Healthy life expectancy: Is there convergence in Europe?

Antonio Marino

N°152218034

Dissertation submitted in partial fulfillment of the requirements for the degree of MSc in Macroeconomic Policy at Católica-Lisbon School of Business & Economics and University of Naples Federico II.

Thesis written under the supervision of Miguel Gouveia

2021

Abstract

To advance our knowledge on convergence in global health, this study employed data from 31 European countries to test the convergence hypothesis using Healthy life expectancy (HLE) data from 2004 to 2016.

Convergence was examined through different perspectives in the literature. First, we studied the absolute and relative disparities of HLE across countries using the DMM (Dispersion measure of mortality) and the Gini coefficient. Second, we tested the hypothesis of absolute convergence analyzing the catching-up process across the European countries by β-convergence together with their degree of dispersion (measured by standard deviation).

The female population registered a decreasing trend of DMM and Gini coefficient in contrast with the male population that met in the last year an increase in absolute and relative inequality. Although this positive result (speaking only for females), and the results of β-convergence that show a positive speed of convergence equal to 0.0201 for the females and 0.0163 for the males, the differences in standard deviations estimated do not show statistical significance in accepting the hypothesis of absolute convergence.

In the third place, we looked at the how HLE evolved in Europe. We estimated the transition probabilities from a certain HLE value to another, and on this basis, we constructed a long-run distribution of HLE.

Looking at the long-run distributions of HLE, estimated using the method proposed by Quah(1996), we observed a bimodal distribution for both genders that confirm the hypothesis of club convergence submitted to the probabilities estimated.

Keywords: Healthy life expectancy based on self-perceived health(HLE), Healthy inequality, Dispersion measure of mortality, Gini coefficient, β - convergence, σ-convergence, kernel density estimation, absolute convergence, club convergence.

Resumo

Para avançar nosso conhecimento sobre convergência em saúde global, este estudo empregou dados de 31 países europeus para testar a hipótese de convergência do uso de dados de expectativa de vida em saúde (HLE) de 2004 a 2016.

A convergência foi examinada por meio de diferentes perspetivas na literatura. Primeiro, estudamos as disparidades absolutas e relativas de HLE entre os países usando o DMM e o coeficiente de Gini. Em segundo lugar, testamos a hipótese de convergência absoluta analisando o processo de catching-up entre os países europeus por β-convergência juntamente com seu grau de dispersão (medido pelo desvio padrão).

A população feminina apresentou tendência de diminuição do DMM e do coeficiente de Gini em oposição à população masculina que registou no último ano um aumento da desigualdade absoluta e relativa. Embora este resultado positivo (falando apenas para mulheres), e os resultados de convergência β que mostram uma velocidade positiva de convergência igual a 0,0201 para as mulheres e 0,0163 para os homens, as diferenças nos desvios padrão estimados não mostram um significância estatística para aceitar a hipótese de convergência absoluta.

Em terceiro lugar, vimos como o HLE evoluiu na Europa. Estimamos as probabilidades de transição de um determinado valor de HLE para outro e, com base nisso, construímos uma distribuição de longo prazo de HLE.

Observando as distribuições de longo prazo do HLE, estimadas pelo método proposto por Quah (1996), observamos uma distribuição bimodal para ambos os sexos que confirma a hipótese de convergência do clube submetida às probabilidades estimadas.

Palavras-chave: Expectativa de vida saudável com base na autopercepção de saúde (HLE), Desigualdade saudável, Medida de dispersão da mortalidade, coeficiente de Gini, β-convergência, σ-convergência, estimativa de densidade de kernel, convergência absoluta, convergência de clube.

Acknowledgements

First off I would like to say “thanks” to Prof. Miguel Gouveia, my supervisor, that grants me the opportunity to work in this field of economics and for his support in the entire process.

Importantly, I would like to thank the University of Naples Federico II and all other institutions that gave me the possibility to be involved in this experience. I would like to thank Católica-Lisbon and all its faculty to let me pass with serenity the last year of my academic career. A special appreciation goes to my Academic Directors, Professor Catarina Reis and Marco Pagnozzi, for their support during the years of my master’s. Then, I want to give all my gratitude to all my master teachers, who lectured me with especially interesting topics and presented me with most of the tools used in this dissertation.

I would also thank Prof. Dal Bianco Silvia to make available the commands on MATLAB used for the estimation of the stochastic and ergodic kernel distribution.

Lastly, my family and, my parents gave me the opportunity to pursue academic studies and to continue them in a foreign country. Without them, all this would not have happened.

Table of Contents

List of figures

List of tables

List of abbreviations

AMISE – Asymptotical mean integrated squared error

DMM – Dispersion measure of mortality

HLE – Healthy life expectancy based on self-perceived health

LEB – Life Expectancy at birth

MISE – Mean integrated squared error

NLS – Non-linear least squares

1. Introduction

In the European panorama, the population lives longer and healthier than ever before in human history. Modern knowledge and technologies allow for the healing of diseases that were considered hard to treat successfully and prevent in the past. Instead, the education, the socio-economical, and legal conditions changed the people's behavior guaranteeing a healthier lifestyle. All that lead to an increase in the longevity and in the wellbeing of the European population.

The Healthy life expectancy based on self-perceived health describes how many years a person is expected to live in good perceived health at a certain age. It differs from Life Expectancy because that represents the mean number of years still to be lived by a person who has reached a certain exact age if subjected throughout the rest of his or her life to the current mortality conditions based on age-specific probabilities of dying and other demographic condition like gender. Anyone can deny the utility of the LEB, but the answer to what is the expected lifetime of a person is not enough knowledge to make decisions related to the job, pension, or health care system. In few words, if people live longer than before or longer with respect other countries but their health status is compromised they can't extend their waiting period for retirement. For this, we decided to use HLE. To take care of the living duration of people and the direction of the health trend in the European countries.

From 2004 to 2016 HLE based on self-perceived health has progressed, not only at the time of birth but also across all age groups and different countries. In Europe the male population has registered an HLE based on self-perceived health at birth of 73.1 years, encountering a growth of 2.38% from 2010 to 2016. The women has experienced, instead, an increase of 1.87% from 2010 to 2016. The annual assessment of the European health is a key component of the public policy process to improve health levels and to diminish inequalities. Then, it is essential to learn how the health index is converging or diverging so as address efficiently the resources in a public policy process that has also the aim to provide the well-being of people.

The term convergence can mean many different things (see Quah,1993)[15]:

a) Countries originally better than average are more likely to turn below average eventually, and vice-versa - the cycle repeats;
b) Whether a country HLE is eventually above or below average is independent of that country's original position;

c) HLE disparities between countries have neither unit-roots nor deterministic time trends; and
d) Each country eventually gets to an HLE as high as all the others - the cross-section dispersion diminishes over time.

Cases (a) and (b) vaguely correspond to the notions of mixing and ergodicity in econometrics; see e.g. White (1984)[23]. Case (c) is one formulation of persistence in HLE disparities. Case (d) is closest to the notion of countries with a lower HLE eventually catching up with countries with a higher level of HLE. This dissertation aims to examine all the definitions of convergence verifying the conditions of their occurrence. In cases (a) and (b) we will study the transition dynamics of the distribution of HLE, estimated using a kernel density distribution. This approach, adopted by Quah,(1996)[16] for his research means to examine the evolution of distribution experienced in a certain period. Our study intends to estimate a distribution that describes the transitions from an HLE status to another registered in Europe for the considered period. The distribution will be estimated by multivariate kernel density estimation using the common procedure in the literature, Wendy L. Martinez Angel R. Martinez (2002)[12]. Finally, we are going to project such distribution into the future to observe the long-run distribution of HLE. The HLE disparities across countries mentioned in point (c) will be scrutinized looking carefully at the trend of two measures of inequality: the DMM(Dispersion measure of mortality) and the Gini coefficient that are commonly used to investigate convergence in absolute and relative dispersion, Srinivas Goli and Perianayagam Arokiasamy (2013)[5]. In the end, as Barro and Sala-i-Martin(1992)[1] did in their study we adopted the same regression-based approach for our research. We used the **β-convergence** to look at the catching-up process between short-lived populations and long-lived populations and then we look at the trend of HLE standard deviations of both genders to make a hypothesis on a manifestation of absolute convergence.

2. Data analyses

HLE based on self-perceived health describes how many years a person is expected to live in good perceived health at a certain age. We have conducted our research to discover what is the current trend in Europe for the years a new birth will live in good health. For this reason, we used HLE[1] at birth.

This study used data from Eurostat that involves the Sullivan method[2] to calculate the HLE at birth of both genders for each year. The method is based on two main parts:

- A life table that enables calculation of the life expectancy for each age x_i;
- The observed prevalence of the population in a healthy or unhealthy status;

And for its calculations it considers two simple assumptions:

- the maximum age class is set at 85+ for all countries, both sexes and all years;
- deaths are assumed to occur halfway through the year except for age 0;

The death rate (m_{xi}) corresponding to each x_i is calculated by dividing the number of deaths by the average population.

It helps us to calculate the probability of dying at age x_i (q_{xi}):

$$q_{xi} = m_{xi} / [1+(1-a_{xi})m_{xi}]$$

where m_{xi} is the death rate and a_{xi} is the fraction of the year that a person has lived in addition to the x_i complete years (with the second decision above, $a_{xi} = 0.5$ but for age 0 the coefficient $a_{xi} = 0.2$).

With this probability q_{xi} we calculate the number l_{xi+1} of people still alive at age x_{i+1}.

$$l_{xi+1} = (1-q_{xi}) * l_{xi}$$

For the first age, l_{xi} is fixed at 1.

---

[1] The terms HLE and HLE based on self-perceived health are used indifferently for convenience.

[2] The procedure of Sullivan's method is taken from the website of Eurostat : https://ec.europa.eu/eurostat/cache/metadata/Annexes/hlth_silc_17_esms_an1.pdf

Based on the assumption that people dying at age x_i will live in average half of the year between age xi and age x_{i+1} except for age 0, it is possible to determine the total number of years L_{xi} that survivors of age x_i will live between age xi and x_{i+1}.

$L_{xi} = (l_{xi+1} + l_{xi}) / 2$

The total number of years that survivors of age 0 will live between 0 and the first year of birth is:

$L_{x0} = 0.2*l_{x0} + 0.8 * l_{x1}$

For the last stage ω, since the death rate (m_{xi}) corresponding to each x_i is calculated by dividing the number of deaths by the average population (thus $m_{xi}=(l_{xi}-l_{xi+1})/L_{xi}$), the total number of years correspondent to survivors can be derived.

$L_\omega = l_\omega/m_\omega$ ($\omega = 85$)

Then, the formula for the total number of years that the survivors of age x_i will still live before dying is as follows:

$Tx_i = \sum^{\omega}_{xi} L_{xi}$

By dividing this number of years by the number of survivors at age x_i, we obtain the life expectancy e_i for a person still alive at age x_i.

$e_{xi} = T_{xi} / l_{xi}$

Now, you need to take into consideration the prevalence of the population in healthy or unhealthy status. The prevalence is used to divide the hypothetical years of life lived by people at different ages into years with and without good perceived health. It consists of applying the prevalence π_i (j) of each of the different states of health j[3] in an age category i on the years lived Li in this age interval. Indeed, by multiplying the proportion π_i (j) of being in a state of health j

[3] The health states categories j are two. One composed by "bad" and "very bad", the other one by "Very good", "Good" and "Fair".

by L_i, it is obtained the number L_i (j) of person-years spent by survivors of age x_i in health state j in the age interval.

$L_i (j) = L_i *(\pi_i (j))$

For each age category, by summing all the L_k (j) from this interval to the eldest one, it results the total number T_i (j) of person-years still to spend in a state of health j for persons alive at age x_i.

$T_i (j) = \Sigma L_k (j)$ where $k \geq i$

Finally, the life expectancy in state j at age x_i is obtained by dividing the total number of person-years spent in state j by the number of survivors at age x_i:

$e_i (j) = T_i (j) / l_{xi}$

To define the proportion of the population in good health (prevalence data) the statistics rely on strongly harmonized and widely accepted health concepts and comparable data.
Then, the proportions π_i (j) is obtained by a survey that asks a representative sample of the population how they assess their health in general by using several health states:

- Very good
- Good
- Fair
- Bad
- Very bad

Finally, the proportion $\pi_i(j)$[4] of people with less than good health is determined by age categories of 5 years. Before proceeding to calculations some assumptions are necessary. The prevalence for the first age group, 16-19, has been applied to the 15-19 population group. Another assumption has been to consider that prevalence of people before the age of 15 years is half of the prevalence of the next age interval (16-19 years).

[4] "Less than good perceived health" is based on the answer category 4 (bad) and 5 (very bad).

The interest of Sullivan's method lies in its simplicity, the availability of its basic data, and its independence of the size and age structure of the population.

The full availability of the data for the entire period is compromised since there is some missing data in several years. Data from 2004 onwards is available for Belgium, Denmark, Estonia, Ireland, Greece, Spain, France, Italy, Luxembourg, Malta, Austria, Portugal, Finland, Sweden and Iceland; Czechia, Germany, Cyprus, Lithuania, Hungary, Malta, Netherland, Poland, Slovenia, Slovakia, The United Kingdom miss data for 2004; Norway and Iceland miss data for 2006 and 2016 respectively and finally for the remaining countries we don't have data for several years. For Bulgaria we don't have data for the years 2004 and 2005 - Romania – 2004, 2005 and 2006 - Croatia from 2004 to 2009 and Switzerland from 2004 to 2007 and 2015. The HLE of both genders increased for countries with a long-lived population as for countries with a short-lived population. We can notice it from the minimum and maximum values along years in Table 1.

