (2014), Maynou *et al.* (2015), Maynou and Saez (2016), Stańczyk (2016) and Hrzic *et al.* (2021), all of them focused on the European case, and Bremberg (2017) for the OECD case. Their conclusions are similar and the use of the standard concept of β-convergence suggests evidence in favor of convergence.

However, the results of Quah (1993) and Islam (2003) warn us against the use of this procedure, given that negative values of the parameter β from the growth-initial level regression do not necessarily imply a reduction in this dispersion. Then, these authors suggest it is better to focus on the analysis of the dispersion (σ-convergence). Furthermore, we should also take into account the results of Baumol (1986) and Galor

(1996), showing that the β- and σ-convergence analysis may be biased by the existence of convergence or groups of countries/regions that tend to the same steady-state equilibrium. These problems can be easily overcome by using the statistics recently designed by Phillips and Sul (2007, 2009), PS hereafter, which allows us to test the null hypothesis of convergence and, provided we reject it, determine the existence of convergence clubs.

Against this background, the aim of the paper is to analyse the evolution of premature mortality in OECD countries, measured by PYLL. More precisely, we apply the previously-mentioned PS methodology to determine whether this health indicator converges for the OECD countries or, by contrast, several convergence clubs co-exist. We also try to determine whether the Great Recession may have altered the convergence results. In this regard, we should note that this recent crisis dealt a severe blow to public policies (Keegan *et al.*, 2013; Karanikolos *et al.,* 2013; Basu and Bundick, 2017; Parmar *et al.*, 2016), which may have increased the differences in health outcomes between the countries studied (Rajmil and Fernández de Sanmamed, 2019). Then, we consider it of interest to analyse its possible effect. Finally, following the suggestions of Hrzic *et al.* (2020), we try to identify the forces that help us to understand the creation of the convergence clubs.

The rest of the paper is organized as follows. Section 2 describes the data, methodology and analysis framework. Section 3 presents the empirical results of the convergence analyses, a description of the clubs, and a discussion of the results, which includes economic implications and policy recommendations. Finally, Section 4 draws the most important conclusions.

2. Data and methods

2.1. Data

The paper focuses on the analysis of the PYLL. This is quite an interesting measure of health outcome given that if offers information on the premature mortality of a population, which has the capacity of reflecting the effectiveness of the health policies of a country better than alternative indicators such as life expectancy or mortality rate, as discussed in Romeder and McWhinnie (1977) and Gardner and Sanborn (1990). PYLL considers deaths that occur at an early age and may be avoidable. To that end, we first need a reference age limit. In the present case, this reference age limit is 70 years. Subsequently, deaths occurring at each age are summed and then multiplied by the number of years remaining until this reference age limit. The OECD standardizes the PYLL indicator for each country and each year to make comparisons between countries easier. It takes 2010 as the reference population for age standardization and is measured in years lost per 100,000 population, aged 0-69 years. Our sample considers 33 OECD countries for the period 1990-2017 for which we have a complete data set which allows us to disaggregate by gender.

Table 1 presents some descriptive statistics of the PYLL variable. We can observe that there exist some gender differences, with female data showing the lowest values, as Asiskovitch (2010), Stefko *et al.* (2020) and the World Health Organization (WHO, 2021) also point out. Besides this, we can also observe that PYLL has shown a negative growth rate across the sample, although countries do not evolve in the same way. For instance, we note that the United States (female) and Costa Rica (male) exhibit the smallest decline in PYLL, with values close to 0 (-0.7% and -0.9%, respectively). By contrast, Korea shows the highest reduction in PYLL, with average growth rates of -3.6% and -3.8% for female and male, respectively. We can also observe that European countries have had relatively similar decreasing rates. This is not the case of Asian countries (Korea doubles the decreasing rates of Japan) or Latin American countries, where the behaviour of Mexico and Brazil is clearly different to that of Chile.

The analysis of the time evolution of the growth rates also offers us some interesting insights. To that end, we have split the sample in several segments in order to capture the possible effect of the Great Recession on the evolution of this variable. The first segment shows that the growth rates always take negative values for the female and the male cases, although the variation rank is somewhat different. The average growth rates go from -4.2% (Luxembourg) to -1.1% (United States) for the female case, whilst this interval goes from -4.1% (Luxembourg) to -1.4% (Costa Rica) for the male case. If we now consider the post Great Recession sample (2008-2017) we observe that some countries exhibit positive growth rates, especially in the male case, and that the variation of this average growth rate also increases. In particular, the behaviour during the last considered segment is especially heterogenous. We can observe the coexistence of some countries that increased their reduction rate, some others with positive growth rates (especially for the male case), whilst most of the countries experience a slowdown in their decreasing rates.

The results of this initial descriptive analysis point to the existence of considerable variation in the evolution of the PYLL. These differences are summarized in Figures 1a and 1b, which reflect the time evolution of the variation coefficient for the female and the male case, respectively. As we can observe, the cross-sectional dispersion has clearly increased over time, being even higher for the male case than for the female. Furthermore, the dispersion is more marked in the last part of the time-series, which suggests that countries did not exit from the Great Recession in a similar way, with some of them suffering serious losses in the health levels of their inhabitants.

This heterogeneity implies an increment in health inequalities, making difficult the definition and the achievement of global health goals. This can be interpreted as the inexistence of convergence in the health status of the OECD countries, given that convergence implies that the dispersion must decrease. In fact, the results of the analysis shown in Figure 1 do not suggest the existence of σ-convergence. Rather, we should conclude that both male and female PYLL diverge. However, this is not the most appropriate method to determine convergence, given that we do not test this hypothesis directly. The following section presents the methodology that we will employ to that end.

2.2. Convergence methodology

Convergence has been defined in the economic literature as a process where the dispersion of a variable reduces for a group of countries or regions. At the limit, when the cross-sectional variance is 0, all the components of this group show the same value of the variable. The interest in this type of analysis grew due to the seminal paper by Barro and Sala-i-Martin (1992), who focused their analysis on the per capita Gross Domestic Product (GDP). This paper was the seed of a very large literature, mainly focused on the use of the β- and σ-convergence concepts. However, the criticism of the use of the β-convergence notion made by Quah (1993) and Friedman (1992), based on the results of Galton (1877), suggested the use of alternative methods to analyse convergence. To that end, Carlino and Mills (1993, 1996) and Bernard and Durlauf (1995) developed a somewhat different approach, based on the use of unit root test statistics to test the time series properties of the ratio of the per capita GDP (or an alternative socioeconomic measure) of two geographical units (i.e. countries, regions, provinces). The presence of a unit root in the above-mentioned ratio was considered as evidence against convergence.

However, none of these papers develop or use a statistic that focuses on testing the null hypothesis of convergence. This problem is considered in Phillips and Sul (2007, 2009), PS hereafter, who designed a very popular statistic that has been extensively employed to test for convergence. Due to its capacity to analyse the existence of disparities in the evolution of these variables, it comes as no surprise that this statistic has been recently applied to the convergence analysis of some health variables. For instance, Panopoulou and Pantelidis (2013) and Clemente *et al.* (2019a, 2020) studied health care expenditure in the United States and Clemente *et al.* (2019b) studied the case of Spanish regional health care expenditure. Closely related to this, Duncan and Toledo (2019), Kasman and Kasman (2020) and González-Álvarez *et al.* (2020) analysed the international evolution of obesity indexes. Finally, Panopoulou and Pantelidis (2012) examined some health outcomes for a group of OECD countries, whilst Christopoulos and Eleftheriou (2020a) focused on the potential years of life lost in the United States. As we can observe, the use of this methodology is becoming a very useful tool to determine the existence of different patterns of behaviour of some health variables. Therefore, it seems appropriate to apply it to the present case.

The methodology is defined as follows. Let us consider that X_{it} represents the log of the potential years of life lost, with i=1, 2, …, 33 (the considered OECD countries) and t= 1990, …, 2017. This variable can be decomposed as $X_{it} = \delta_{it}\ \mu_t$, where μ_t is the single common component and δ_{it} is the time-varying factor loading coefficient that measures the idiosyncratic distance between the common trend components μ_t and X_{it}. PS suggest testing for convergence by analysing whether δ_{it} converges for some δ. To do so, they first define the relative transition parameter, as follows:

$$h_{it} = \frac{X_{it}}{N^{-1}\sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{N^{-1}\sum_{i=1}^{N} \delta_{it}} \qquad (1)$$

This variable describes the transition path for the i-th country relative to the panel average. In the presence of convergence, δ_{it} converges towards δ and, therefore, h_{it} should converge towards 1, while its cross-sectional variation, H_{it}, which is defined as follows:

$$H_t = N^{-1} \sum_{i=1}^{N}(h_{it} - 1)^2 \to 0, as\ T \to \infty \quad (2)$$

should go to 0 when T goes towards infinity. Then, PS test for convergence by estimating the following equation:

$$log\frac{H_1}{H_t} - 2log[\log(t)] = \alpha + \beta \log(t) + u_t, t = T_o, \ldots, T \quad (3)$$

with $T_o = [rT]$, and r=0.3, and u_t being the perturbation of the model. Equation (3) is commonly known as the log-t regression. The null hypothesis of convergence is rejected whenever parameter β is lower than 0. PS suggest estimating model (3) by methods that take into account the possible presence of autocorrelation and heteroskedasticity and, later, building the corresponding t-statistic to test the null hypothesis β=0. The use of these robust methods ensures that this t-ratio converges towards a standard N(0,1) distribution and, therefore, we will reject the null hypothesis of convergence whenever this t-statistic takes values lower than -1.65.

If we reject convergence, PS propose the following robust clustering algorithm for identifying clubs in a panel:

i. Order the N countries according to their final values of the variable under analysis (PYLL in our case).

ii. Form the core group. This core group consists of the k countries that maximize the value of the convergence t-statistic, subject to the restriction that it is greater than -1.65 and that 2<=k<N.

iii. Add one country at a time of the remaining countries to the core group, and re-estimate model (3) for each formation. Use the t-statistic to decide whether a country should join the core group.

iv. Repeat steps (ii)–(iii) iteratively and stop when clubs can no longer be formed. If the last group does not have a convergence pattern, conclude that its members diverge.

More details on this procedure can be found in Du (2017).

PS recommend performing club merging tests after running the algorithm using equation (3) in order to avoid an over-estimation of the number of clubs. Finally, we have followed the suggestion of PS and extracted the trend components of the series by filtering them using the Hodrick and Prescott (1997) filter, applying the standard value λ=400.

3. Results and discussion

3.1. Convergence

The results of the of the Phillips-Sul methodology are presented in Table 2. Panel I shows that the null hypothesis of convergence is rejected for both male and female PYLL. This implies that this health indicator presents great differences in its evolution, such that we cannot observe a common behaviour across the considered sample. However, it is still possible that some countries show similarities and, consequently, the existence of some convergence clubs. We should note that a convergence club is a statistically created grouping to capture the intrinsic PYLL trajectory shared by several countries, as Liou *et al.* (2020) state. The use of the clustering algorithm of PS allows us to estimate the existence of 4 convergence clubs for both female and male cases, as Panel II reflects. It is relevant to recall the ordering nature of these estimated clubs to better interpret them. Then, we should note that Club 1 includes the countries with the highest values of PYLL. By contrast, Club 4 is made up of the countries that perform best, in the sense that they show the lowest PYLL values. Panel III of Table 2 shows the results of the PS club merger tests performed to avoid overestimating the number of clubs, but the results indicate that there is no convergence between adjacent clubs, so the final number of estimated clubs remains unaltered.

To make the interpretation of the results easier, Figures 2a and 2b present the maps with the estimated clubs for the female and the male case, respectively. If we analyse the composition of these clubs, we can observe that the countries included in Club 1 are Brazil, Mexico and the United States for the female case, and Brazil, Hungary and Mexico for the male case. Similarly, Club 2 includes Canada, three Latin American countries (Chile, Colombia and Costa Rica) and three European countries (Hungary, Poland and United Kingdom) for the female case. For the male case, the composition slightly varies, in that Club 2 includes the United States, the three Latin American countries (Chile, Colombia and Costa Rica) and three European countries (Czech Republic, Greece and Poland). Club 3 and Club 4 are composed of European countries and the Asian and Oceanic countries (Korea, Japan and Australia). We can see that Australia, Belgium, France and Germany are included in Club 3 for both cases, whilst Ireland, Israel, Italy, Korea, Luxembourg, Norway, Spain, Sweden and Switzerland always belong to Club 4.

As we have seen, there are some differences in the composition of the estimated convergence clubs when comparing the female and the male case. To analyse whether the differences are significant, we have constructed the variables CF_i and CM_i, which take the value m if the i-th country is included in the m-th estimated club, with m=1, 2, 3, 4 and CF and CM reflecting the female and the male case, respectively. Later, we have calculated the statistics proposed by Kruskal and Wallis (1952) and Van der Waerden (1953) for these variables in order to test for the equality of the median of the samples. These statistics take the values 0.22 and 0.27, respectively. They asymptotically follow a χ^2 of n-1 degrees of freedom, with n being the number of variables considered (2 in the present case). Then, we should conclude that the differences are not significant and the cluster algorithm provides similar results for the two gender cases.

If we now focus on some particularly interesting countries, we find that the case of the United States is noteworthy, since this country is included in the estimated clubs with the largest values of PYLL. The case of Germany is also noticeable given that it is never included in the best estimated club (Club 4), whilst this country belongs to Club 3 for both genders. By contrast, we observe that the Mediterranean countries such as Spain and Italy perform as well as Scandinavian countries such as Denmark, Iceland, Norway and Sweden.

Figures 3a and 3b reflect the evolution of average values of the PYLL for the different estimated clubs. It is observed that the evolution of the PYLL clearly declines for all the cases. We can also appreciate some gender and time differences. The gender comparison shows that the decline in the PYLL is slightly greater for the male case than for the female case. For instance, the average growth rate of Club 1 is -1.7% and -1.4% for the male and the female case, respectively. Similarly, these growth ratios are -2.9% and -2.5% for Club 4.

If we now focus on the time evolution, we can also draw some interesting insights. To that end, we can split the sample into two segments, 1990-2008 and 2008-2017, in order to capture the effect of the Great Recession on the evolution of the PYLL. For both the female and the male cases, we can observe that the Great Recession slowed down the process of improvement in the health indicator. For instance, the average growth rates of Club 1 for the periods pre and post Great Recession are -1.7% and -0.8%, for the female case, and -2.2% and -1.3%, for the male case. Similarly, the growth rates for the estimated Club 4 are -3.1% and -2.1%, for the female case, and -3.6% and -2.6%, for the male case. The values for the Clubs 2 and 3 are qualitatively similar, although slightly lower in quantitative terms.
We can also draw some insights by analysing the evolution of the estimation of the parameter β in (3). Figure 4 shows its estimated values when the sample goes from 1990 to k, with k=2004-2017. As we can appreciate, the estimated value remains almost unaltered until the onset of the Great Recession. However, it clearly decreases after 2010. This simple analysis, in the sense that it is just based on the evolution of the log t-ratio statistic, is quite useful to show that disparities did not cease growing after that period, although cannot help us to understand why this has occurred. In any event, it clearly suggests the existence of an effect of the Great Recession on the evolution of the health outcomes, and quite probably on some other socioeconomic variables, supporting the existence of a strong relationship between economics and health.

3.2. Factors driving Club formation

The previous section has shown how OECD countries do not exhibit similar levels of PYLL. Rather, we have estimated the existence of 4 convergence clubs for the female and the male case. It is sensible now to investigate which factors may drive the formation of these estimated clubs. To that end, we have estimated an ordered logit model. This type of model has been selected because the dependent variables are ordinal and ranked in descending order according to the different steady states. The initial model specification is as follows:

$$y_i = x_i'\beta + u_i \qquad (i=1,2,\dots,33) \tag{4}$$

where the dependent variable y_i may have various possible outcomes, each of them related to the number of clubs that the PS methodology has estimated. Then, y_i= m if the i-th country is included in the m-th estimated club, with m=1, 2, 3, 4. On the other hand, the vector of independent variables x_i' reflects the different factors that could explain the drivers of the estimated clubs.

As a previous step to the estimation of the ordered logit, we have considered a set of general socio-economic variables that are commonly used in the literature as possible explanatory variables in the model. For this purpose, we have followed the papers of Lynch *et al.* (2004), López-Casasnovas and Soley-Bori (2014), James (2015) and Gavurova *et al.* (2020), amongst some others, where the determinants of health outcomes are analysed.

The set of potential explanatory variables is presented in Table 3, whilst their definition and the source of the data are reflected in the Appendix. Table 3 also includes the average values of these variables for the different estimated clubs, with these averages based on the use of the last observed value of each explanatory variable. In what follows, we present these variables and analyse the average values in order to better understand their powers of discrimination.

The first group of variables is composed of some socioeconomic indicators. In this regard, we should note that the results of some previous studies, such as those of Varbanova and Beutels (2020) and Gavurova *et al.* (2020), reveal the importance of variables such as the per capita Gross Domestic Product (GDP), the Gini index, the Human Capital index and the unemployment rate. The results shown in Table 3 confirm this, and we can observe a clear relationship between economy and health in such a way that the worse the economic conditions, the worse the health outcomes. To see this, we should simply note that the average Gini index indicates that Clubs with the highest premature deaths show the highest degree of income inequality among the population. Similarly, the per capita GDP and the Human Capital Index have an inverse relationship with PYLL. The higher the economic and social development, the lower the premature deaths for both female and male populations (López-Casasnovas and Soley-Bori, 2014).

The unemployment ratio, however, does not provide any clear discrimination between clubs. A possible explanation is given by the fact that unemployment, especially structural unemployment, is influenced by institutional factors, such as the performance of trade unions and employers' organizations in terms of collective bargaining and dismissals (Van Gool and Pearson, 2014).

Another interesting group of variables is that which represents the status of the health system of each country. Then, following Panopoulou and Pantelidis (2012), Park and Nam (2019) and González-Álvarez *et al.* (2020), we considered variables such as the degree of population immunization, life expectancy, mortality, the prevalence of some diseases and per capita health expenditure.

Table 3 shows that measles immunization and polio immunization are directly related to premature mortality (the higher the immunization rate of the population, the more probable the country will be included in clubs with the lowest values of PYLL). We also observed that measles immunization discriminates better against clubs than polio immunization. By contrast, mortality caused by cardiovascular disease, cancer or diabetes has a different sign, but a similar interpretation, given that the higher the average values of these variables, the more probable the country will be included in the clubs with the highest premature mortality. Both results are expected, especially if we take into account the results of Meslé *et al.* (2002) and Vallin and Meslé (2004) who consider that the capacity of countries to prevent mortality caused by these diseases (via the adoption of new technologies) is key to understanding the differences in mortality between countries.

The analysis of the means of the prevalence of diabetes, anemia and overweight persons indicates that as the prevalence of these diseases decreases, premature mortality is lower (we are in better clubs). It has to be pointed out that overweight discriminates less for males than for females; however, anemia and diabetes discriminate equally for both.

By contrast, we cannot observe a linear link between Current Health Expenditure and PYLL. Rather, we can observe that, for the female population, the average of this variable in Club 1 is the highest, then it decreases between Club 1 and Club 2, and increases again between Club 2 and Club 3, before decreasing again in Club 4. The behaviour of this variable for the male case is the opposite: the Current Health Expenditure increases for the first Clubs as we move from the first Club to the third Club but decreases again in Club 4 (Table 3).

Finally, we have also considered some variables that can offer a view on the political stability of the country, as Mackenbach *et al.* (2013), Zare *et al.* (2015) and Woolf and Schoomaker (2019) have done. Following these papers, we have used the government effectiveness, control of corruption and voice accountability variables.

The inclusion of some governance variables is relevant in order to reveal the influence of institutions on healthcare. In our case, the results of Table 3 suggest that the lower the control of corruption, the higher the premature mortality. However, the higher the effectiveness of governance, the lower the premature mortality. Similarly, we observe that the greater the degree of participation of a country's citizens in the election of their government, as well as freedom of expression, freedom of association and freedom of the media, the lower the premature mortality of its inhabitants.

Finally, we consider that the environmental policy may clearly influence the health of the citizens of a country. Consequently, our set of the potential drives should include some environmental variables. Following the results of Leksell and Rabl (2001) and Lelieveld *et al.* (2015) or Samoli *et al.* (2019), who have highlighted the importance of the impact of air pollution on health, we have additionally considered carbon dioxide damage and greenhouse gas emissions.