We can also notice that the females' HLE reaches larger values than males. It means that the steady-state or steady states in the case of club convergence will have a higher value. The time series of the males and females' HLE over the period 2004-2016 are depicted in Figure 1. (a) and (b) and Figure 2. (a) and (b). We can imagine in the correspondent graphs to both genders two trend lines different in steepness that let emerge two groups of countries. In Figure 1. It appears clearly that the countries which have the HLE equal to 65.8 or below in 2005 have a trend line steeper than those countries that had superior HLE levels initially.

Table 1.

		Mean		Standard deviation		MIN		MAX	
Year	N° obs.	Males	Females	Males	Females	Males	Females	Males	Females
2004	15	70,23	73,38	3,97	3,75	59,8	63,2	75,1	77,8
2005	27	68,40	71,63	5,95	5,54	56,6	61,6	76,2	78,6
2006	27	68,54	71,75	5,91	5,56	57,7	62,7	76,2	79,4
2007	28	68,84	72,30	5,80	5,40	58,3	62,1	76,6	79,6
2008	29	69,85	73,44	5,78	5,19	59,2	64,6	77,8	82,1
2009	29	70,10	73,86	5,29	4,97	60,1	65,5	77,9	81,7
2010	31	70,27	73,80	5,43	5,08	60,5	65	78,0	81,8
2011	31	70,55	73,87	5,26	4,92	59,9	64,9	78,2	82,1
2012	31	70,89	74,29	5,29	4,78	60,3	65,3	78,6	82,3
2013	31	71,26	74,46	5,13	4,47	61,3	66,5	78,3	81,9
2014	31	71,52	74,95	5,20	4,53	61,6	66,6	78,8	81,9
2015	30	71,52	74,78	4,94	4,30	62,8	67	78,4	80,6
2016	30	72,07	75,48	4,88	4,07	62,2	67,7	79,0	81,9

This shows a greater average growth rate for countries that had lower HLE initial values[5]. The same happens for the time series of HLE for the female population. Indeed, the European countries that have the HLE under 70 years have a higher average growth rate. This presents good conditions for the occurrence of convergence in average, and thus for the examination of **β-convergence**.

Moreover, we found a negative trend of the standard deviation of HLE for both genders. Precisely the standard deviation related to the male population decreased by 17.98%, from 5.95 in 2005 to 4.88 in 2016. Meanwhile, the standard deviation of females' HLE dropped from 5.54 in 2005 to 4.07 in 2016, a reduction of 26.53%[6]. Based on the common convergence literature, Barro Robert J. and Xavier Sala-i-Martin (1992) and Xavier Sala-i-Martin(1996)[1][18] the catching-up process between short-lived and long-lived European populations and the decreasing standard deviation are good requirements for the occurrence of absolute convergence. Naturally, these conditions must be verified properly. The next section provides to explain the mathematical tools as the regression-based approach of Barro and Sala-i-Martin(1992) used in this research to uncover the future evolution of HLE levels. However, we cannot exclude the possibility that the HLE of European countries can diverge or can converge to different steady states in the future. For this reason, according to the results that emerged from the study by Quah(1996), we used his approach based on distributions estimated by the kernel functions. Looking at the long-run distribution of the HLE of both groups we could evaluate if in Europe we are going to face the case of convergence, polarization, or stratification with the number of modes depicted by the distribution.

[5] Excluding the case of Spain and Cyprus. These two countries evidenced a trend line steeper than usual in both gender groups.

[6] The first year, 2004, was not considered in the trend analyses of the standard deviation because of the lack of data in that year. To recover from these deficiencies we started from 2005 to assess the evolution of the standard deviations.

Figure 1a. The Males' HLE path of 12 European countries that have an initial HLE value below 66.

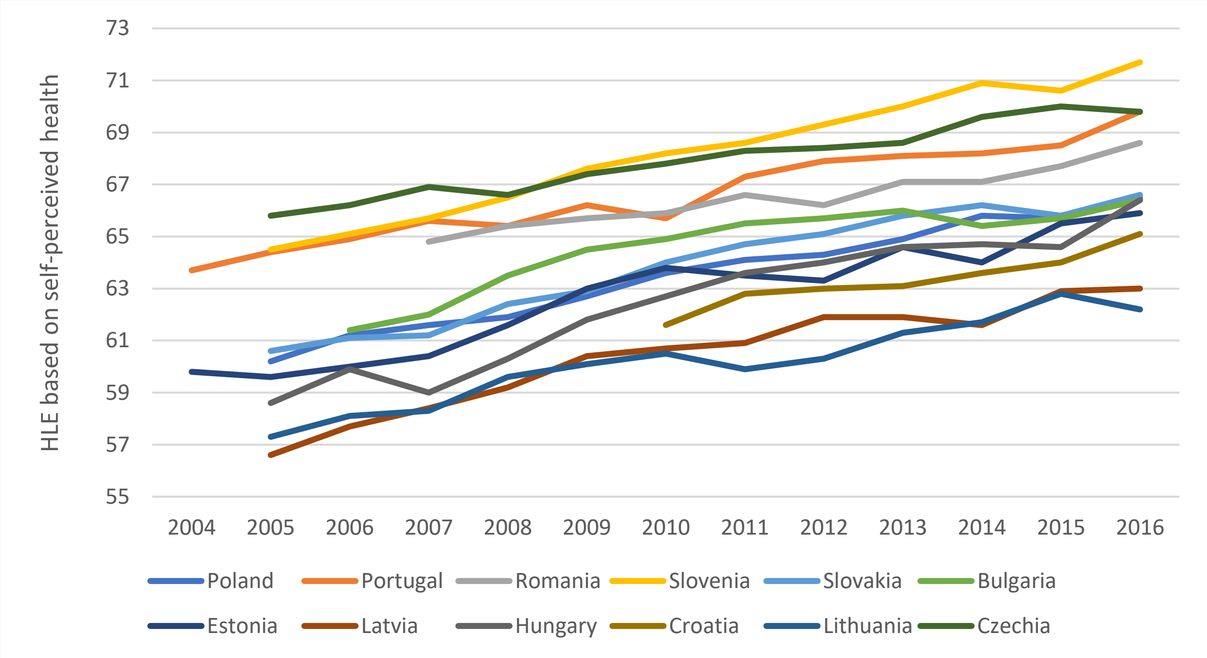


Figure 1b. The Males' HLE path of 19 European countries that have an initial HLE value above 66 .

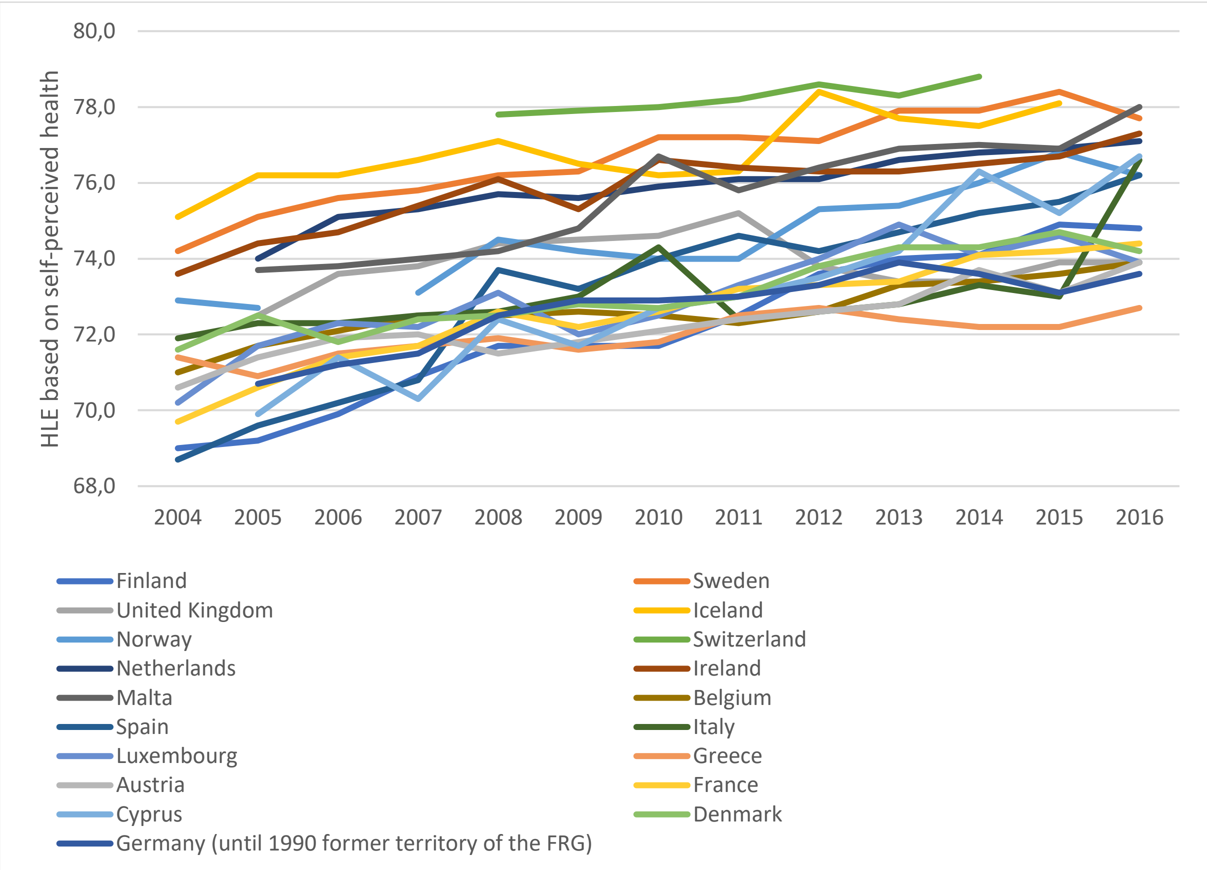


Figure 2a. The Females' HLE path of 12 European countries that have an initial HLE value below 70.

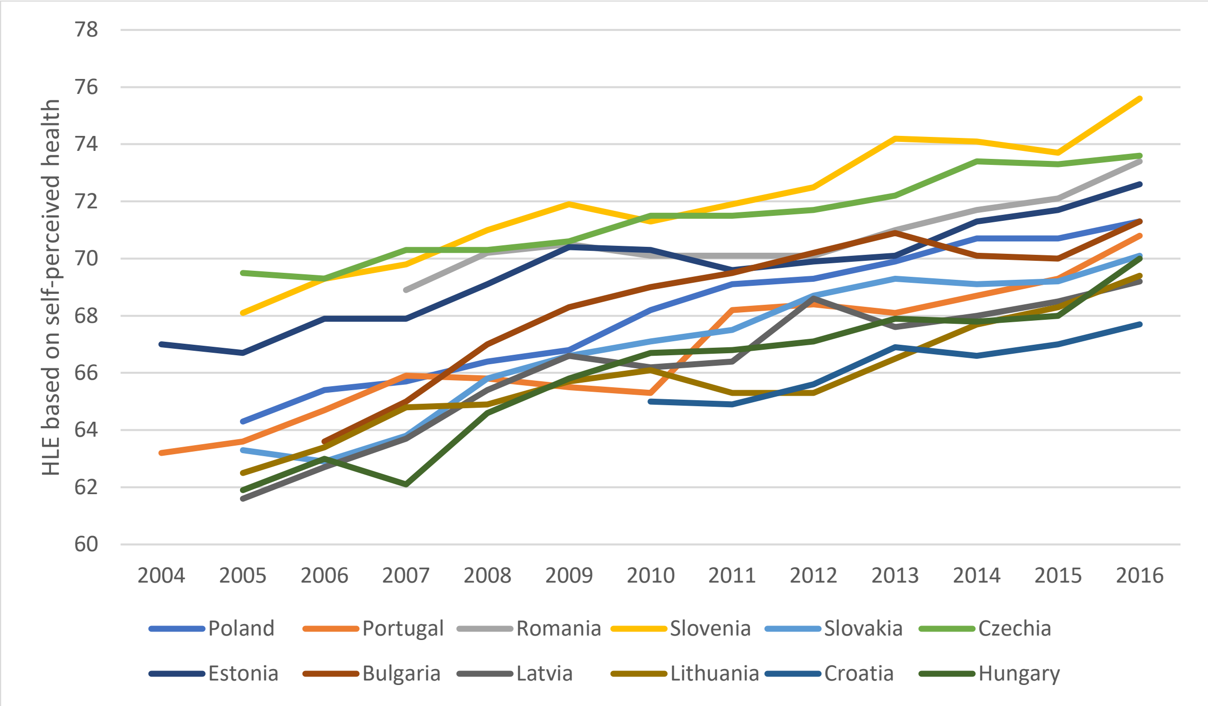


Figure 2a. The Females' HLE path of 19 European countries that have an initial HLE value above 70.