The results of Table 3 are quite useful for providing solid evidence of the link between environment and health. For both males and females, the discriminatory power of this variable is noticeable, and those countries included in the clubs with the lowest PYLL values exhibit the lowest levels of CO_2 emissions. This result is better understood if we take into account the connection between mortality caused by cardiovascular disease and air pollution. According to the World Health Organisation (WHO, 2021), 9 out of 10 people are exposed to levels of air pollution that put them at increased risk of cardiovascular disease. These types of diseases are the leading cause of death worldwide, claiming an estimated 17.9 million lives each year. By contrast, greenhouse gas emissions seem to exhibit a lower predictive capacity, as can be observed in Table 3.

Once we have presented the group of potential explanatory variables, we should employ them to estimate the best possible logit model. To that end, we have first pre-selected those variables that can help us to discriminate the inclusion of a country in the different estimated clubs based on the average values included in Table 3. Later, we have used them in a general-to-particular strategy, removing the non-significant variables. The final estimated model is presented in Table 4.

The male model includes the following variables: the Gini index, per capita GDP, measles immunization, mortality from cardiovascular diseases, cancer or diabetes, and carbon dioxide damage. The female model is somewhat different. It does not include per capita GDP and carbon dioxide damage but does include health expenditure. The explanatory power of both models is high: the female model correctly classifies 76% of the countries, whilst this figure for the male model is even slightly higher (79%).

The results of the ordered logit estimation show that the Gini index for both female and male has a negative effect on premature deaths. Then, if the level of inequality increases, there is a greater probability of being in the worst club and therefore that the population loses more potential years of life. The importance of inequality is greater for the female model, indicating a different incidence in the effect of inequality on premature deaths depending on the gender considered. This result is better understood if we take into account that, in spite of considering developed countries, females still have, on average, less access to education, earn less than males and enjoy less economic independence in these countries. This translates into females having fewer health resources and less access to health care (Arber and Thomas, 2001; Bartley, 2004; Kawachi and Kennedy, 1999; Kawachi *et al.*, 1999), and the gender gap still exists in this respect.

By contrast, we can observe that the per capita GDP is positively related to the creation of male clubs and we can conclude that the economic development of countries allows the male population to lose fewer potential years of life. This result is consistent with several previous studies which showed that countries' development is positively related to national levels of health (Lynch *et al.*, 2004; Torre and Myrskylä, 2014; Christopoulos and Eleftheriou, 2020a). Our results indicate that greater availability of resources and improvements in lifestyle and nutrition, as Ashraf *et al.* (2008) note, translates into reduced premature mortality and increased life expectancy.

The negative coefficient that accompanies Current Health Expenditure in the female estimation confirms the results obtained from our initial descriptive analysis. Then, as current health expenditure increases, there is a greater possibility of being in a worse club and therefore that the female population, in this case, loses more potential years of life. This variable is relevant and can explain why some countries are placed in unexpected clubs, as in the case of the United States. It is interesting to note that OECD countries have the highest health expenditure in the world (Mujtaba and Shahzad, 2021), so one would expect that they should also show the best health outcomes. However, the negative sign suggests that this is not necessarily true, opening the door to using efficiency as a key factor in this regard, as we will discuss later.

We can also observe the importance of the variables that measure the health status of the country. In particular, measles immunization appears in the final estimated model. In this regard, we should note that vaccination against diseases mainly affects the first years of life, having a positive and long-term effect on the life expectancy of the population and a negative effect on infant mortality (Park and Nam, 2019). According to the World Health Organization (WHO, 2021), immunization is a global health and development success story. Vaccines are available to prevent more than 20 life-threatening diseases, helping people of all ages to live longer, healthier lives. For example, vaccination against diphtheria, tetanus, pertussis and measles prevents between two and three million child deaths each year, and many more future deaths in older age groups (Duclos *et al.*,2009). Our results show that measles immunization has a positive coefficient for both males and females, and premature deaths are reduced when the percentage of the population vaccinated against measles increases. Measles immunization is the type of vaccination that allows us to analyse total vaccination coverage, since it is a prototypical health service.

Mortality due to cardiovascular disease, cancer or diabetes has a direct effect on the increase in premature deaths, as already indicated by the averages. It is noteworthy that the environmental variable is only relevant for the male estimate and with a negative effect on clubs. This result is supported by the research of Abbey *et al.* (1999) who found a positive correlation of air pollution with mortality for men, but not for women. These results may be due to limitations in the way data on air pollution deaths are measured, as the World Health Organization (WHO) uses national estimates of solid fuel use to calculate indoor air pollution deaths that do not consider possible differential risks for women within societies or within the household (Austin and Mejia, 2017). However, some studies along these lines also estimate that more men die from air pollution than women (Lim *et al.,* 2012; WHO, 2021).

3.3. Discussion of results and implications/recommendations

The results we have obtained lead us to draw very useful insights. First, we should note that the use of the PS methodology has allowed us to reject the convergence hypothesis, which statistically confirms the existence of different patterns of behaviour. Moreover, we can also observe that this disparity has grown since 2008. Then, the supra institutions such as the United Nations or the European Union should consider these disparities when setting their goals.

We have subsequently seen that we can group countries according to their PYLL evolution. The analysis of the forces that create the estimated clubs also produces some interesting implications. If we begin by considering the effect of per capita GDP on the creation of the estimated clubs, we can observe that this variable has a positive effect on the probability of being included in the club with the lowest premature male mortality. Previous researches have already pointed out that the higher the economic development, the more advanced the welfare state of a country. However, Widding-Havneraas and Pedersen (2020) and Asiskovitch (2010) show that a higher wealth status does not necessarily lead to improvements in people's life expectancy if this is not accompanied by a more equitable distribution of material resources that directly or indirectly improve health. Low-income individuals appear to be more vulnerable to health limitations (López-Casasnovas and Soley-Bori, 2014), so reducing income inequality would improve the health of the most disadvantaged population and help reduce health inequalities (Bremberg, 2017).

In connection with this point, we should note that the United States is the country where income inequality is most closely related to population health because the composition of United States health spending is mainly from private sources (Lynch *et al.,* 2004). An examination of how the public-private composition of health spending affects health outcomes shows that when health care is financed more by public spending, it is associated with lower rates of premature mortality for both genders (Or, 2000).

Another interesting insight is that our estimations show that health care expenditure only appears in the female model. Furthermore, this estimated parameter takes a negative value, which suggests that increments in the level of health care expenditure raise the probability of a country being included in Club 1. In order to explain this (apparently) counterintuitive result, we should consider the existence of decreasing returns of health care spending in developed countries, as Nixon and Ulmann (2006) point out. These authors suggest the presence of a saturation point for health spending in developed countries, in such a way that increments in the levels of health care expenditure do not generate proportional variations in the health of the population. This is also supported by the results of Poullier *et al.* (2002), Potrafke (2010) and Chansarn (2010), who note that health care expenditure in high-income countries brings only marginal improvements in life expectancy. The existence of a problem of efficiency in the use of health expenditure is also concluded in López-Casasnovas and Soley-Bori (2014), Ravangard *et al.* (2014), Blázquez-Fernández *et al.* (2017), Linden and Ray (2017), Costa (2021) and Costa *et al.* (2021). The results of Christopoulos and Eleftheriou (2020b) are useful in this regard, showing that the marginal benefit of every additional dollar spent by the OECD countries on healthcare is pretty close to 0.

The presence of an environmental variable also deserves a comment, given that this provides additional evidence of the direct connection between health and the environment. We should note that, according to WHO results, ambient air pollution is responsible for 4.2 million premature deaths each year as a result of exposure to fine particulate matter. To better understand the effect of this “silent killer”, to use the denomination employed in Austin and Mejia (2017) and Zhang (2017), we should note that air pollution is estimated to cause 6-9 million premature deaths per year and will cost 1% of OECD countries' GDP by 2060 (OECD, 2021). Therefore, controlling pollution is a key factor not only to achieve greener and more sustainable economic growth, but also to avoid unnecessary damage to people's health.

Finally, we recognise that the estimation of the ordered logit model is not free of problems, mainly due to the scantiness of the sample, as Ogundimu *et al.* (2016) reflect. In spite of this, the combination of the results presented in Tables 3 and 4 are very helpful to identify which the most relevant factors are. We should also note that a causality analysis would be a good complement to the analysis carried out here, which would require the use of somewhat different econometric tools. Therefore, this is left for future research.

4. Conclusions

The aim of this paper has been to analyse the existence of disparities in the evolution of some international health indicators. For this purpose, we have selected the potential years of life lost variable, given that it can summarize the general health status of a society better than alternative indicators such as life expectancy or the mortality rate. The considered database covers the period 1990-2017 and includes information of 33 OECD countries, with data being disaggregated by gender.

We have employed the methodology designed by Phillips and Sul (2007, 2009). Its use has allowed us to reject the null hypothesis of convergence for the female and the male case. This implies that the 33 OECD countries do not have a common pattern of behaviour in terms of potential years of life lost. Subsequently, we have analysed whether some convergence clubs may exist, by applying the clustering algorithm proposed by Phillips and Sul. We have estimated the existence of 4 convergence clubs that exhibit small and non-statistically significant gender differences. The analysis of the time evolution of these estimated clubs shows that the PYLL has declined for the sample, although the speed of this decline has clearly slowed down since the Great Recession.

We have also analysed the drivers of club formation. The explanatory variables finally included in our estimated models are the Gini index, current health expenditure, per capita GDP, measles immunization, mortality from cardiovascular disease, cancer or diabetes, and carbon dioxide damage. We find gender differences in the estimation of the model:

for females, the Gini index, current health expenditure, measles immunization and mortality from cardiovascular disease, cancer or diabetes are important; while for males the Gini index, per capita GDP, immunization rates, mortality from cardiovascular disease, cancer or diabetes, and carbon dioxide damage are relevant.

Per capita GDP and measles immunization have a positive effect on potential years of life lost. The economic development of a country is positively related to good national health levels, while vaccination against diseases that mainly affect the first years of life has a positive and long-term effect on the life expectancy of the population. Thus, growth must be accompanied by public health policies that maintain adequate levels of health coverage and consider the health benefits derived from advances in vaccines.