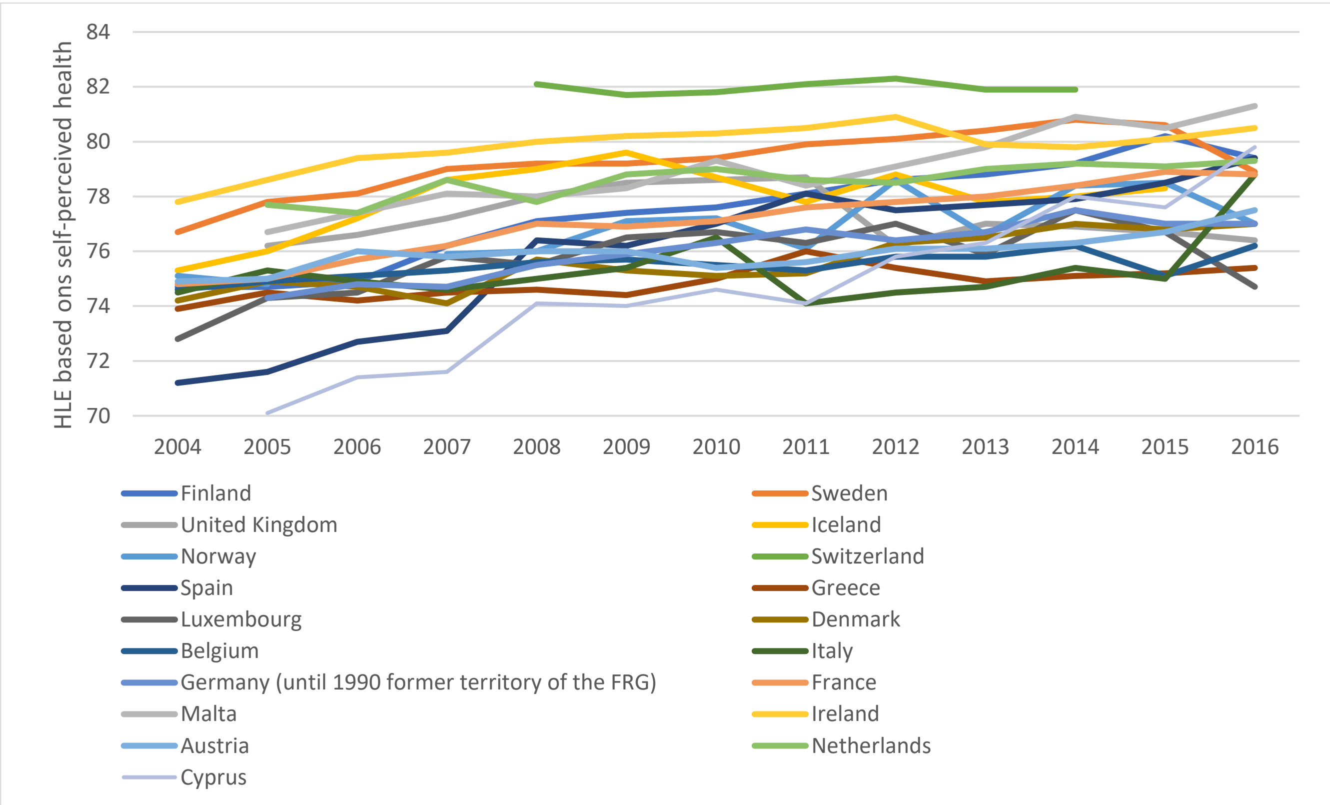
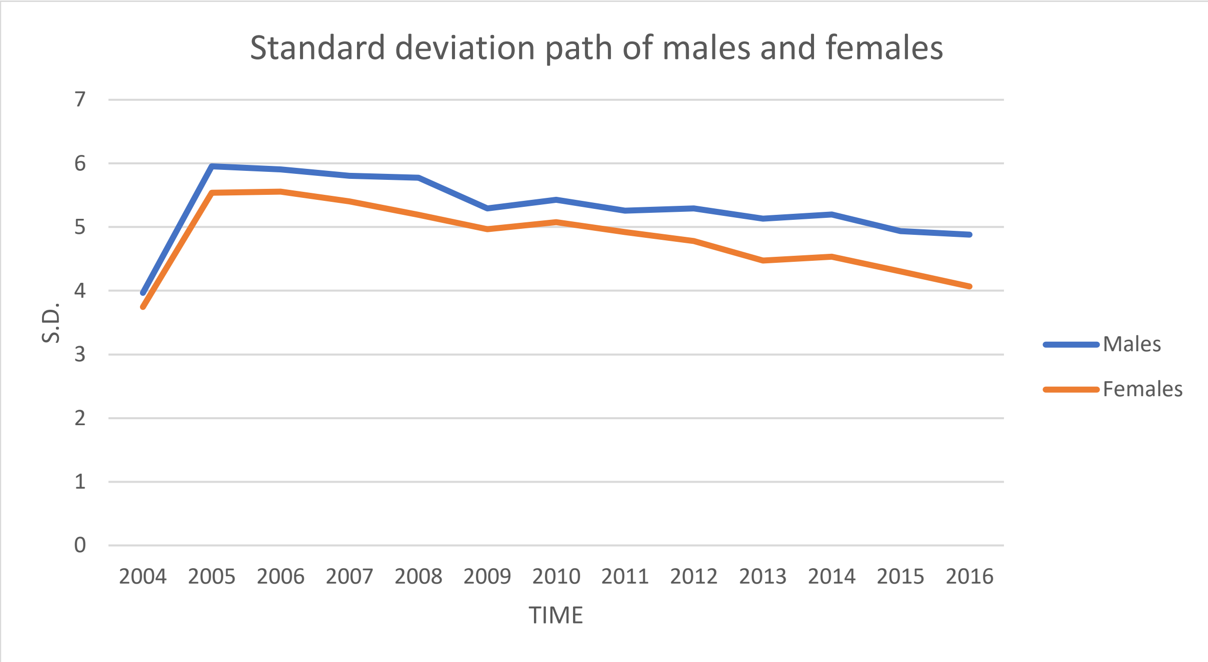


Figure 3.


Standard deviation path of males and females
S.D.
7
6
5
4
3
2
1
0
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
TIME
Males
Females

3. Methodology

3.1 Measuring health inequality

3.1.1 Gini coefficient

The Gini coefficient measures the inequality among values of a frequency distribution that in our case are the values of HLE across countries. It varies between 0 and 1. The value of 0 corresponds to perfect equality of health and, in the opposite, a value of 1 corresponds to perfect inequality.

From John Creedy (2015) and Lerman and Yitzhaki (1989)[3][11] it is confirmed that the covariance-based expression for the Gini coefficient can easily be modified to deal with the sample weights, that in our case are the population sizes for each country.

For incomes x_i i= 1, 2, …, n and incomes ranked in ascending (strictly, non-decreasing) order, the covariance expression for the Gini, G, is:

$$G= (2/x)*Cov(x, F(x)) \tag{1}$$

Where F(x) is the distribution function and x is the weighted average of income and Cov(x, F(x)) is the covariance.

Lerman and Yitzhaki (1989) assume that if each observation has a weight, w_i the sum of the weights is equal to 1, F(x) is obtained, where $w_0 = 0$, as:

$$\mathbf{F(x)= w_i /2 + \sum_{i=0}^{i-1} w_i} \tag{2}$$

The estimate of the Gini coefficient is thus :

$$\mathbf{G= (2/x) \sum_{i=1}^{n} w_i (x_i - x)(F(x)- \dot{F})} \tag{3}$$

Where $\dot{F}$ is the weighted mean of F(x).

3.1.2 Dispersion measure of mortality

The DMM measures the degree of dispersion that exists at any point of time in the mortality experiences of the countries in the world or states in a country. It is calculated as the average absolute intercountry mortality difference, weighted by population size, between every pair of countries. Change in the DMM over time indicates whether mortality is becoming more or less similar across Europe; a decrease indicates convergence, whereas an increase indicates divergence (Moser et al., 2004)[13]. It is also used to investigate the dispersion in LEB. In our case, we adopt it to see the degree of HLE dispersion existing in Europe. Since the Gini coefficient can be also assumed to be the ratio between the DMM and the weighted average, Goli S., Moradhvaj, Chakravorty S., Rammohan A. (2019)[6] we can easily determine the DMM by multiplying the Gini coefficient by the average of HLE weighted by the population size. Without computing the average absolute interstate HLE difference in Europe.

3.2. Absolute Convergence

3.2.1 β-convergence

One of the many definitions of convergence explains how its occurrence comes up with the reduction of dispersion across individuals. Although, a part of convergence literature sustained that **β-convergence** is irrelevant for convergence analyses, Danny Quah(1993)[15], another part showed that is a necessary condition, Xavier Sala-i-Martin(1996)[18] and that it gives us the possibility to observe how the units will converge. Sala-i-Martin in his study "Regional cohesion: Evidence and theories of regional growth and convergence" makes clear the importance of **β-Convergence.** Consider two economies, A and B, with the same degrees of income inequality over 50 years. Economy A is mainly agricultural. The scarce land is controlled by the small group of privileged who bequeaths it to their children. Hence, the children of the landowners remain rich, and the children of the poor end up being poor. Instead, economy B is centered around the industrial sector. A few skillful entrepreneurs that have good ideas and can implement them are the rich owners of the company. The rest of the population works for them. Then, some of the workers' children have good ideas and entrepreneurial skills, so they decide to run their own companies. Some of the entrepreneurs' children don't have the same ability as their parents, so eventually, they lose their fortunes. After 50 years the degree of income inequality remains constant, but the wealth is held by different families. Economy B corresponds to what we call β-Convergence, so when the average growth rate of an economy is inversely correlated to its starting value. Hence, **β-Convergence** becomes crucial to answer these questions: which economy we are observing? How fast do the rich become poor and the poor rich, or Is it possible to convert economy A to economy B and vice versa?

Speaking about HLE, **β-Convergence** can help us to see: what kind of country we are observing, how quickly the countries with a lower HLE reach larger values, and the countries with larger values reach lower values.

To carry out our analyses we will use data on HLE based on self-perceived health, y_{it} , where the t denotes the year and i denotes the region. The **β-Convergence** equation describes how the annual average growth rates per each country over the period 2004-2016 are related to the log of initial values of HLE for each country at time t_0 as follows:

$$1/T * ln\, [\, y_{i,t0+T} / y_{i,t0}\,] = \gamma - (1-e^{-sT})/T \cdot ln\, y_{i,t0} + (1-e^{-sT})/T \cdot ln\, \hat{y}_i + \varepsilon_{i,t0,\, t0+T}$$

where $\varepsilon_{i,t0,\, t0+T}$ is a distributed lag of the error term ε_{it} between t_0 and T. The terms γ and $\hat{y}_i$, that in Barro and Sala-i-Martin(1992)[1] correspond to the rate of technological progress and the steady-state, are equal for all the countries. The greater s the greater the responsiveness of the average growth rate to the gap between $ln\, y_{i,t0}$, and $ln\, \hat{y}_i$, that is the more rapid the convergence toward the steady-state. The model implies conditional convergence, given that, the average growth rate is higher for a lower $ln\, y_{i,t0}$ when γ and $\hat{y}_i$ are constant. Because of that, we can substitute $(1-e^{-sT})/T \cdot ln\, \hat{y}_i$ and γ in one unique term α which joins both terms.

$$1/T * ln\, [y_{i,t0+T} / y_{i,t0}] = \alpha - (1-e^{-sT})/T \cdot ln\, y_{i,t0} + \varepsilon_{i,t0,\, t0+T} \qquad (4)$$

In the convergence literature Xavier Sala-i-Martin(1996) and Srinivas Goli and Perianayagam Arokiasamy(2013)[18][5], s is called the **speed of convergence**, thus the annual rate of convergence. It measures how fast the populations' HLE of each country converges toward the steady-state.

To estimate s we adopted the non-linear least-squares method because as already shown in Quah(1996) this avoids some errors. Indeed, if we rearrange equation 4 we will obtain :

$$1/T * ln\, [y_{i,t0+T}] = \alpha + b/T \cdot ln\, y_{i,t0} + \varepsilon_{i,t0,\, t0+T}$$

Where b is equal to e^{-sT}.

Performing Monte Carlo experimentation Quah confirmed that nothing can guarantee that the OLS estimate for b is positive, whatever the true data generating model, and first of all when b is negative, s is undefined.

In the second place, linearizing the equation 4, thus substituting β for - $(1-e^{-sT})/T$ we will have a distortion when we are going to calculate the speed of convergence[7] since it will have been affected by the size of the time interval.

Instead, using the non-linear least squares we will determine directly the speed of convergence taking into account the different time intervals used to calculate the annual average growth rate of each country.[8]

3.2.2 σ- convergence

Young et al. (2008)[24] and Sala-i-Martin(1996)[18] in their research proved that β-convergence is a necessary but not sufficient condition for σ-convergence; on the contrary, σ-convergence is sufficient but not a necessary condition for β- convergence. σ-convergence occurs when the dispersion of health indicator decreases. So, we expect that the standard deviation of the HLE based on self-perceived health decreases over time. We assess the existence of σ-convergence by testing:

$$\sigma_t > \sigma_{t+k}$$

where σ_t is the standard deviation of the indicator at time t. If σ_{t+k} is lower than σ_t implies convergence. To test the homogeneity of the variance, we adopted Levene's test, Levene, H. (1960)[10]. Because it allows us to prove the homogeneity of the variances at the median that is appropriate against many types of non-normal data and given the non-normality resulted from Shapiro-Wilk test, *Shapiro, S. S.;* Wilk, M. B. *(1965)*[21], the Levine's test becomes an appropriate instrument to our case.

[7] When we use linear regression the speed of convergence s corresponds to $-ln(1+T\cdot\beta)/T$.

[8] This is due to the lack of annual information for some European countries like Bulgaria, Germany, Croatia and the rest of countries mentioned before in the description of data.

3.3. Convergence clubs

3.3.1 Dynamic approach

In estimating s by NLS, as explained in the preceding paragraph, the specification of the disturbances $\varepsilon_{i,t0,\,t0+T}$ can be relatively arbitrary without affecting the conclusion. They can have different variances[9]. They can be serially correlated[10]. None of these changes the conclusion. Then, β-convergence will produce a biased estimate of the speed of convergence. Quah(1996) showed these and other types of deficiencies of the regression-based approach used by Barro and Sala-i-Martin for the study on convergence. He explained that disturbances are an ongoing series and not one-shot, observed uniquely at the beginning or at the end of the considered period. So, β-convergence estimates an average behavior and not the behavior of the entire distribution. Moreover, if the disturbances are ongoing the possibility that σ will tend to 0 vanishes, Danny T. Quah(1996)[16].