On the other hand, the Gini index, mortality due to cardiovascular disease, cancer or diabetes, per capita current health expenditure and the cost of damage due to carbon dioxide emissions have a negative effect on premature mortality. Health inequalities arise from unequal income distribution and discrepancies related to access to health care, education, work and living conditions. The United States is the country where income inequality is most related to the health of the population; even though current United States healthcare spending has increased in recent years, its impact on improving the population's health is not apparent.

To reduce the incidence of premature mortality from cardiovascular disease, cancer or diabetes, the development of cost-effective, accessible and equitable healthcare innovations is necessary for their management. Finally, the cost of damage due to carbon dioxide emissions has a negative effect on the health of the population. In view of the results and for future air pollution scenarios, public agencies should choose to further restrict air pollution, as it directly affects the health of the population.

The results obtained are relevant for future health scenarios, especially for countries where air pollution continues to increase and where suitable emission levels are not expected to be reached to prevent the population from developing cardiorespiratory diseases. The development of comprehensive strategies, cost-effective health innovations and equitable management, such as improving the capacity of health systems to care for patients with cardiovascular issues and monitoring disease patterns and trends, is needed to reduce the incidence, morbidity, and mortality of the disease. Likewise, to avoid the adverse effects of air pollution, governments need to develop policies that support effective environmental actions as a key aspect of future health improvement. Governments should consider that to reduce premature mortality, growth must be accompanied by reductions in health and economic inequality.

References

Abbey, D. E., Nishino, N., McDonnell, W. F., Burchette, R. J., Knutsen, S. F., Beeson, W. L., & Yang, J. X. (1999). Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *American Journal of Respiratory and Critical Care Medicine, 159*, 373–382.

Arber, S., & Thomas, H. (2001). From women's health to a gender analysis of health. *The Blackwell companion to medical sociology*, 94-113.

Ashraf, Q. H., Lester, A., & Weil, D. N. (2008). When does improving health raise GDP?. *NBER macroeconomics annual*, 23(1), 157-204.

Asiskovitch, S. (2010). Gender and health outcomes: The impact of healthcare systems and their financing on life expectancies of women and men. *Social Science & Medicine*, *70*(6), 886-895.

Austin, K. F., & Mejia, M. T. (2017). Household air pollution as a silent killer: women's status and solid fuel use in developing nations. *Population and Environment*, *39*(1), 1-25.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), 223-251.

Barthold, D., Nandi, A., Mendoza Rodríguez, J. M., & Heymann, J. (2014). Analyzing whether countries are equally efficient at improving longevity for men and women. *American journal of public health*, *104*(11), 2163-2169.

Bartley, M. (2004). Health inequality: An introduction to theories, concepts and methods. *Cambridge: Polity Press.*

Basu, S., & Bundick, B. (2017). Uncertainty shocks in a model of effective demand. *Econometrica*, *85*(3), 937-958.

Baumol, W. J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review,* 1072-1085.

Bernard, A. B., & Durlauf, S. N. (1995). Convergence in international output. *Journal of Applied Econometrics*, 10(2), 97-108.

Blane, D., Smith, G. D., & Bartley, M. (1990). Social class differences in potential years of life lost: size, trends, and principal causes. *British Medical Journal, 301*(6749), 429-432.

Blázquez-Fernández, C., Cantarero-Prieto, D., & Pascual-Sáez, M. (2017). Health expenditure and socio-economic determinants of life expectancy in the OECD Asia/Pacific area countries. *Applied Economics Letters, 24*(3), 167-169.

Bremberg, S. G. (2017). Mortality rates in OECD countries converged during the period 1990–2010. *Scandinavian journal of public health*, *45*(4), 436-443.

Carlino, G. A., & Mills, L. (1996). Testing neoclassical convergence in regional incomes and earnings. *Regional Science and Urban Economics*, 26(6), 565-590.

Carlino, G. A., & Mills, L. O. (1993). Are US regional incomes converging? A time series analysis. *Journal of monetary economics, 32*(2), 335-346.

Chansarn, S. (2010). Labor Productivity Growth, Education, Health and Technological Progress: A Cross-Country Analysis. *Economic Analysis & Policy*, *40*(2).

Christopoulos, K., & Eleftheriou, K. (2020a). Premature mortality in the US: A convergence study. *Social Science & Medicine*, *258*, 113141.

Christopoulos, K., & Eleftheriou, K. (2020b). The fiscal impact of health care expenditure: Evidence from the OECD countries. *Economic Analysis and Policy*, 67, 195-202.

Clemente, J., Lázaro-Alquézar, A., & Montañés, A. (2019a). US State health expenditure convergence: A revisited analysis. *Economic Modelling*, *83*, 210-220.

Clemente, J., Lázaro-Alquézar, A., & Montañés, A. (2020). Does the Great Recession contribute to the convergence of health care expenditures in the US States? *International journal of environmental research and public health*, *17*(2), 554.

Clemente, J., Lázaro-Alquézar, A., and Montañés, A. (2019b). Convergence in Spanish Public health expenditure: Has the decentralization process generated disparities? *Health Policy, 123*(5), 503-507.

Costa, E. (2021). Modeling Health Spending Financial Sustainability. *Available at SSRN: https://ssrn.com/abstract=3886591*

Costa, E., Santos, R. & Barros, P.P. (2021). The Financial Sustainability of the Portuguese Health System, Baltagi, B.H. and Moscone, F. (Ed.) *The Sustainability of Health Care Systems in Europe (Contributions to Economic Analysis, Vol. 295),* Emerald Publishing Limited, Bingley, 209-229.

Dickinson, F. G., & Welker, E. L. (1948). What is the leading cause of death? Two new measures. Bureau of Medical Economic Research, *American Medical Association. Bulletin (Chicago)*, 64.

Du, K. (2017). Econometric convergence test and club clustering using Stata. *The Stata Journal*, 17(4), 882-900.

Duclos, P., Okwo-Bele, J. M., Gacic-Dobo, M., & Cherian, T. (2009). Global immunization: status, progress, challenges and future. *BMC international health and human rights, 9*(1), 1-11.

Duncan, R., & Toledo, P. (2019). Inequality in body mass indices across countries: Evidence from convergence tests. *Economics & Human Biology*, 33, 40-57.

Friedman, M. (1992). Do old fallacies ever die?. *Journal of Economic Literature, 30*(4), 2129-2132.

Galor, O. (1996). Convergence? Inferences from theoretical models. *The economic journal*, *106*(437), 1056-1069.

Galton F. (1877). Typical laws of heredity. *Proceedings of the Royal Institution, Vol.8,* 282–301.

Gardner, J. W., & Sanborn, J. S. (1990). Potential years of life lost (PYLL)—what does it measure? *Epidemiology*, 322-329.

Gavurova, B., Khouri, S., Kovac, V., & Ferkova, M. (2020). Exploration of Influence of Socioeconomic Determinants on Mortality in the European Union. *International Journal of Environmental Research and Public Health*, *17*(13), 4699.

González-Álvarez, M. A., Lázaro-Alquézar, A. & Simón-Fernández, M. B. (2020). Global Trends in Child Obesity: Are Figures Converging? *International Journal of Environmental Research and Public Health, 17*(24): 9252.

Hadad, S., Hadad, Y., & Simon-Tuval, T. (2013). Determinants of healthcare system's efficiency in OECD countries. *The European journal of health economics*, *14*(2), 253-265.

Haenszel, W. (1950). A standardized rate for mortality defined in units of lost years of life. *American Journal of Public Health and the Nations Health, 40*(1), 17-26.

Hertz, E., Hebert, J. R., & Landon, J. (1994). Social and environmental factors and life expectancy, infant mortality, and maternal mortality rates: results of a cross-national comparison. *Social Science & Medicine*, *39*(1), 105-114.

Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 1-16.

Hrzic, R., Vogt, T., Brand, H., & Janssen, F. (2021). The Short-Term Effects of European Integration on Mortality Convergence: A Case Study of European Union's 2004 Enlargement. *European Journal of Population, 37*(4), 909-931.

Hrzic, R., Vogt, T., Janssen, F., & Brand, H. (2020). Mortality convergence in the enlarged European Union: a systematic literature review. *European Journal of Public Health, 30*(6), 1108-1115.

Islam, N. (2003). What have we learnt from the convergence debate? *Journal of economic surveys, 17*(3), 309-362.

James, J. (2015). Health and education expansion. *Economics of Education Review*, 49, 193-215.

Jaworska, R. (2014). Health inequalities across the European Union regions: A beta-convergence approach. *Comparative Economic Research, 17*(4), 71-86.

Karanikolos, M., Mladovsky, P., Cylus, J., Thomson, S., Basu, S., Stuckler, D., Mackenbach, J. P. & McKee, M. (2013). Financial crisis, austerity, and health in Europe. *The Lancet*, *381*(9874), 1323-1331.

Kasman, S., & Kasman, A. (2020). Convergence in obesity and overweight rates across OECD countries: Evidence from the stochastic and club convergence tests. *Empirical Economics*, 1-34.

Kawachi, I., & Kennedy, B. P. (1999a). Income inequality and health: pathways and mechanisms. *Health services research*, *34*(1 Pt 2), 215.

Kawachi, I., Kennedy, B. P., Gupta, V., & Prothrow-Stith, D. (1999b). Women's status and the health of women and men: a view from the states. *Social Science & Medicine, 48*, 21–32.

Keegan, C., Thomas, S., Normand, C., & Portela, C. (2013). Measuring recession severity and its impact on healthcare expenditure. *International journal of health care finance and economics*, *13*(2), 139-155.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association, 47*(260), 583-621.

Leksell, I., & Rabl, A. (2001). Air pollution and mortality: quantification and valuation of years of life lost. *Risk analysis*, *21*(5), 843-843.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, *525*(7569), 367-371.

Lim, S. S.,Vos, T., Flaxman,A.D., Danaei,G., Shibuya,K., Adair-Rohani,H., *et al.* (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010. *The Lancet, 380*(9859), 2224–2260.

Linden, M., & Ray, D. (2017). Life expectancy effects of public and private health expenditures in OECD countries 1970–2012: Panel time series approach. *Economic Analysis and Policy, 56*, 101-113.