Quah sustains that the natural way to study convergence empirics is to provide an empirical model for how the distribution evolves. This framework studies the evolution of the distribution of HLE over time by analyzing the kernel density plots of initial, final, and long-run distributions, identifying the formation of convergence-clubs, polarization, or persistent inequality.

Assume that the distribution of the HLE, as for males as for females, is noted at time t as φ_t. Then the simplest form of the dynamic of this distribution can be represented using the Markov chain assumption. This is similar to the first-order autoregressive process and hence the dynamic of the distribution at time t is:

$$\varphi_t = T(\varphi_{t-1}, u_t) \qquad t \geq 1 \tag{5}$$

where u_t is the disturbance term and T is the operator that maps how one part of the distribution evolves to another, from t-1 to t.

Therefore, the dynamic of the distribution can be alternatively written as:

---

[9] Any shocks can have different impacts on the HLE of each European country.

[10] Aghion, P., Howitt, P., and Murtin, F. (2010) and Zhang, J. and Zhang, J. (2005) show that individuals with higher life expectancy are likely to save more, and saving turn into capital accumulation and therefore in GDP growth, and in the second place they invest more in education which in turn should be growth-enhancing. GDP and education are both positive determinants for HLE.

$$\varphi_t = T_{ut}(\varphi_{t-1}) \qquad t \geq 1 \tag{6}$$

When T absorbs the disturbance term it becomes a stochastic kernel.

The above equation indicates that the cross-section distribution changes from its current state to another state according to a certain probability distribution.

The changes of the state are called transitions and the functional relationship of these transitions is called transition probability function or a stochastic kernel.

In equation (6), T is the stochastic kernel that shows how this distribution evolves. It contains information about the shape and the dynamic of the distribution. Iterating the process with equation (6) and the Markov chain assumption, we get:

$$\varphi_{t+1} = M\varphi_t \qquad t \geq 1 \tag{7}$$

or:

$$\varphi_{t+s} = M^s\varphi_t \qquad s \geq 1 \tag{8}$$

Similarly, iterating the system up to infinity we get the long-run (ergodic) distribution.

Therefore, the ergodic distribution of HLE of both groups is as follows:

$$\varphi_\infty = M\varphi_\infty \qquad s \longrightarrow \infty \tag{9}$$

φ_∞ is the long-run limit of the distribution of HLE based on self-perceived health across regions. If the distribution after s periods (φ_{t+s}) and/or the ergodic (long-run) distribution (φ_∞) show a trend towards a point mass, it is indicative of convergence over time. Alternatively, if φ_{t+s} or φ_∞ shows a trend towards bimodality, it can be concluded that the distribution tends towards polarization[11]. If more than two modes are identified, then that is evidence of stratification of the distribution.

In order to operationalize these concepts, the stochastic kernel T_{ut} has to be present in a continuous space.

[11] This means that some regions catch up with one another but only within particular subgroups (Baumol, 1986)

Let us assume that $[(y_1,z_1);(y_2,z_2);\ldots;(y_n,z_n)]$ are the pair of HLE based on self-perceived health observed for each in the two different times t and t+s[12] and n represents the number of countries. Here, y and z represent the initial value of HLE and its value after s periods.
If the cross-section distribution of HLE is represented by the density functions $f_t(y)$ and $f_{t+s}(z)$ at time t and t+s, respectively, and assuming that the process that describes the evolution of distribution is time-invariant respecting the Markov assumptions, then the stochastic kernel is defined by the equation:

$$f_{t+s}=\int g_s(z/y)f_t(y)dy \tag{10}$$

g_s is the conditional probability density function after s times and it is the stochastic kernel.
Similarly, the ergodic distribution is :

$$f_\infty=\int g_s(z/y)f_\infty (y)dy \tag{11}$$

The conditional distribution in equations (10) and (11) is by definition, the ratio between the corresponding joint distributions and the marginal distributions[13].
The joint distribution can be estimated using kernel density functions[14][15]. The supports of y and z were discretized in a set of N equally large intervals, where interval h is denoted as Ω_h. Then, integrating the joint distribution for y we are finding the marginal density of z.
For the settlement of the long-run distribution of HLE, we need to involve matrix operations. Given the estimate of the stochastic kernel $g_s(z/y)$, you can compute the implied long-run or ergodic kernel of the HLE relative to both groups, males and females. The long-run density, $f_\infty(z)$, is the solution of the following expression:

$$f_\infty(z)=\int g_s(z/y)f_\infty (y)dy$$

[12] That, in our case, correspond to 2005 and 2016 because of the missing data in 2004.
[13] To find out the stochastic kernel or the conditional distribution function these following steps are necessary: Transposing the joint pdf, which appear as a 900x1 column vector, in a 30x30 matrix with the command *reshape*, where the rows and the columns correspond to y and z respectively; Then, the summation of all the columns will appear like an integration for z that lead us to estimate the marginal density of y. Finally, to estimate the conditional distribution will need to compute the ratio between the joint pdf and the marginal pdf. All the calculations were done on MATLAB.
[14] We use the product kernel, that you can see in the next section, to estimate the joint distribution between data of 2005 and 2016 by the command *ksdensity* on MATLAB.
[15] The domain is discretized in a set of 30 intervals, equally large. The width of the interval h is calculated using the rule of thumb created by Scott(2015) which will be explained in more detail in the next section.

or in a matrix form

$\varphi_\infty = M\varphi_\infty$

Rearranging it you will get:

$(I-M)\ \varphi_\infty=0$

From the last expression presented in matrix form, you can deduce that the ergodic distribution will be the vector φ_∞ that solves this system of equations. We employed such a discretization of our support to estimate our ergodic distribution as we did already in the estimation of the stochastic kernel.

Each component of the vector φ_∞ represents the probability of HLE[16] assuming a value comprised in a given Ω_h and M is the matrix of transition probabilities m_{hihj}[17] from an interval Ω_h to another.

To discover the solution to our system of equations we must bring some elementary operations on our stochastic kernel M, Dal Bianco Silvia (2010)[4].

Since each column of the matrix, M is a conditional density and, then, the sum of its elements is 1; M does not have full rank and, by consequence, the system does not have a unique solution. To find a unique solution it is standard to simply drop one row of M (to make its columns linearly independent) and then add the restriction that the entries of vector φ_∞ sum to 1.[18]

Then the Matrix (I-M) is rewritten as matrix B, where on the diagonal the entries are equal to $1-m_{hihj}$ for every $i,j=1,2,\ldots,30$ and $i=j$, for all the rest of entries they are equal to $-\ m_{hihj}$ where $i\neq j$.

Now, as the last row of our matrix B corresponds to a row vector of 1s, our system of equation will not be equal to 0, but instead, to b.

[16] ϕ_∞ will be 30x1 vector where 30 corresponds to the number of intervals chosen.

[17] We used the letter h to identify the optimal bandwidth for both variables, y, and z. To avoid confusion it must be said that in the subscript of the transition probability m the first corresponds to the intervals of y and the second corresponds to the intervals of z.

[18] The restriction is due to the definition of probability.

$$B=\begin{pmatrix} 1-m_{h1h1} & \dots & -m_{h1h30} \\ \dots & 1-m_{h2h2} & \dots \\ -m_{h29h1} & \dots & -m_{h29h30} \\ 1 & 1 & 1 \end{pmatrix}$$

Where b is a column vector of 0s except for the last row, because it will be substituted by 1 due to the restriction of the probability.

Then our equations system is going to be:

$$B\varphi_{\infty}=b$$

To calculate the ergodic[19] is extremely simple now since is enough to compute the inverse of B.

$$\varphi_{\infty}=B^{-1}b$$

The concepts described so far are the basis for the distribution dynamic approach.

Now, the analysis of convergence consists of studying the kernel density distributions estimated. The initial and the final distribution of HLE are determined using the univariate kernel density estimation as the HLE based on self-perceived health observed in 2005 and 2016 as variables and the long-run distribution or ergodic distribution. The number of modes on the ergodic distribution determines whether the system is moving towards convergence, polarization, or stratification. Finally, this approach also uses the stochastic kernel to identify the intra-distributional dynamics that are responsible for these above results.

---

[19] To calculate the inverse of B we used the command *inv()* on MATLAB.

3.3.2 Kernel density estimation

According to Parzen(1962)[14] one of the possible estimates of the probability density function can be written as the weighted average over the sample distribution function.
Let $X_1, X_2, \ldots, X_n$ be independent random variable identically distributed as a random variable X, the estimate of a probability density function is :

$$f_n(x) = (1/(nh)) \cdot \sum_{i=1}^{n} K((x-X_i)/h) \qquad (12)$$

where h is a suitable positive number called bandwidth, n is the number of observations and K is the kernel function.
In case of multivariate data, we have a sample of size n where each observation is a d-dimensional vector X_i i=1,2,…,n and the simplest case of multivariate kernel estimator is a product kernel[20], Wendy L. Martinez Angel R. Martinez (2002) [12]. It is expressed in the following formula:

$$f_n(x) = (1/(nh_1 \ldots h_d)) \cdot \sum_{i=1}^{n} \{\Pi_{j=1}^{d} K((x_j-X_{ij})/h_j)\}$$

Where the j component of the X_i observation corresponds to the years which we refer to estimate the joint distribution.
The kernel can assume distinct forms. The most common are: the Epanechnikov kernel density function, the Gaussian, and the Triangular.
Due to its convenient mathematical properties, we used the Gaussian kernel density, although most researchers agree that the estimation is not sensitive to the type of kernel but rather to the bandwidth since the kernel tail behavior is reduced from the averaging process, Wendy L. Martinez Angel R. Martinez (2002)[12].
The bandwidth of the kernel, h, is a free parameter[21] and it will drive how many values are included in estimating the density at each point.

[21] A **free parameter** is a variable in a mathematical model which cannot be predicted precisely or constrained by the model and must be estimated experimentally or theoretically.

The choice of an appropriate bandwidth is – on the other hand – a very important step as it has a significant impact on the density estimates.

The optimal bandwidth is the bandwidth that minimizes the AMISE (the asymptotical mean integrated squared error), thus the bandwidth that minimizes the distance between the estimated density and the true density function of our sample when n goes to infinity. To have the formula for AMISE we need to start from the mean integrated squared error.

It is itself stated as :

$$MISE(f; n, h) = E \int \{f_n(x) - f(x)\}^2 dx \quad (13)$$

This can be stated also as the sum of the variance of the estimated density and its squared bias is:

$$MISE(f; n, h) = \int Var\, f_n(x)\, dx + \int \{Ef_n(x) - f(x)\}^2\, dx \quad (14)$$

Using Taylor expansions of the density will have been estimated by the kernel density estimation, the formulas for the variance and the bias are (Wendy L. Martinez Angel R. Martinez, 2002 and Scott, 2015) [12][20] :

$$Var\, f_n(x) = (R(K)/(nh)) \cdot f(x) + o((nh)^{-1}) \quad (15)$$

$$BIAS = \tfrac{1}{2}\, \mu_2(K)\, f''(x) h^2 + o(h^2{}_n) \quad (16)$$

Where R(K) denotes the squared integral of our kernel and $\mu_2(K)$ represents the second moment of the kernel.

Then, choosing a small h we obtain an estimator with a small bias but with a large variability. In opposite, if h is too large we obtain an estimator with large bias but with small variability. Substituting the equations (15) and (16) in equation (14) we can determine the mean integrated squared error of our estimated probability density function.

$$MISE = R(K)/(nh) + o((nh)^{-1}) + 1/4\, \mu_2{}^2(K)\, R(f''(x)) h^4 + o(h^2{}_n)$$

We don't have any more f(x) in the variance part because we computed the integration in the MISE, and since f(x) is a pdf its integration will be equal to 1.
The same thing occurs for the part concerning the bias because you can notice that there is no more f ''(x) but it is substituted with R(f ''(x)) that defines, as mentioned before, the squared integral.
Asymptotically, remains only the dominating part of the MISE, which is called AMISE.

$$AMISE= 1/4\ \mu_2^2(K)\ R(f''(x))h^4 + R(K)/(nh)$$

Computing the derivative of the AMISE and putting it equal to 0 we discover the optimal bandwidth, h, that minimize the error.

$$h_{AMISE}= \{R(K)/[\mu_2^2(K)\ R(f''(x))]\}^{-1/5}$$

A simple solution to estimate R(f ''(x)) is considering f as a normal density since it is infinitely differentiable[22]. This allows us to determine easily an optimal h.
According to Scott (2015)[20] the optimal h can be assumed as follows:

$$h_{RT}= 1.06n^{-1/5}\sigma$$

Instead, in case of multivariate data is suitable to include the number of dimensions in the choice of h. Properly, Scott(2015) proposed a general normal reference rule ;

$$h_{RT}=\{[4/(d+2)n]^{\wedge}1/(d+4)\}*\sigma$$

To make the h value more robust to fit better for long-tailed, skew distribution, and bimodal mixture distribution, it is better to substitute the estimated value of σ with the minimum between standard deviation and the inter-quantile range/1.34.

[22] As Scott(1992), this assumption of f is not so strong as a parametric Gaussian assumption i.e. use of equation h_{RT} on non-Gaussian data will not result in a distribution that looks Gaussian.

4. Results

4.1 Convergence in absolute and relative dispersion of HLE

As the first objective of our study, we began to study convergence measuring the overall health inequalities in Europe. We did it through the use of two types of index: the DMM and the Gini coefficient for the absolute and the relative inequality. As we said before, if the trend of both measures is decreasing along the years under consideration, we can presume a possible convergence of HLE across European countries. We calculated the Gini coefficient according to Lerman and Yitzhaki (1989) and since it can be defined as the ratio between DMM and the weighted average we determined the dispersion of mortality measure along the years. In 2004, as for males as for females, the Gini and DMM are at a historic minimum, but then, both experienced an immediate enormous growth in the next year. This is due to changes in the sample given the entry in the data of several countries about which there was missing information in 2004.