Liou, L., Joe, W., Kumar, A., & Subramanian, S. V. (2020). Inequalities in life expectancy: An analysis of 201 countries, 1950–2015. *Social Science & Medicine*, 253, 112964.

López-Casasnovas, G., & Soley-Bori, M. (2014). The socioeconomic determinants of health: economic growth and health in the OECD countries during the last three decades. *International Journal of Environmental Research and Public Health, 11(1)*, 815-829.

Lynch, J., Smith, G. D., Harper, S. A., Hillemeier, M., Ross, N., Kaplan, G. A., & Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *The Milbank Quarterly*, *82*(1), 5-99.

Mackenbach, J. P., Hu, Y., & Looman, C. W. (2013). Democratization and life expectancy in Europe, 1960–2008. *Social Science & Medicine, 93*, 166-175.

Martín Cervantes, P. A., Rueda López, N., & Cruz Rambaud, S. (2019). A causal analysis of life expectancy at birth. Evidence from Spain. *International journal of environmental research and public health*, *16*(13), 2367.

Maynou, L., & Saez, M. (2016). Economic crisis and health inequalities: evidence from the European Union. *International journal for equity in health, 15*(1), 1-11.

Maynou, L., Saez, M., Bacaria, J., & Lopez-Casasnovas, G. (2015). Health inequalities in the European Union: an empirical analysis of the dynamics of regional differences. *The European Journal of Health Economics, 16*(5), 543-559.

McCartney, G., Hearty, W., Arnot, J., Popham, F., Cumbers, A., & McMaster, R. (2019). Impact of political economy on population health: a systematic review of reviews. *American journal of public health*, *109*(6), e1-e12.

McDonnell, S., Vossberg, K., Hopkins, R. S., & Mittan, B. (1998). Using PYLL in health planning. *Public health reports, 113*(1), 55.

Meslé, F., Vallin, J., & Andreyev, Z. (2002). Mortality in Europe: the divergence between east and west. *Population, 57*(1), 157-197.

Mujtaba, G., & Shahzad, S. J. H. (2021). Air pollutants, economic growth and public health: implications for sustainable development in OECD countries. *Environmental Science and Pollution Research*, *28*(10), 12686-12698.

Nixon, J., & Ulmann, P. (2006). The relationship between health care expenditure and health outcomes. *The European Journal of Health Economics*, *7*(1), 7-18.

OECD (2021). Topics: Health. Retrieved from: https://www.oecd.org/health/

Ogundimu, E. O., Altman, D. G., & Collins, G. S. (2016). Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of clinical epidemiology*, 76, 175-182.

Or, Z. (2000). Determinants of health outcomes in industrialised countries: a pooled, cross-country, time-series analysis. *OECD Economic studies*, 53-78.

Panopoulou, E. & Pantelidis, T. (2012). Convergence in per capita health expenditures and health outcomes in the OECD countries. *Applied Economics, 44*(30), 3909–3920.

Panopoulou, E., & Pantelidis, T. (2013). Cross-state disparities in us health care expenditures. *Health Economics, 22*(4), 451-465.

Park, M. B., & Nam, E. W. (2019). National Level Social Determinants of Health and Outcomes: Longitudinal Analysis of 27 Industrialized Countries. *SAGE Open, 9*(2), 2158244019854496.

Parmar, D., Stavropoulou, C., & Ioannidis, J. P. (2016). Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *BMJ (Clinical Research edition)*, *354*, i4588-i4588.

Phillips, P.C., & Sul, D. (2007). Transition modelling and econometric convergence tests. *Econometrica*, 75, 1771-1855.

Phillips, P.C., & Sul, D. (2009). Economic transition and growth. *Journal of Applied Econometrics, 24*, 1153-1185.

Potrafke, N. (2010). The growth of public health expenditures in OECD countries: do government ideology and electoral motives matter?. *Journal of health economics*, *29*(6), 797-810.

Poullier, J. P., Hernandez, P., Kawabata, K., & Savedoff, W. D. (2002). Patterns of global health expenditures: results for 191 countries. *Geneva: World Health Organization*.

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics,* 427-443.

Rajmil, L., & Fernández de Sanmamed, M. J. (2019). Austerity policies and mortality rates in European countries, 2011–2015. *American journal of public health*, *109*(5), 768-770.

Ravangard, R., Hatam, N., Teimourizad, A., & Jafari, A. (2014). Factors affecting the technical efficiency of health systems: A case study of Economic Cooperation Organization (ECO) countries (2004–10). *International journal of health policy and management, 3*(2), 63.

Reibling, N., Ariaans, M., & Wendt, C. (2019). Worlds of healthcare: a healthcare system typology of OECD countries. *Health Policy*, *123*(7), 611-620.

Retzlaff-Roberts, D., Chang, C. F., & Rubin, R. M. (2004). Technical efficiency in the use of health care resources: a comparison of OECD countries. *Health policy*, *69*(1), 55-72.

Richardson, E. A., Pearce, J., Mitchell, R., Shortt, N. K., & Tunstall, H. (2014). Have regional inequalities in life expectancy widened within the European Union between 1991 and 2008? *The European Journal of Public Health, 24*(3), 357-363.

Romeder, 1. J., & McWhinnie, J. R. (1977). Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning. *International journal of epidemiology, 6*(2), 143-151.

Samoli, E., Stergiopoulou, A., Santana, P., Rodopoulou, S., Mitsakou, C., Dimitroulopoulou, C., ... & Euro-Healthy Consortium. (2019). Spatial variability in air pollution exposure in relation to socioeconomic indicators in nine European metropolitan areas: A study on environmental inequality. *Environmental Pollution*, 249, 345-353.

Schooling, C.M., Lau, E.W.L., Tin, K.Y.K. & Leung, G.M. (2010). Social disparities and cause-specific mortality during economic development. *Social Science & Medicine*, *70*(10), 1550-1557.

Stańczyk, R. (2016). Convergence of health status in the European Union: a spatial econometric approach. *Athens Journal of Health, 3*, 95-112.

Stefko, R., Gavurova, B., Ivankova, V., & Rigelsky, M. (2020). Gender Inequalities in Health and Their Effect on the Economic Prosperity Represented by the GDP of Selected Developed Countries—Empirical Study. *International journal of environmental research and public health, 17*(10), 3555.

Torre, R., & Myrskylä, M. (2014). Income inequality and population health: an analysis of panel data for 21 developed countries, 1975–2006. *Population studies*, *68*(1), 1-13.

Vallin, J., & Meslé, F. (2004). Convergences and divergences in mortality: a new approach of health transition. *Demographic research, 2*, 11-44.

Van der Waerden, B. L. (1953). Order Tests for the Two-Sample Problem (second communication). *Indagationes Mathematicae (Proceedings), 56*, North-Holland.

Van Gool, K. & M. Pearson. (2014). Health, Austerity and Economic Crisis: Assessing the Short-term Impact in OECD countries. *OECD Health Working Papers, No. 76*, OECD Publishing.

Varbanova, V., & Beutels, P. (2020). Recent quantitative research on determinants of health in high income countries: A scoping review. *PloS one*, *15*(9), e0239031.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, 817-838.

Widding-Havneraas, T., & Pedersen, S. H. (2020). The role of welfare regimes in the relationship between childhood economic stress and adult health: a multilevel study of 20 European countries. *SSM-Population Health*, *12*, 100674.

Woolf, S. H., & Schoomaker, H. (2019). Life expectancy and mortality rates in the United States, 1959-2017. *Journal of the American Medical Association, 322*(20), 1996-2016.

World Bank. (2020). World Bank Dataset. Retrieved from: https://databank.worldbank.org/home

World Health Organization. (2021). Cardiovascular Diseases. Retrieved from: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

Xu, W., Engelman, M., & Fletcher, J. (2021). From convergence to divergence: Lifespan variation in US states, 1959–2017. *SSM-population health*, 16, 100987.

Zare, H., Gaskin, D. J., & Anderson, G. (2015). Variations in life expectancy in Organization for Economic Co-operation and Development countries–1985–2010. *Scandinavian journal of public health, 43(8),* 786-795.

Zhang, J. (2017). Low-level air pollution associated with death: policy and clinical implications. *Journal of the American Medical Association*, *318*(24), 2431-2432.

Table 1. PYLL descriptive analysis (1990-2017)