Figure 4.

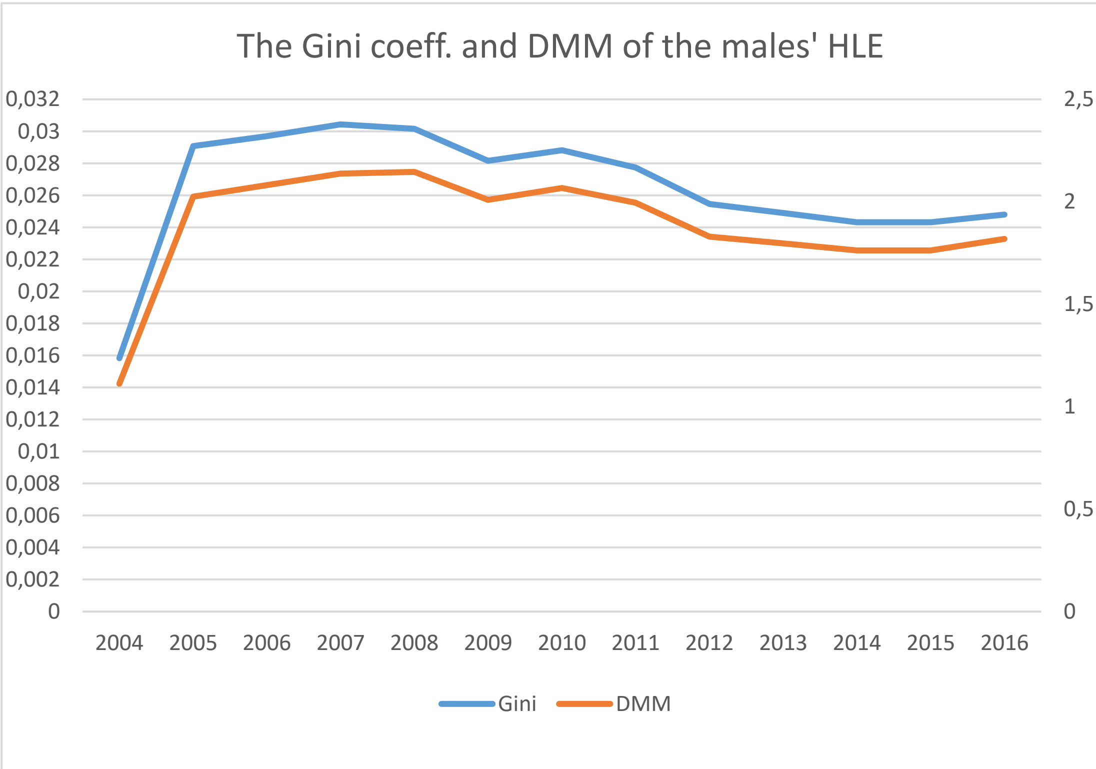
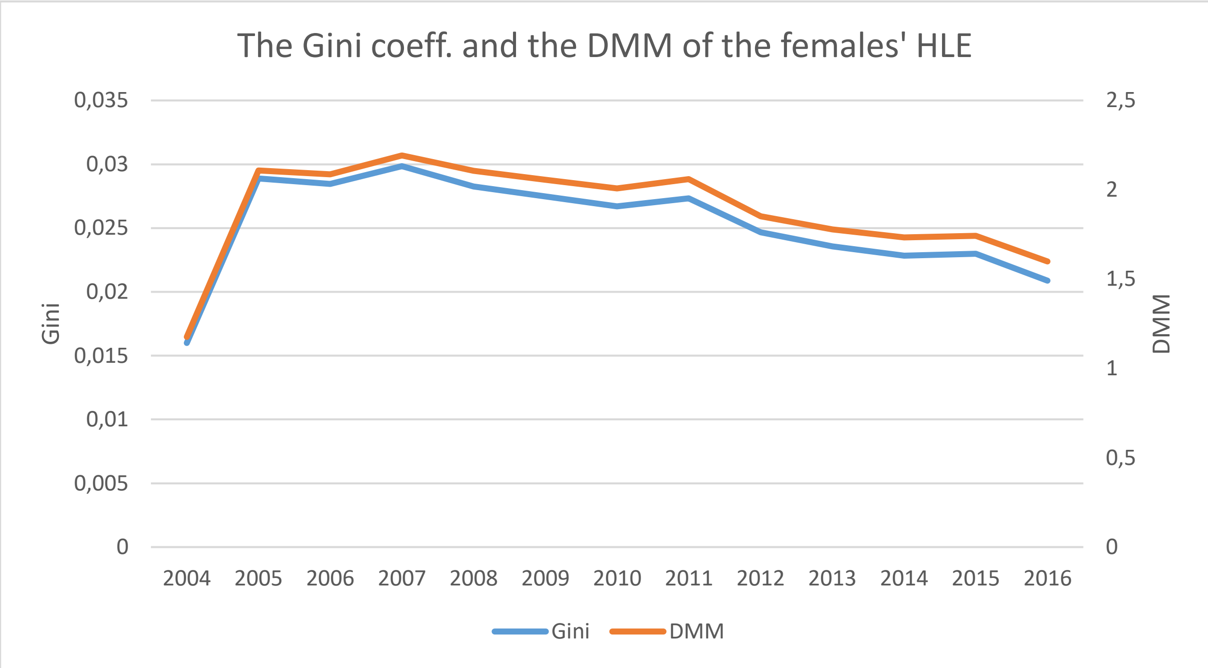


Figure 5.



The males' DMM and Gini coefficients estimated display fluctuations until 2010 then there is a decreasing period that stopped in 2014 followed by an increase in the last years. The same occurs for the women which experienced several changes in absolute and relative inequality in the first years, until 2007. Then, a fall of HLE inequality persists as far as 2011 where the DMM and Gini coefficient arrives at 2.058 and 0.027 respectively.

From this analysis emerges a different trend in DMM and Gini coefficient between males and females. As you can see from Figures 3 and 4 in the last year of our considered period, the males' DMM and Gini index recorded an increasing trend revealing a possible divergence in the future. Meanwhile, the DMM and the Gini index observed in the female population of the European countries registered a decreasing trend in 2016. This lead us to conclude the females' HLE in Europe will converge. Convergence in health inequalities is a sufficient condition for convergence in average but it is not a necessary condition, Srinivas Goli and Perianayagam Arokiasamy(2013)[5]. In the next paragraph, we will see the results provided by β-Convergence and σ-convergence to confirm if convergence in average will take place.

4.2 Absolute convergence

Given the positive results obtained in the analysis of convergence in absolute and relative dispersion of HLE about the female population, the research aims to go deep to uncover if this corresponds to convergence in average.
It led us to know if the HLE will converge in the future toward a unique steady-state in Europe. The purpose of this section is to verify if the conditions of absolute convergence are satisfied. To examine them, we used the most common approach adopted by Barro and Sala-i-Martin whose adopted in their research the **β-convergence** and the **σ-convergence**. The absolute convergence is used where the gap between countries with an inferior level of HLE and countries with a higher level of HLE shrinks. This happens when the first one increases more rapidly the HLE than the second one. We analyzed this using NLS on a regression model that assumes the average growth rates as the dependent variable and the logarithm of the starting value of HLE as the independent variable.

Table 2.

a) Results of the NLS using the males' HLE data

N° obs.	**Avgr_M**	**coefficient**	**standard errors**	**t-stat**	**p-value**	**R^2** (adjusted)
31	LnYM_0	.0163165	.003373	4.84	0.000	0.4801
	constant	.067765	.0116726	5.81	0.000	

b) Results of the NLS using the females' HLE data

N° obs.	**Avgr_M**	**coefficient**	**standard errors**	**t-stat**	**p-value**	**R^2** (adjusted)
31	LnYF_0	.0201423	.0043778	4.60	0.000	0.4375
	constant	.0813892	.0146338	5.56	0.000	

Table 2 (a) and (b) have shown the speed of convergence experienced from males and females on the basis of the average growth rate of each country calculated using data recorded from 2004 to 2016. The speed of convergence related to HLE of male population equals 0.016, on the other side that referred to female population is equal to 0.02. Although we know that in the base of the results emerged from the analyses of convergence in dispersion the HLE of males won't converge in the future, these results lead us to find that the HLE of the female population converges faster than males' HLE. This is due to the fact that the catching-up process faced in the female population is much more evident than in the male population. In a few words, this means that the HLE of female population of European countries that have a smaller starting value increases more quickly on average than male population that starts at the same value. This is a good condition to forecast a reduction in the standard deviation of HLE. Figure 3. already showed how the HLE standard deviation trend of both genders decreased until 2016, but we need to assess the statistical significance of this reduction. To do that, we involved a statistical test evaluating the homogeneity of variances across years.

Before, we looked at the shape of distributions that the cross-sectional data of HLE assume in each year to apply the most fitting test to check the validity of the variance equality hypothesis. To control if the cross-sectional data are normally distributed we opted for using the Shapiro-Wilk test.

Table 3. Results of Shapiro Wilk test by genders.

YEAR	Obs.	W		V		p-value	
		Males	Females	Males	Females	Males	Females
2004	15	0.85982	0.78023	2.718	4.261	0.02399	0.00207
2005	27	0.87572	0.85717	3.654	4.199	0.00389	0.00160
2006	27	0.87850	0.87457	3.572	3.687	0.00446	0.00368
2007	29	0.88683	0.91078	3.507	2.765	0.00481	0.01792
2008	30	0.88991	0.90816	3.499	2.919	0.00480	0.01338
2009	30	0.90578	0.91291	2.995	2.768	0.01166	0.01763
2010	31	0.90955	0.91673	2.946	2.712	0.01258	0.01934
2011	31	0.91614	0.93877	2.731	1.994	0.01867	0.07631
2012	31	0.91764	0.93817	2.683	2.014	0.02044	0.07346
2013	31	0.91437	0.93428	2.789	2.141	0.01678	0.05739
2014	31	0.90978	0.92647	2.939	2.395	0.01276	0.03518
2015	30	0.90964	0.92267	2.872	2.458	0.01457	0.03148
2016	30	0.92062	0.95029	2.523	1.580	0.02783	0.17211

Table 3. reports the test statistic W, the determinant of the variance-covariance matrix V, which is referred to as an order statistic[23] of identically distributed and independent random variable sampled from the standard normal distribution and the p-value of the test statistic, all of them for both genders. In addition to p-values the semi-parametric test provides V which has a value

[23] In statistics, the *k*th **order statistic** of a statistical sample is equal to its *k*th-smallest value.

approximately equal to 1 for samples from a normal distribution.[24] Large values of V indicate non-normality, Roystone,1991d [17]. Given the results in the table we find that the cross-section data of HLE does not come from a population normally distributed for all the years of the period taken into consideration. For this reason, we cannot apply the F-test since it requires two samples that derive from a population normally distributed. To take care of this shortage we decided to employ the Levine's test since it is capable to prove the homogeneity of the variances at mean, using the 10% trimmed mean and at the median that is appropriate for the cases of involvement of non-normal distribution[25].

Table 4. Results of Levine's test by genders.

p-value	Males	Females
at mean	0.91695671	0.46768775
at median	0.99986804	0.98965448
at trimmed mean	0.95228839	0.55219035

We are interested to look at the p-value at median because the test preserves its robustness against data that are not normally distributed. The null hypothesis of Levine's test assumes the homogeneity of the variances across the groups that in our case coincides with years of our time interval. The p-value is so high to not reject the null hypothesis under every significance level usually used in statistics[26]. Finally, we can conclude that although the trends of HLE standard deviations of both genders tend to go down over time, the trends themselves are not statistically significant in order to ascertain a possible future convergence of HLE among European countries.

[24] The significance level of the test is equal to 0.05.
[25] We tested the homogeneity of the variances on a sample of 31 European countries. We used data starting from 2005 since in the first year of our time interval we recorded only 15 real observations.

[26] Usually, the significance level α is fixed at 5% or below, but the distribution tables show a significance level until 0.99.

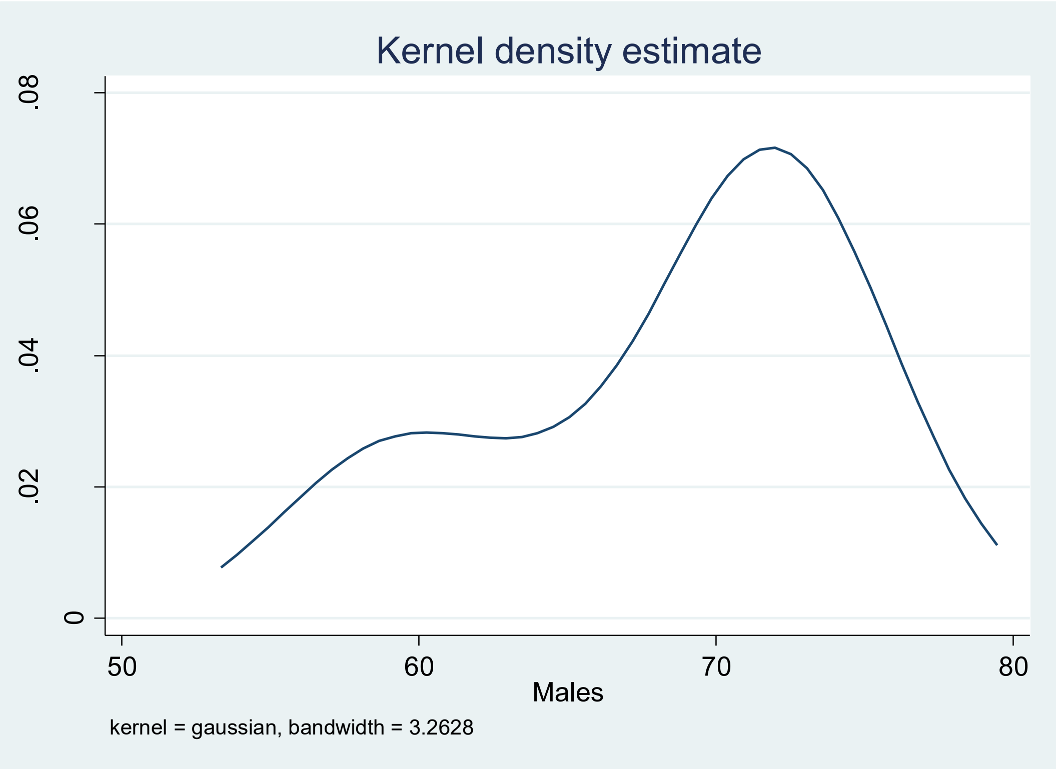
4.3 Convergence clubs

In the last section of our research, we saw that there is not an absolute convergence relative to the HLE registered in Europe. The last thing to do to make predictions on the future evolution of HLE is to verify the hypothesis of club convergence[27]. The analyses based on the study conducted by Quah(1996)[16], consists of pursuing the following steps. In the first step, we estimate the density functions of HLE of the first[28] and the last year of our period considered. This procedure is a way to test if the unimodality in HLE distribution is present or not. In the second step, we estimate transition probabilities to analyze the mobility within the HLE distribution. This means we are going to examine how a given individual in a given point in time transits to another part of the distribution in the future. In the third step, we calculate the ergodic kernel density of HLE under the basis of the transitions probabilities matrix. This permits us to make long-run predictions on HLE in Europe. The distributions of the first and the last year are determined according to the principles of the kernel density estimation that we see in Parzen(1962)[14]. They were estimated based on data of HLE recorded in 2005 and 2016 with an appropriate bandwidth chosen following the normal reference rule by Scott(2015).

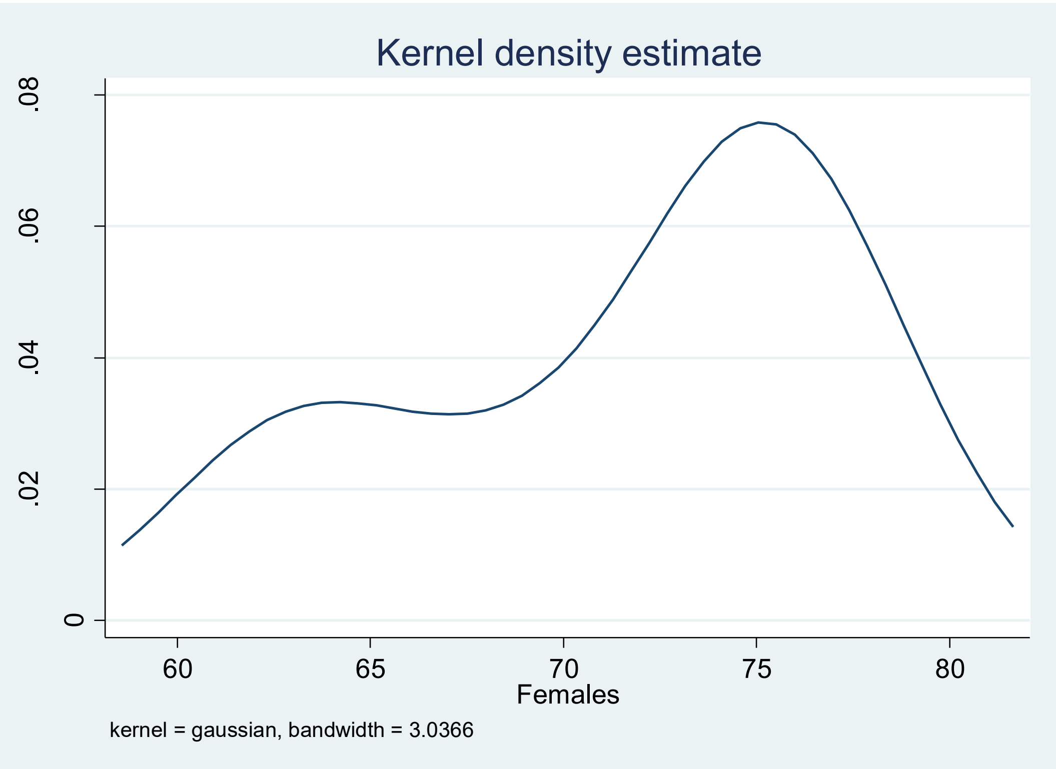
[27] It is different from conditional convergence since does not require a higher growth rate of HLE for countries with short-lived population.

[28] Although the first year corresponds to 2004 we chose 2005 as the first year for the quantity registered in that year since we know that in 2004 there is missing data.

Figure 6. a) The density distribution of the males' HLE in 2005.

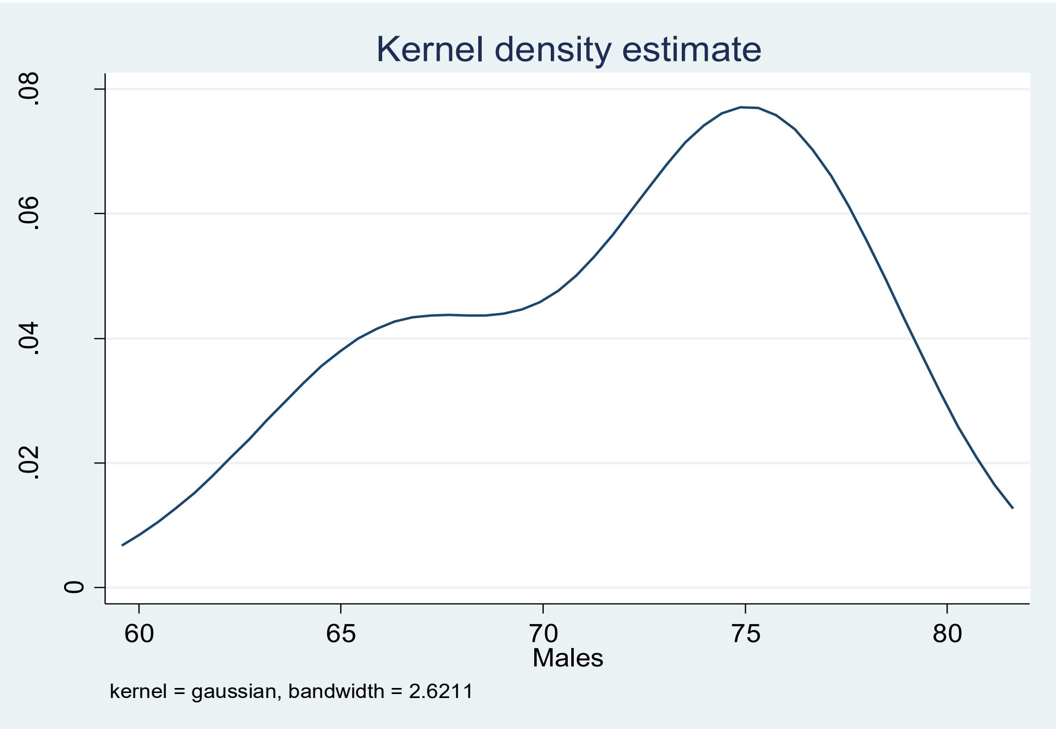


b) The density distribution of females' HLE in 2005.

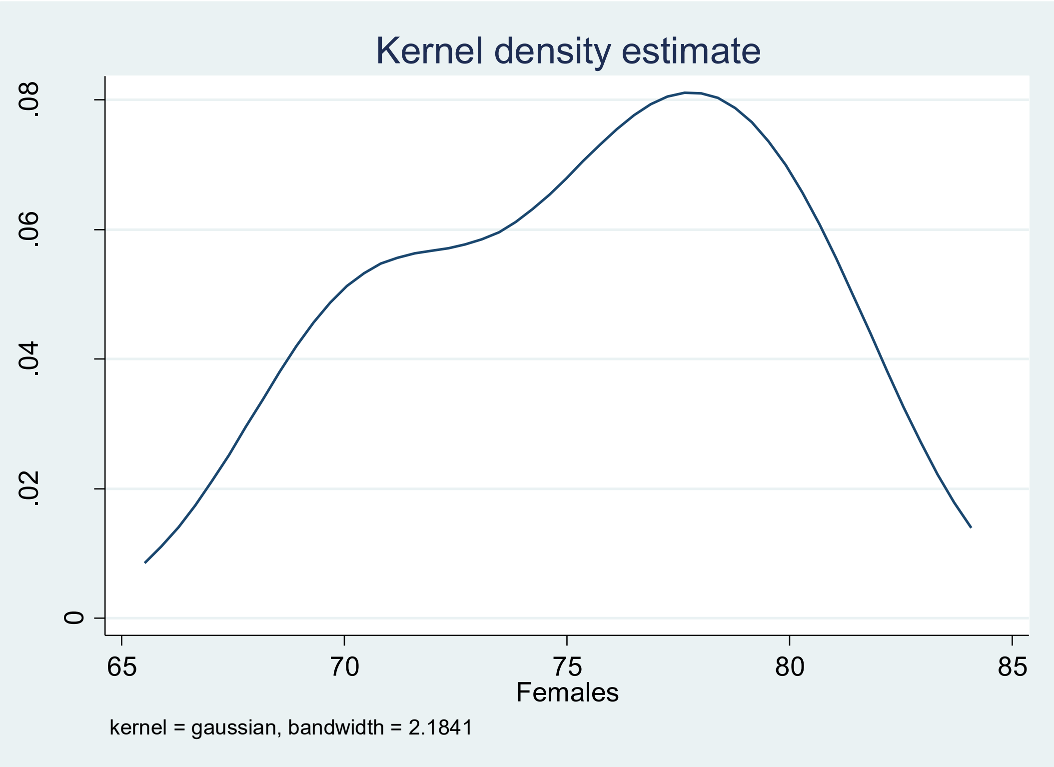


The distribution of HLE related to male data presents two peaks around 60 and 70. The last one is the highest describing a major frequency of countries in Europe where male people live longer and healthier. The distribution related to females data of HLE is always bimodal but it shifts to the right getting greater values for both peaks. The highest one corresponds to 75 years while the smallest is positioned around 65 years. The same here, the female data describes a higher frequency of countries that have female people that live longer and with a health status superior to bad.

Figure 6 a) The density distribution of the males' HLE in 2016.



b) The density distribution of the females' HLE in 2016.



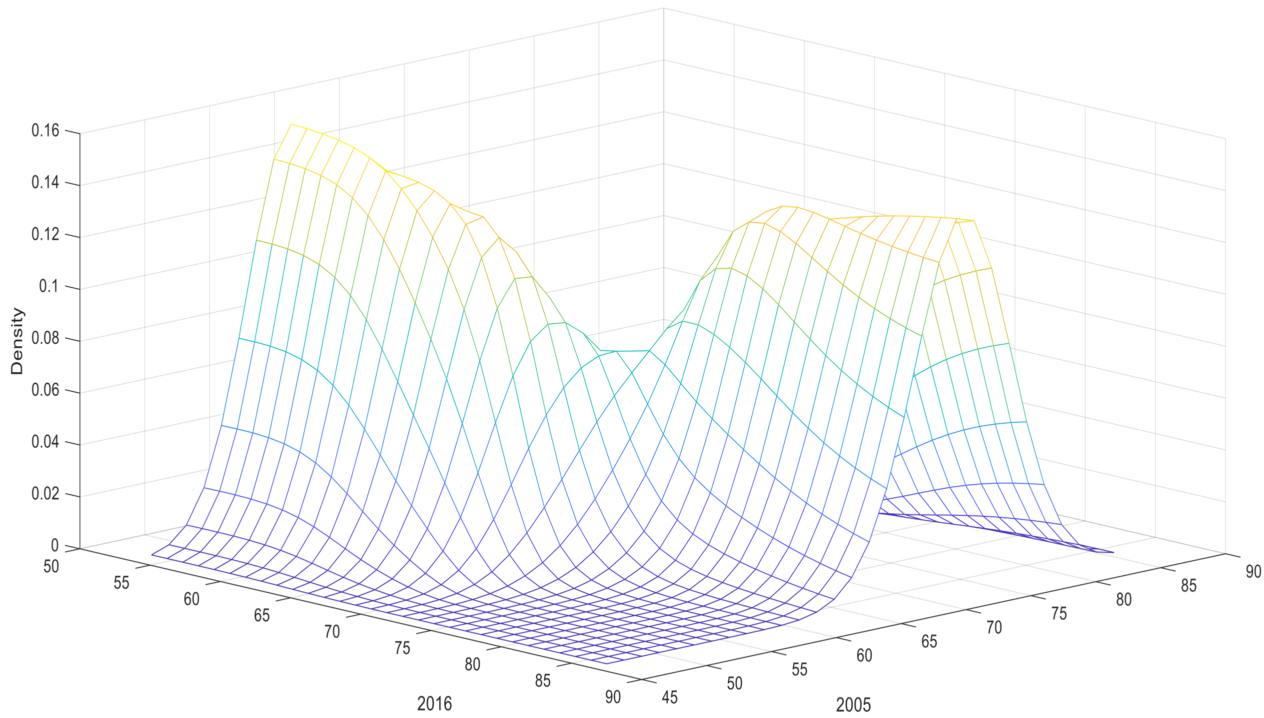
In 2016 the distribution of data related to the males' HLE presents few changes since it preserves the bimodality. This time, the distribution appears almost like the distribution of the females' HLE in 2005 presenting two peaks at 65 and 75 years with a mode at 75 years with the only difference, that is a higher frequency of countries that have the HLE at 65 years. The same occurs for the estimated distribution of the females' HLE. It is still bimodal, but it increases the probability to have larger values of HLE. The two peaks are allocated around 70 and 80 years with a higher frequency than those values registered before. So far we have analyzed the external shape of HLE of both genders for two years. We have seen that the distributions of both genders experiences fluctuations from the beginning of the period considered, but this doesn't say anything about movements of individual countries in the HLE distribution. However, for describing convergence it is important to have information on how units move within the distribution. Generally, a broad range of intra-distribution dynamics is possible, for example, over time there are some initially rich regions falling behind; poor regions overtaking the rich; and groups of regions, beginning at similar levels of development, eventually diverging (Quah, 1996)[29]. Properly, we are going to analyze intra-distribution

[29] In our study the terms "rich" and "poor" are referred to high and low levels of HLE.