	Female					Male				
	min	**max**	**90-17**	**90-08**	**08-17**	**min**	**max**	**90-17**	**90-08**	**08-17**
Australia	2,608	4,516	-2.0%	-2.4%	-1.3%	4,415	8,367	-2.3%	-2.7%	-1.6%
Austria	2,606	4,829	-2.3%	-2.7%	-1.5%	4,736	9,948	-2.7%	-2.9%	-2.4%
Belgium	2,960	4,894	-1.8%	-1.9%	-1.8%	5,035	9,229	-2.2%	-2.0%	-2.7%
Brazil	5,913	9,334	-1.7%	-1.8%	-1.5%	11,920	17,377	-1.4%	-1.5%	-1.2%
Canada	3,177	4,621	-1.4%	-1.5%	-1.0%	5,002	8,640	-1.9%	-2.4%	-0.9%
Chile	3,537	7,051	-2.5%	-3.0%	-1.6%	6,273	13,438	-2.8%	-3.1%	-2.2%
Colombia	4,402	8,066	-2.1%	-2.2%	-2.0%	8,267	17,057	-2.4%	-2.4%	-2.3%
Costa Rica	3,772	5,832	-1.4%	-2.1%	0.1%	6,790	9,404	-0.9%	-1.4%	0.3%
Czech Republic	3,083	6,455	-2.7%	-3.0%	-2.1%	6,506	14,699	-2.9%	-3.2%	-2.5%
Denmark	2,856	5,921	-2.7%	-2.4%	-3.2%	4,590	9,710	-2.7%	-2.4%	-3.4%
Finland	2,558	4,668	-2.1%	-2.2%	-1.9%	5,311	11,300	-2.8%	-2.4%	-3.4%
France	2,669	4,259	-1.7%	-1.7%	-1.7%	5,364	9,982	-2.3%	-2.3%	-2.2%
Germany	2,900	5,023	-2.0%	-2.5%	-1.0%	5,110	9,937	-2.4%	-2.9%	-1.6%
Greece	2,577	4,238	-1.8%	-2.6%	-0.2%	5,507	8,005	-1.3%	-1.4%	-1.2%
Hungary	4,517	8,707	-2.3%	-2.4%	-2.3%	9,547	20,292	-2.6%	-2.2%	-3.4%
Iceland	2,235	4,494	-2.5%	-3.0%	-1.4%	3,557	7,628	-2.2%	-3.4%	0.3%
Ireland	2,614	5,392	-2.6%	-2.5%	-3.0%	4,149	9,173	-2.9%	-2.5%	-3.6%
Israel	2,403	5,170	-2.8%	-3.2%	-1.9%	4,108	7,675	-2.2%	-2.1%	-2.5%
Italy	2,276	4,200	-2.2%	-2.5%	-1.5%	3,971	8,720	-2.8%	-3.0%	-2.4%
Japan	2,069	3,442	-1.9%	-1.7%	-2.2%	3,924	6,670	-1.9%	-1.6%	-2.6%
Korea	2,071	5,643	-3.6%	-3.4%	-3.9%	4,511	13,140	-3.8%	-3.7%	-3.8%
Luxembourg	1,944	5,362	-3.3%	-4.2%	-1.5%	3,562	10,725	-3.8%	-4.1%	-3.1%
Mexico	6,047	9,618	-1.7%	-2.1%	-0.8%	10,564	15,743	-1.1%	-1.9%	0.6%
Netherlands	2,857	4,413	-1.6%	-1.5%	-1.8%	4,073	7,779	-2.4%	-2.7%	-1.8%
Norway	2,369	4,367	-2.2%	-2.1%	-2.6%	3,747	8,529	-3.0%	-2.8%	-3.4%
Poland	3,729	7,228	-2.4%	-2.5%	-2.2%	9,197	17,599	-2.3%	-2.0%	-2.7%
Portugal	2,604	5,734	-2.9%	-3.2%	-2.2%	5,685	12,370	-2.7%	-2.8%	-2.5%
Slovenia	2,558	5,633	-2.5%	-2.8%	-1.8%	5,172	13,052	-3.2%	-2.9%	-3.7%
Spain	2,198	4,189	-2.3%	-2.6%	-1.8%	4,314	9,432	-2.9%	-2.7%	-3.1%
Sweden	2,465	4,148	-1.9%	-2.0%	-1.6%	3,931	7,322	-2.3%	-2.6%	-1.7%
Switzerland	2,138	4,200	-2.5%	-2.4%	-2.5%	3,614	8,544	-3.0%	-3.3%	-2.4%
U. Kingdom	3,190	5,208	-1.8%	-1.9%	-1.6%	5,027	8,760	-2.0%	-2.1%	-1.7%
United States	4,648	5,954	-0.7%	-1.1%	0.1%	7,758	11,271	-1.1%	-1.7%	0.1%

This table presents some descriptive statistics of the PYLL variable. Min and Max are the sample minimum and maximum values, respectively, whilst the columns 90-17, 90-08 and 08-17 show the average growth rate for these periods, respectively.

Table 2: Testing for convergence for PYLL

	Female		Male	
Panel I: Testing for convergence	$\hat{\beta}$	Log t-ratio	$\hat{\beta}$	Log t-ratio
	-10.442	-829.828*	-10.421	-876.720*
Panel II: Initial estimated Clubs	$\hat{\beta}$	Log t-ratio	$\hat{\beta}$	Log t-ratio
Club 1:	0.083	1.980	0.127	0.577
	Brazil. Mexico. United States		Brazil. Hungary. Mexico	
Club 2:	0.135	2.213	0.074	1.190
	Canada. Chile. Colombia. Costa Rica. Hungary. Poland. United Kingdom		Chile. Colombia. Costa Rica. Czech Republic. Greece. Poland. United States	
Club 3:	0.676	6.385	0.415	5.033
	Australia. Belgium. Czech Republic. Denmark. France. Germany. Greece. Netherlands		Australia. Austria. Belgium. Canada. Finland. France. Germany. Japan. Portugal. Slovenia. United Kingdom	
Club 4:	0.164	5.456	0.644	6.876
	Austria. Finland. Iceland. Ireland. Israel. Italy. Korea. Luxembourg. Japan. Norway. Portugal. Slovenia. Spain. Sweden. Switzerland		Denmark. Iceland. Ireland. Israel. Italy. Korea. Luxembourg. Netherlands. Norway. Spain. Sweden. Switzerland	
Panel III: Club merging analysis	$\hat{\beta}$	Log t-ratio	$\hat{\beta}$	Log t-ratio
Club 1+2	-0.685	-122.655*	-0.504	-28.262*
Club 2+3	-0.296	-7.839*	-0.607	-50.400*
Club 3+4	-0.482	-11.688*	-0.276	-7.505*

The estimated β and the Log t-ratio are obtained from the estimation of equation (3). '*' Denotes statistical significance at the 5% level, rejecting the null hypothesis of convergence.

Table 3. Average values of factors driving Club formation.

	Female				Male			
	Club 1	**Club 2**	**Club 3**	**Club 4**	**Club 1**	**Club 2**	**Club 3**	**Club 4**
GDP per capita	31,424	30,198	46,933	53,558	21,282	30,683	45,890	58,423
Gini Index	44.50	39.59	30.23	31.53	43.10	38.03	31.61	31.85
Human capital Index	0.64	0.70	0.77	0.78	0.62	0.69	0.79	0.77
Unemployment	6.86	6.24	7.60	6.62	6.79	8.22	6.26	6.46
Immunization measles	86.33	93.86	95.25	95.60	88.67	94.57	94.36	95.50
Immunization polio	84.66	93.57	95.87	95.00	86.33	94.00	94.09	95.75
Mortality CVD, cancer, or diabetes	15.63	14.59	11.64	9.79	18.43	14.34	10.70	9.63
Maternal mortality	37.33	22.00	5.12	4.86	35.00	21.42	6.45	4.50
Diabetes prevalence	11.56	7.08	6.35	6.12	10.26	7.67	6.40	6.00
Prevalence anemia	20.60	21.34	14.58	13.71	26.63	21.40	13.80	12.67
Prevalence overweight	63.10	61.62	59.76	54.34	61.00	62.07	56.46	56.27
Current Health Expenditure	10.68	8.17	9.70	9.01	7.29	8.92	10.00	8.98
Control corruption	-0.26	0.81	1.36	1.38	-0.45	0.52	1.55	1.48
Government effectiveness	0.41	0.77	1.34	1.45	0.07	0.64	1.50	1.49
Voice accountability	0.47	0.88	1.25	1.24	0.24	0.82	1.29	1.30
Greenhouse gas emissions	9.48	7.87	11.19	10.99	5.65	8.64	10.69	11.91
Carbon dioxide damage	1.01	1.04	0.79	0.58	1.11	1.16	0.70	0.52

This table presents the average values of the considered explanatory variables of each one of the estimated clubs. These average values are based on the use of the last observable value of the explanatory variables.

Table 4. Ordered logit estimation.

	Female			Male		
	Coefficient	**t-Ratio**		**Coefficient**	**t-Ratio**	
Gini Index	-0.48	-4.20	***	-0.21	-2.01	**
Current Health Expenditure	-1.28	-3.73	***			
GDP per capita				0.0001252	2.79	***
Immunization measles	0.57	3.67	***	0.3061579	1.96	**
Mortality from CVD, cancer or diabetes	-1.87	-3.41	***	-1.65	-3.77	***
Carbon dioxide damage				-2.73	-1.89	*
Pseudo R^2	0.68			0.68		
Correctly predicted cases	76%			79%		

This table shows the results of the ordered logit estimation of equation (4). The dependent variable (y_i) takes the value m when the i-th country is included in the m-th estimated club, with m=1, 2, 3, 4. The t-ratios for testing the single significance of the parameters are robust to the presence of heteroskedasticity by using the Huber-White method, see White (1980) in this regard.

'*' Denotes statistical significance at the 10% level. '**' at the 5% level. and '***' at the 1% level.

Figure 1: PYLL coefficient of variation

Figure 1a. Female

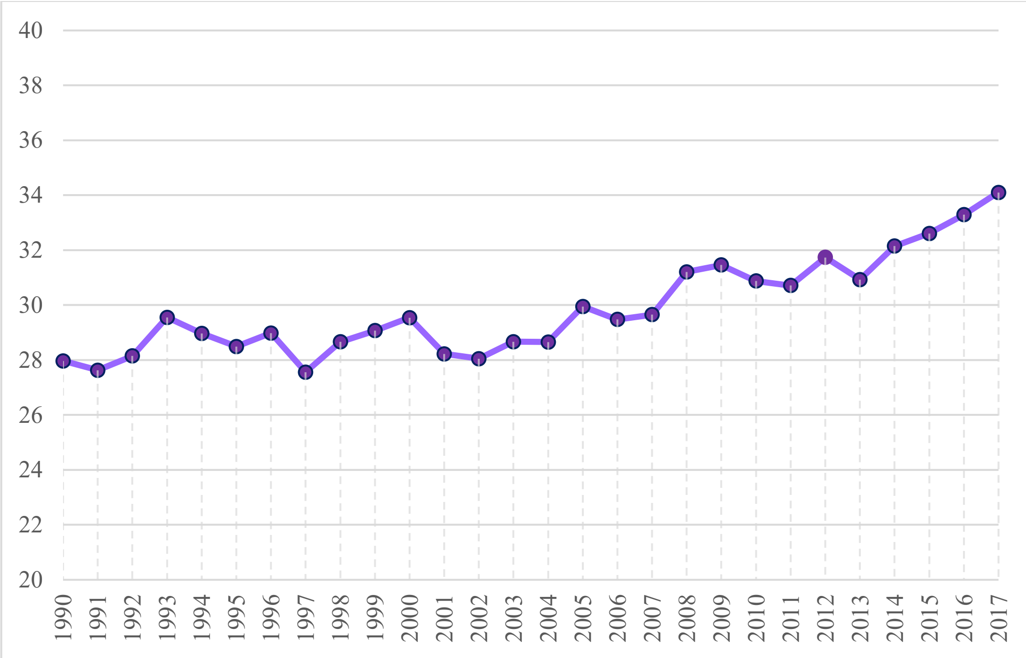
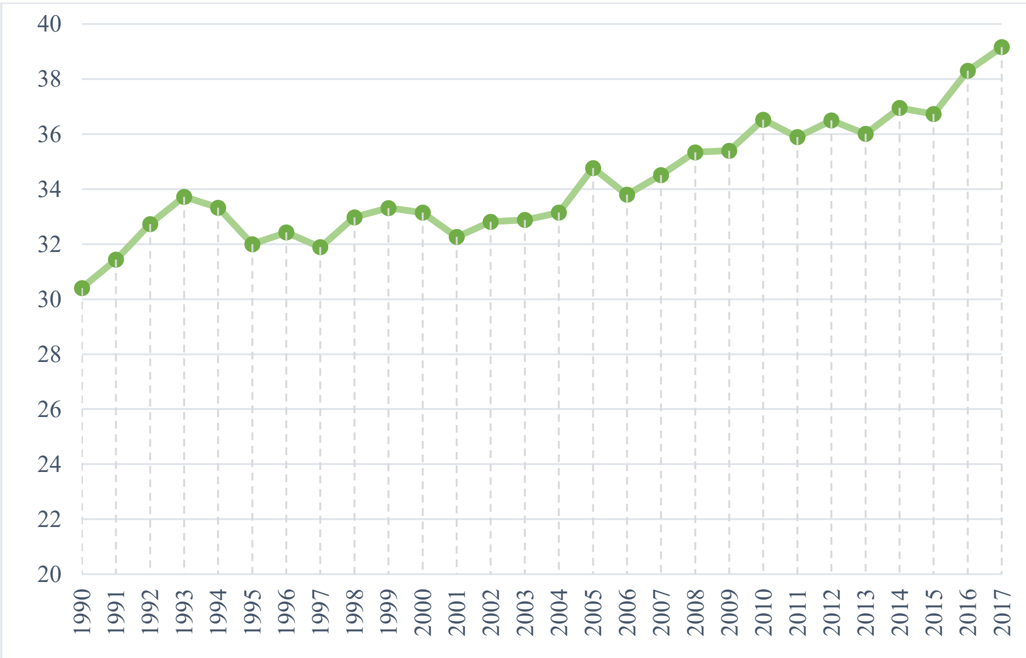


Figure 1b. Male



These figures present the evolution of the coefficient of variation of the female and male PYLL of the countries included in the sample for 1990-2017.

Figure 2a. Geographical location of countries included in the estimated clubs: Female case

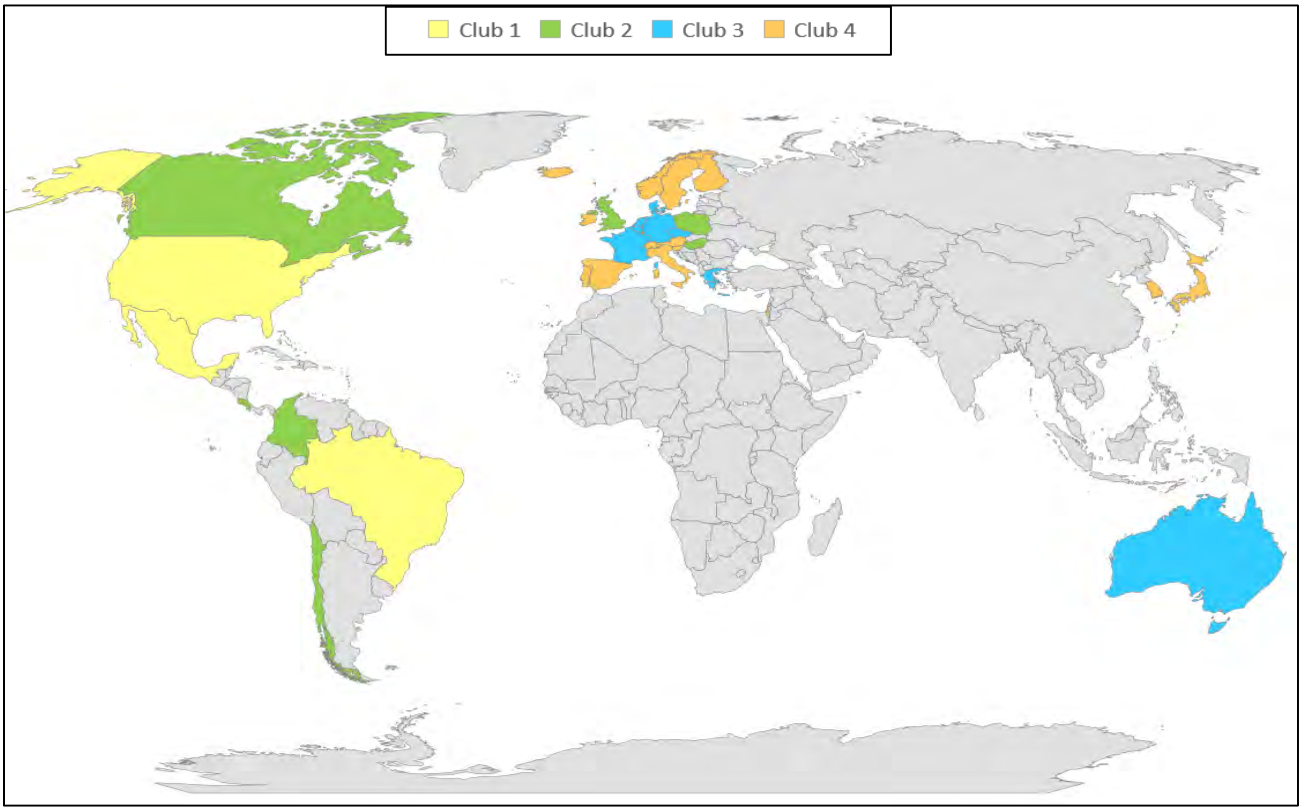


Figure 2b. Geographical location of countries included in the estimated clubs: Male case