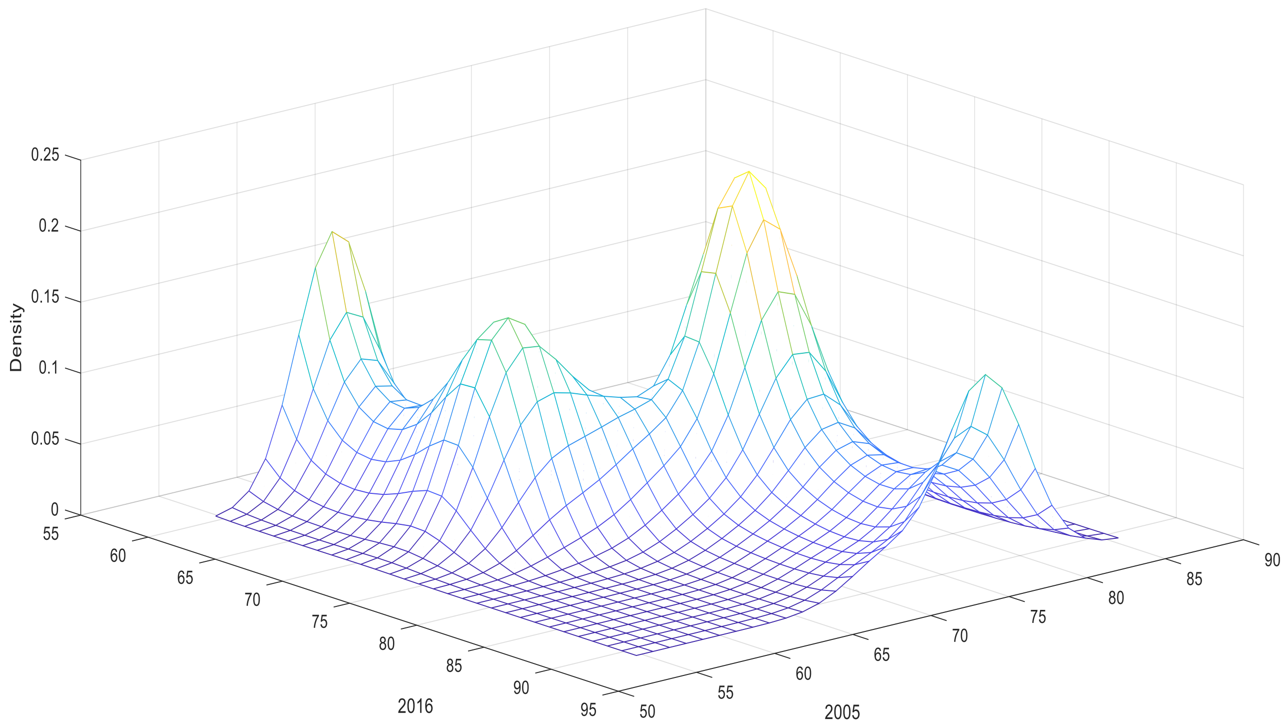
mobility by developing a probability model of transitions which captures the distribution's law of motion. We involved the stochastic kernel to examine how a given individual of the distribution at time t (e.g. 2005) transits to another part of the distribution by the time $t + s$ (e.g. 2016). It is estimated by the ratio between the joint distribution of starting and final values of HLE and the marginal density of y. The stochastic kernel related to the males' group illustrated in Figure 7a depicts a low frequency of those countries that in 2005 has HLE at 65 and in 2016 reach a level of HLE of 70. The mode of the distribution corresponds to those countries that remain around 57 years, while with not so much difference in terms of frequency we can distinguish other three areas of our distribution that represent three types of groups of countries: those that start from 57 and arrive at a level around 65 years, countries that in 2005 has an HLE of 70 years arriving around 75 in 2016, in the end, those countries that have an HLE around 75 in 2005 and finish to stay at an HLE level between 80 and 85 in the last year. The transitions from a state to another of female HLE are distributed differently. Figure 7b shows a majority of countries that remain stable around 77 years without experiencing fluctuations. In the second place, in terms of frequency, some countries are stuck on the starting level around 60 followed by the group of countries that increases their HLE from 64 in 2005 to 72 in 2016. Finally, they are few the countries that take place at 90 years beginning from an HLE of 77 years in 2005.

Figure 8.

a) The stochastic kernel of the males' HLE



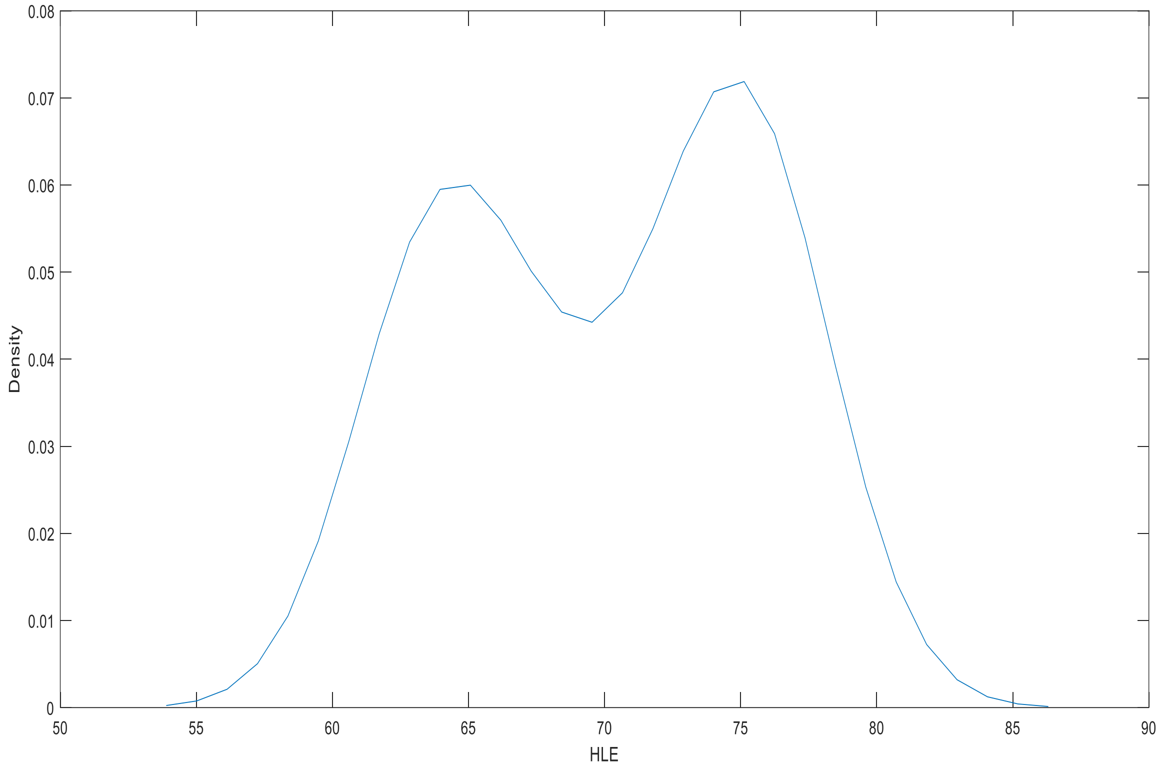
b) The stochastic kernel of the females’ HLE



According to conditional distributions estimated and Markov assumptions, we can forecast the future evolution of HLE. If we suppose that the law of motion is affected by regional policy we can interpret the time-invariance assumption by Markov as "unchanged" regional policy expenditures. By saying that, we saw in stochastic kernel in Figure 8a that the European countries tend to increase their level of HLE from 2005 to 2016 following different paths except for a part of countries that starts from an HLE value equals to 57 years, which continue to stay there. But since the transition probabilities are stable over years the same countries in the future can move upwards. Precisely, this is the phenomenon that happens in the long-run that we can see in Figure 8a.

In the same manner evolves the HLE of the female population. As you can see the stochastic kernel illustrates a majority of countries that are immobile at 77 years and a smaller probability to grow from a status to another, in particular from 64 to 72 and/or from 77 to 90.

Figure 9. the ergodic kernel of the males' HLE

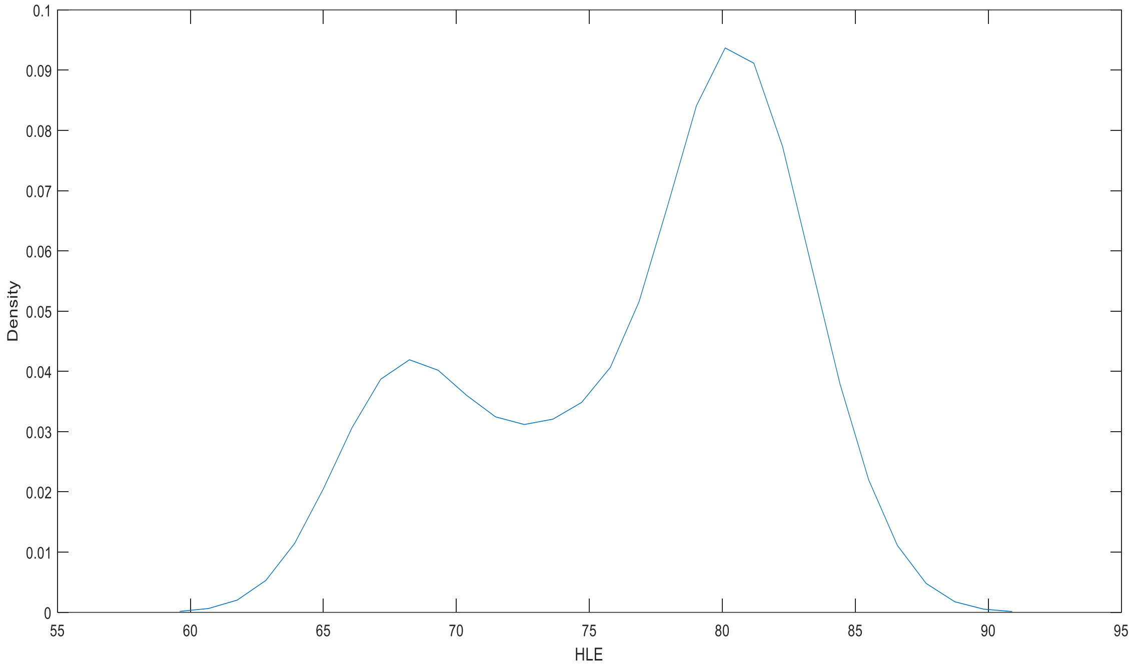


But repeating the evolution process more times in the future the convergence values can change. For instance, the countries that belong to the class of country that is stable at 77 years in the future can be part of that grows toward 90 years. The long-run distributions are depicted in Figures 9 and 10.

Watching the long-run distributions immediately we can see that they assume a bimodal shape confirming the convergence club hypothesis and specifically polarization. From the males' ergodic kernel emerged two distinct groups of countries.

In the future, we could face a wider group of countries with an HLE of 75 years and a smaller with 65 years. In the case of females, we distinguish a group of countries that has an HLE of 80 which corresponds to the mode of our distribution, and the other group that has an HLE around 68.

Figure 10. The ergodic kernel of the females' HLE



5. Conclusion

The main goal of our research has been to see how the HLE of European countries developed from 2004 to 2016. While we used the regression-based approach by Barro and Sala-i-Martin to analyze the catching-up process between countries with smaller and larger initial values, on the other side we involved the dynamic approach formulated by Quah to see how HLE in Europe evolves focusing on the distributions estimated. According to the first model, the HLE growth rate seems to be higher for countries that start from a lower level. Although this lead to an absolute convergence hypothesis, the results showed exactly the opposite. The test of homogeneity of variances showed that the variances estimated don't present a difference statistically significant across years.

Since the results had excluded the case of absolute convergence, what remained to see was if about some distinctive traits the European countries will converge to different steady states. The long-run distributions of HLE of males and females exhibit a bimodality that reveals a higher frequency at 75 years for males and 80 for females, and a less frequency around 65 and 68 years for males and females respectively. Furthermore, we see that the mode evidenced in the females' distributions is higher than that evidenced in the males' distribution counterposed with a higher frequency for the lowest. Seeing the stochastic kernel distributions of both genders, this is justified from a greater probability of males estimated for countries that remain fixed at 57 years together with those countries that from 57 moves on around 70. Indeed, in Figure 8b the probability to continue to have an HLE around 60 years in 2016 is less than to have HLE values superior to 70. The ergodic kernel distributions were estimated assuming the time invariance hypothesis by Markov. To validate this assumption we estimated the conditional distributions of both genders of HLE data related to years 2010 with data recorded in 2005 and those related to data of 2016 conditioned with data of 2010. Both of the distributions related to male data are quite similar to that disclosed before, instead, the distributions related to female data present a shape of the distribution different from that we saw before. Then the assumption, although it doesn't represent reality, is fundamental for the realization of our predictions.