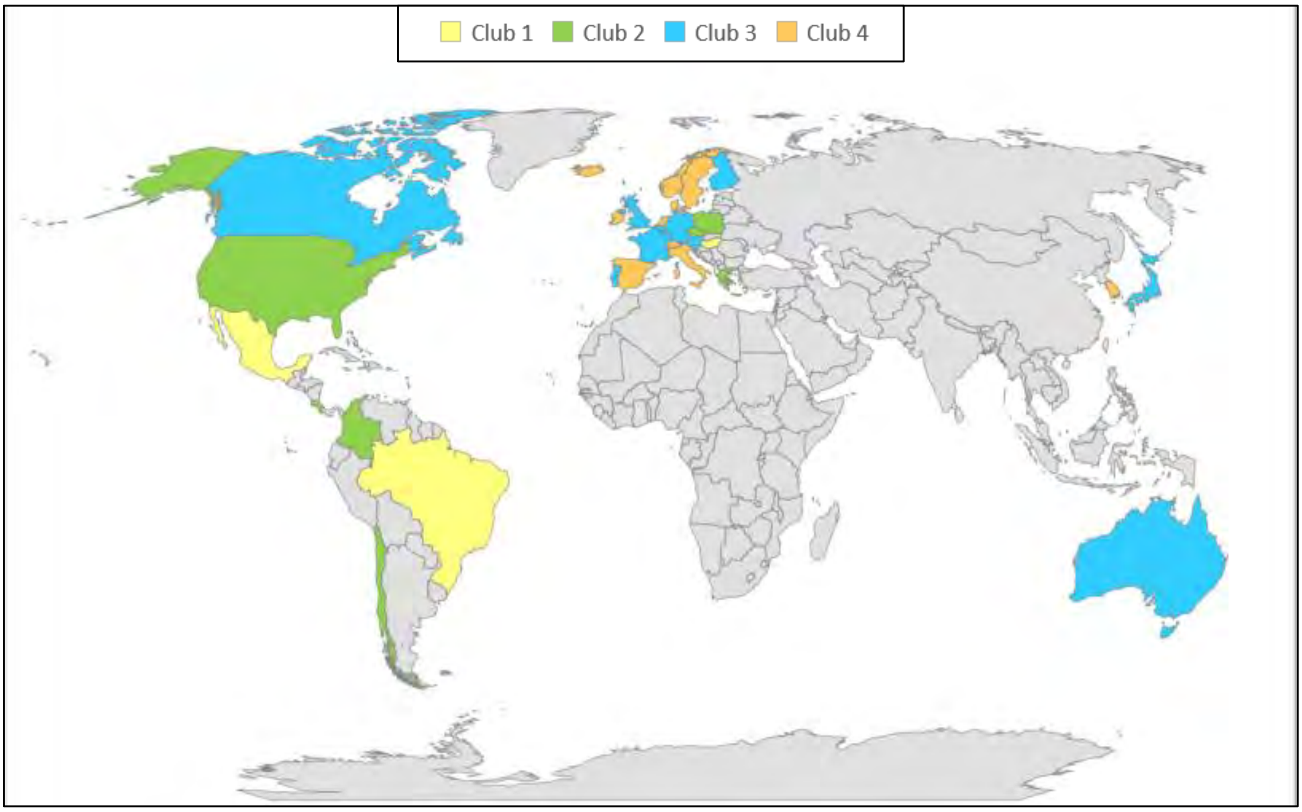


Figure 3: Average values of PYLL by estimated Clubs

Figure 3a. Female PYLL evolution

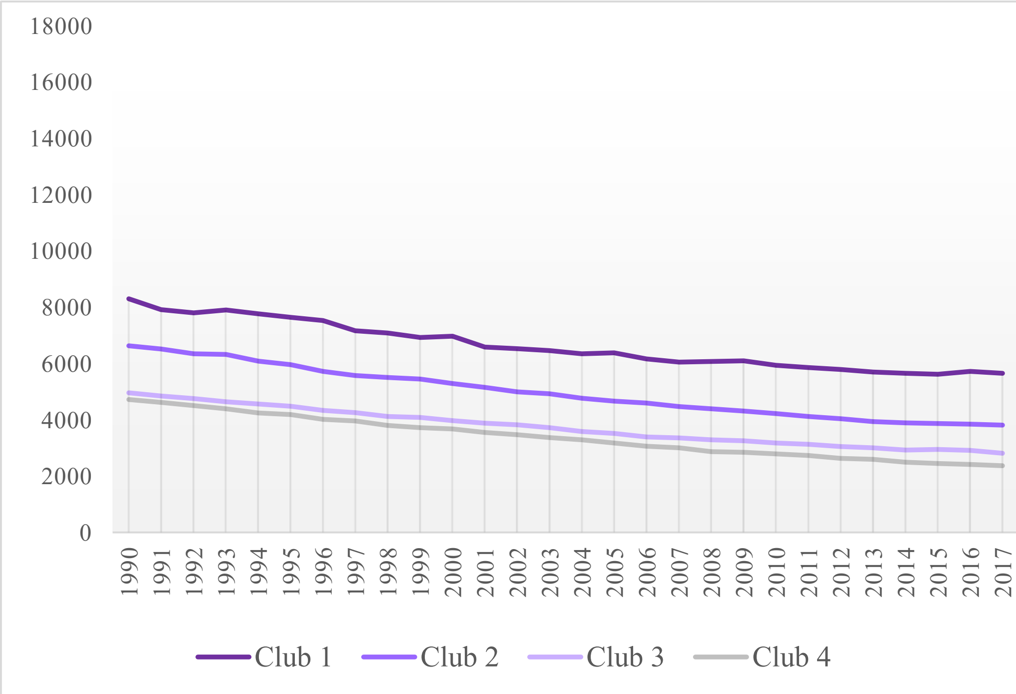
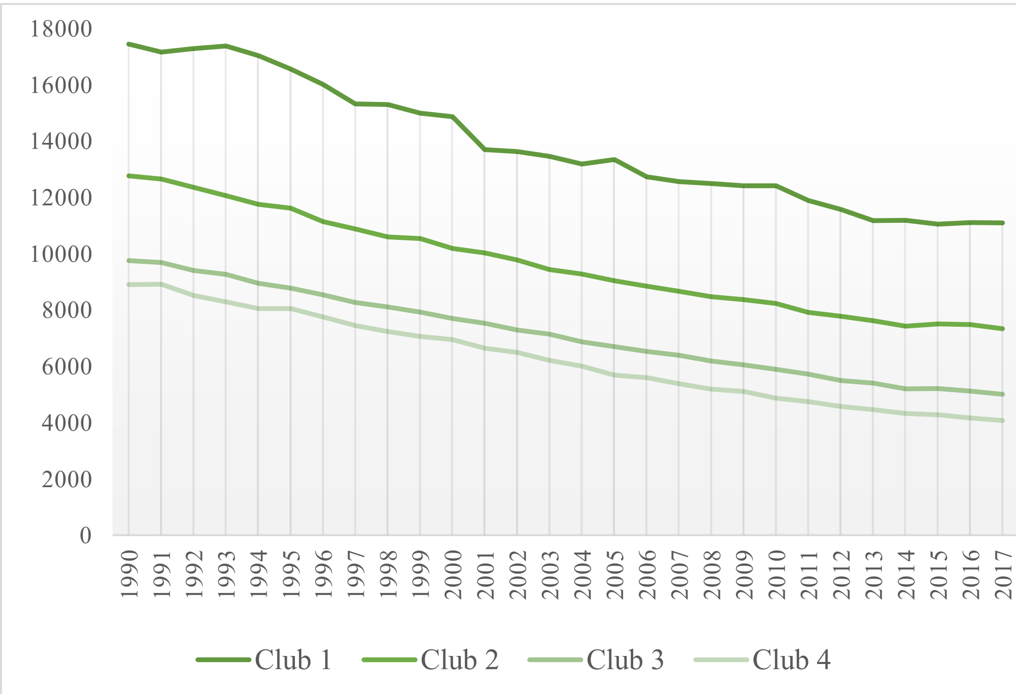
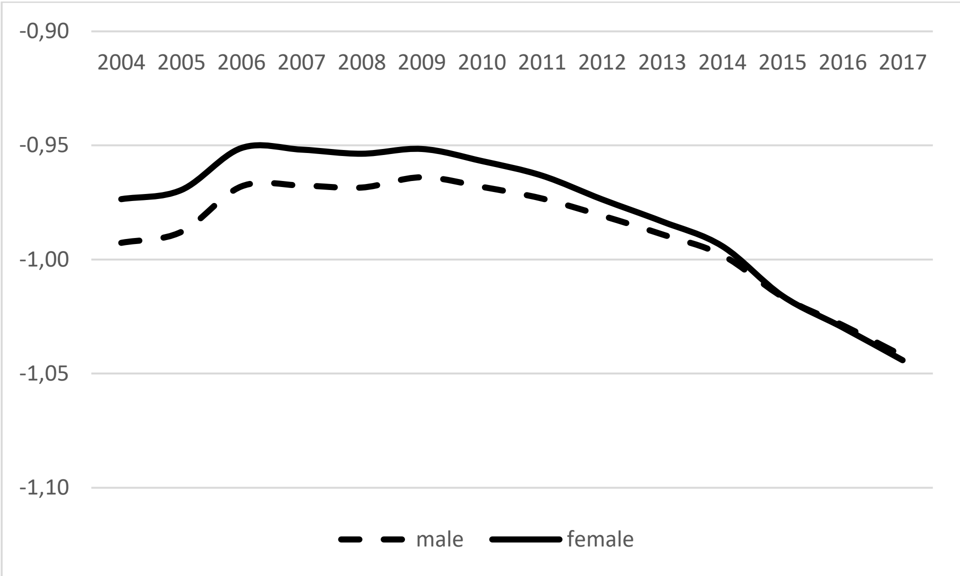


Figure 3b. Male PYLL evolution



These figures reflect the average values of the PYLL of different estimated clubs.

Figure 4. Log t-ratio evolution



This figure presents the values of the t-ratio for testing the null hypothesis $\beta = 0$ in (3) when this model is recursively estimated for the sample t=1990,…, k, where k=2004, 2005,…, 2017.

Appendix I

Definition of variables

Potential years of life lost is an indicator that summarizes premature mortality, providing an explicit way of weighting deaths occurring at younger ages, which may be preventable. The calculation of Potential Years of Life Lost (PYLL) involves summing up deaths occurring at each age and multiplying this with the number of remaining years to live up to a selected age limit (age 70 is used in OECD Health Statistics). In order to assure cross-country and trend comparison, the PYLL is standardized for each country and each year. The total OECD population in 2010 is taken as the reference population for age standardization. This indicator is presented as a total and per gender. It is measured in years lost per 100, 000 inhabitants (total), per 100, 000 men and per 100, 000 women, aged 0-69. Source: OECD Data

GDP per capita (constant 2017 international $), based on purchasing power parity, is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the country plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2017 international dollars. Source: World Data Bank

Gini Index is the index of inequality among the population of a country or region. The value of the Gini index is between 0 and 1, with zero being the maximum equality (all citizens have the same income) and 1 being the maximum inequality (all income is held by a single citizen). Source: World Data Bank

The Human Capital Index calculates the contributions of health and education to worker productivity. The final index score ranges from zero to one and measures the productivity as a future worker of a child born today relative to the benchmark of full health and complete education. Source: World Data Bank

Unemployment refers to the share of the labor force that is without work but available for and seeking employment. The series is part of the ILO estimates and is harmonized to ensure comparability across countries and over time by accounting for differences in data source, scope of coverage, methodology, and other country-specific factors. The estimates are based mainly on nationally representative labor force surveys, with other sources (population censuses and nationally reported estimates) used only when no survey data are available. Source: World Data Bank

Immunization measles represents child measles immunization and measures the percentage of children aged 12-23 months who received the measles vaccination before 12 months or at any time before the survey. A child is considered adequately immunized against measles after receiving one dose of vaccine. Source: World Data Bank

Immunization polio is the rate of child immunization against polio measured as the percentage of children aged 12-23 months who received polio vaccinations before 12 months or at any time before the survey. A child is considered adequately immunized after three doses. Source: World Data Bank

Mortality CVD, cancer, or diabetes is the percentage of 30-year-olds who would die before age 70 from any cardiovascular disease, cancer, diabetes, or chronic respiratory disease, assuming he/she would experience current mortality rates at each age and would not die from any other cause of death (e.g., injury or HIV/AIDS). Source: World Data Bank

Maternal mortality is the number of women who die from pregnancy-related causes while pregnant or within 42 days of pregnancy termination per 100,000 live births. The data are estimated with a regression model using information on the proportion of maternal deaths among non-AIDS deaths in women aged 15-49, fertility, birth attendants, and GDP measured using purchasing power parities (PPPs). Source: World Data Bank

Diabetes prevalence refers to the percentage of people aged 20-79 who have type 1 or type 2 diabetes. It is calculated by adjusting to a standard population age-structure. Source: World Data Bank

Prevalence anemia among children is the proportion of children under age 5 whose hemoglobin level is less than 110 grams per liter, adjusted for altitude. Source: World Data Bank

Prevalence overweight is the percentage of adults aged 18 and over whose Body Mass Index (BMI) is more than 25 kg/m2. Body Mass Index (BMI) is a simple index of weight-for-height, or the weight in kilograms divided by the square of the height in meters. Source: World Data Bank

Current Health Expenditure per capita is current health expenditure converted to international dollars using purchasing power parity (PPP) rates. Level of current health expenditure expressed as a percentage of GDP. Estimates of current health expenditures include healthcare goods and services consumed during each year. This indicator does not include capital health expenditures such as buildings, machinery, IT, and stocks of vaccines for emergency or outbreaks. Source: World Data Bank

Control of corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests. The estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5. Source: World Data Bank

Government effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. The estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5. Source: World Data Bank

Voice accountability captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media. The estimate gives the country's score on the aggregate indicator in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5. Source: World Data Bank.

Greenhouse gases emissions represents trends in man-made emissions of major greenhouse gases and emissions by gas. Data refer to total emissions of CO_2 (emissions from energy use and industrial processes, e.g. cement production), CH_4 (methane emissions from solid waste, livestock, mining of hard coal and lignite, rice paddies, agriculture and leaks from natural gas pipelines), nitrous oxide (N_2O), hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), sulphur hexafluoride (SF_6) and nitrogen trifluoride (NF_3). The variable is expressed in per capita terms. Source: OECD Data

Carbon dioxide damage measures the cost of damage due to carbon dioxide emissions from fossil fuel use and cement manufacturing. It is estimated at $30 per ton of CO2 (the unit damage in 2014 US dollars for CO2 emitted in 2015) multiplied by the number of tons of CO2 emitted. Source: World Data Bank