The main result of this approach is the rejection of the hypothesis of the occurrence of the catching-up process between countries that have a lower HLE with countries that have a higher HLE. From the stochastic kernel of males and females, we saw that countries with a lower initial value of HLE may not even grow. They can remain stable at their initial level. Moreover, the countries that start with an HLE superior to 70 may increase much more than countries that have an initial value inferior to 60. In our model, such kind of growth is not justified by specific factors that allow us to describe why the European countries converge toward different steady states. It is an interesting origin for another study that would broaden the views on the convergence of HLEs in Europe to explain the determinants that bring toward a specific steady state[30].

[30] Such kind of methodology was involved by by Quah(1996a), "Convergence empirics across economies with (some) capital mobility," Journal of Economic Growth, 1, 95–124 and Silvia Dal Bianco(2010) that followed his procedure.

6. References

[1] Barro Robert J. and Xavier Sala-i-Martin. 1992. "Convergence". Journal of Political Economy 100(2): 223-251.

[2] Baumol, W. (1986): "Productivity growth, convergence and welfare: what the long run data show," American Economic Review, 76(5), 1072–85.

[3] John Creedy (2015). "A Note on Computing the Gini Inequality Measure with Weighted Data". Victoria Business School.

[4] Dal Bianco Silvia (2010), "Going Clubbing in the Eighties: Convergence in Manufacturing Sectors at a Glance", University of Pavia.

[5]Srinivas Goli and Perianayagam Arokiasamy."Trends in health and health inequalities among major states of India: assessing progress through convergence models Health Economics", Policy and Law / FirstView Article / February 2013, pp 1 26.

[6] Goli S., Moradhvaj, Chakravorty S., Rammohan A. (2019) World health status 1950-2015: Converging or diverging. PLoS ONE 14(3): e0213139. https://doi.org/10.1371/journal.pone.0213139.

[7]Osayomor Ikpotokin(2019)." Asymptotic mean integrated squared error's reduction techniques in Kernel density estimation". International Journal of Computational and Theoretical Statistics, May 2019.

[8] Falko Juessen(2005). "A distribution dynamics approach to regional income convergence in reunified Germany". University of Dortmund, Germany.

[9]Sabyasachi Kar; Debajit Jha; Alpana Kateja (2011). "Club-convergence and polarization of states. A nonparametric analysis of post-reform India". Indian Growth and Development Review :(PDF) Club-convergence and polarization of states A nonparametric analysis of post-reform India (researchgate.net)

[10] Levene, H. 1960. "Robust tests for equality of variances. In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling", ed. I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, and H. B. Mann, 278–292. Menlo Park, CA: Stanford University Press.

[11] Lerman, R.I. and Yitzhaki, S. (1989) Improving the accuracy of estimates of Gini coefficients. Journal of Econometrics, 42, pp. 43-47.

[12] Wendy L. Martinez Angel R. Martinez (2002). "Computational Statistics Handbook with MATLAB". CHAPMAN & HALL/CRC Boca Raton London New York Washington, D.C.

[13] Moser, K., V. Shkolnikov and D. A. Leon (2004), "World mortality 1950–2000: divergence replaces convergence from the later 1980s", Bulletin of the World Health Organization, 83(3): 202–209.

[14] E. Parzen(1962). " On estimation of a probability density function and mode". Standford University.

[15] Danny Quah(1993). "Galton's Fallacy and Tests of the Convergence Hypothesis". The Scandinavian Journal of Economics, Vol. 95, No. 4, Endogenous Growth (Dec., 1993), pp. 427-443.

[16] Danny T. Quah(1996)."Empirics for economic growth and convergence". European Economic Review 4(3 (1996) 1353-1375.

[17] Roystone,1991d. Estimating departure from normality. Statistics in Medicine 10: 1283–1293.

[18] Xavier Sala-i-Martin(1996). "Regional cohesion: Evidence and theories of regional growth and convergence". European Economic Review.

[19] David W. Scott ."On Optimal and Data-Based Histograms", Biometrika, Vol. 66, No. 3 (Dec., 1979), pp. 605-610;

[20] David W. Scott(2015) "Multivariate density estimation : Theory, practice and visualization" – John Wiley and Sons, Inc.; New Jersey.

[21] Shapiro, S. S.; *Wilk, M. B.* (1965). "An analysis of variance test for normality (complete samples)". *Biometrika*. 52 (3–4): 591–61.

[22] Solow, R., 1956, "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics* 70 (1), 65-94.

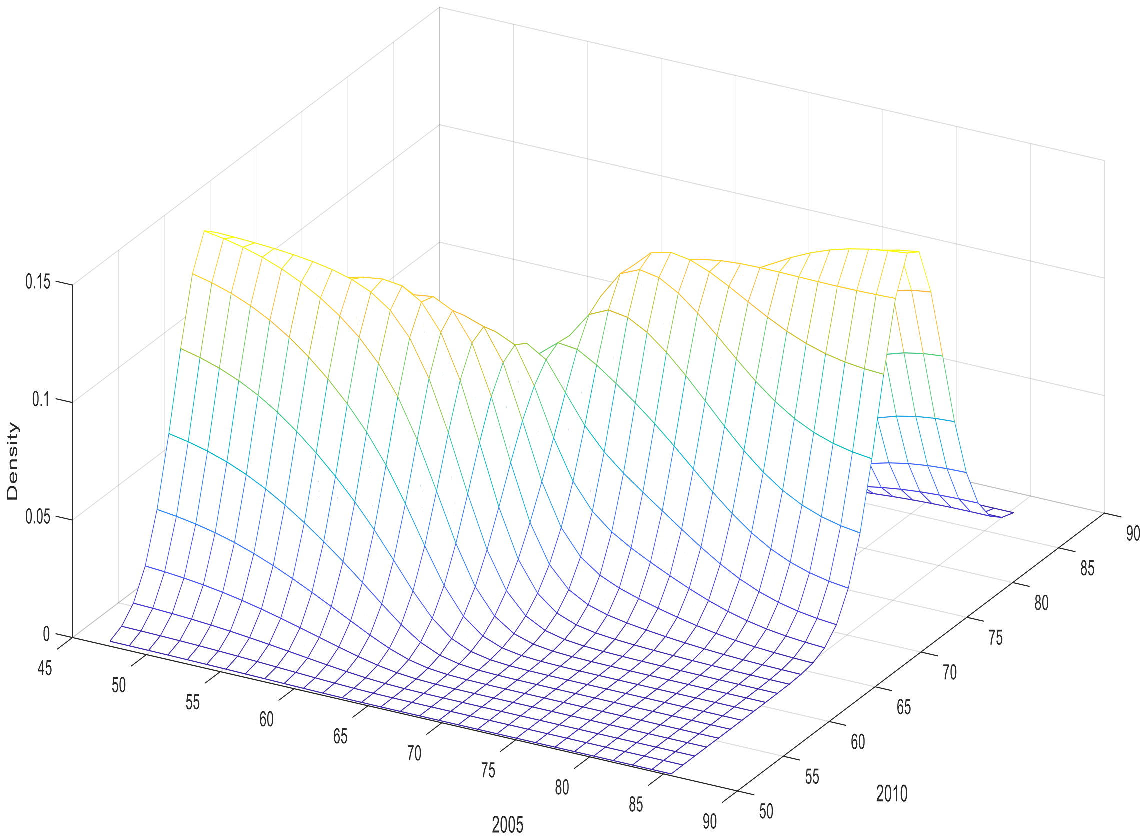
[23] White, H.: Asymptotic Theory for Econometricians. Academic Press, New York, 1984.

[24] Andrew T. Young. "Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data". Journal of Money, Credit and Banking, 2008

7. Appendix

Figure 11.

a) The distribution of HLE related to male data in 2010 conditioned to HLE data of 2005.



b) The distribution of HLE related to male data in 2016 conditioned to HLE data of 2010.

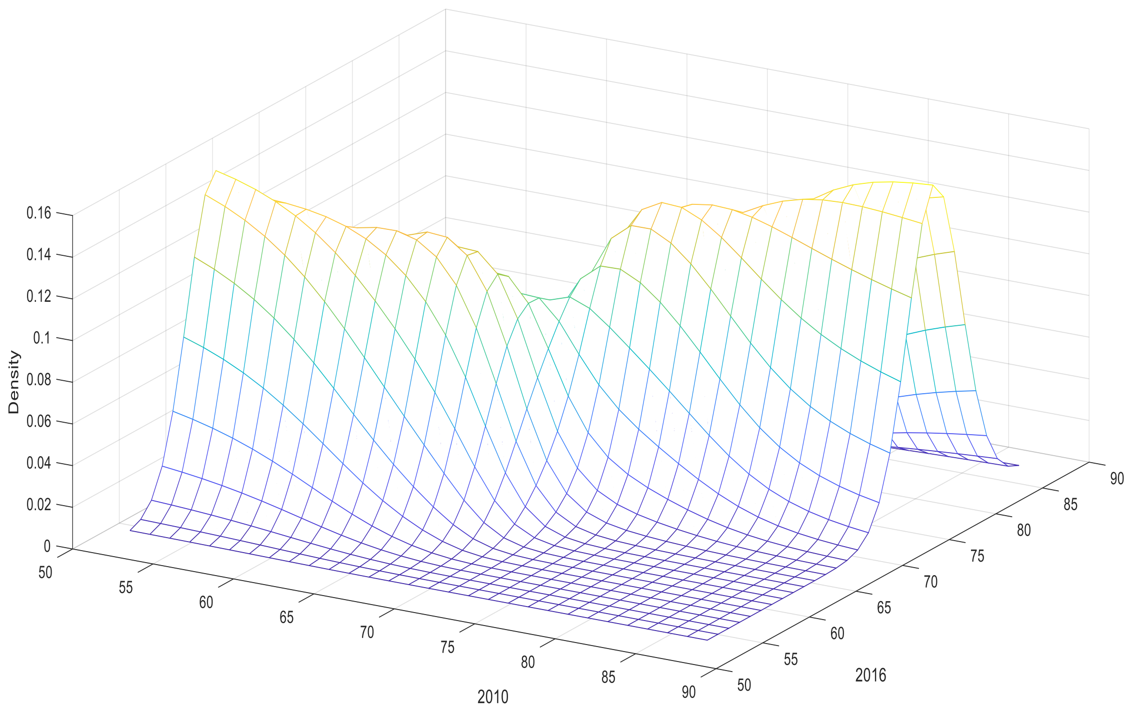
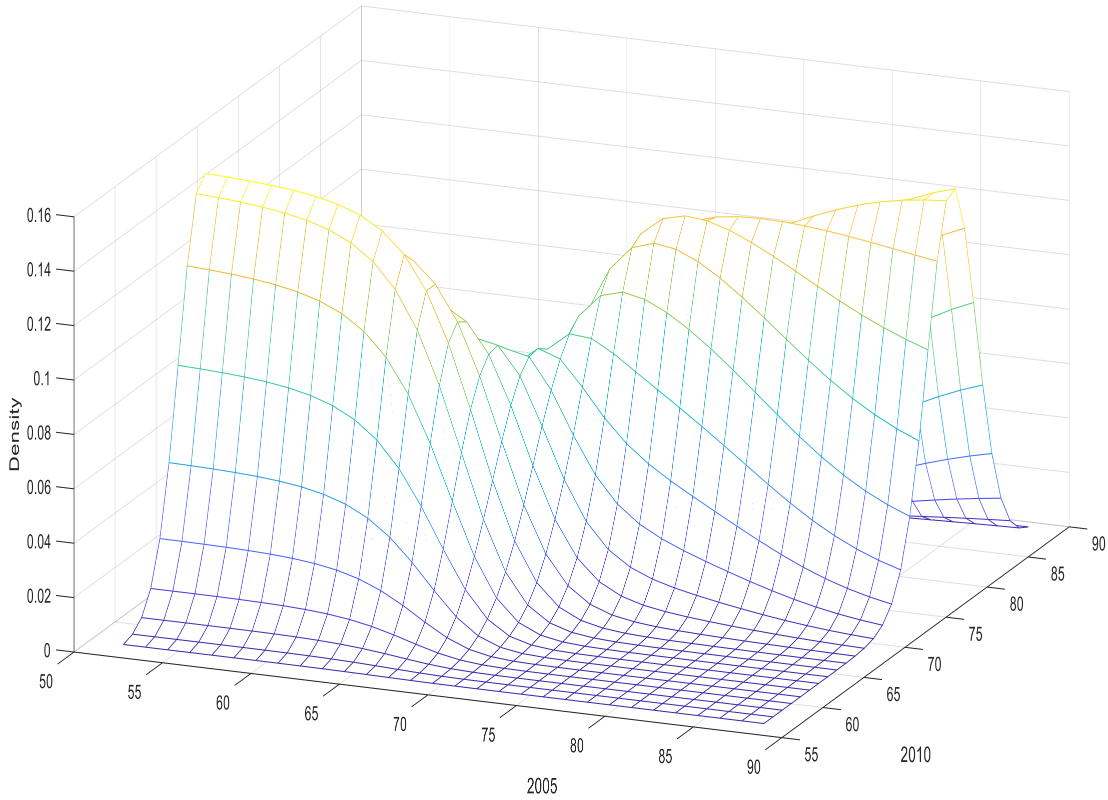


Figure 12.

a) The distribution of HLE related to female data in 2010 conditioned to data of 2005.



b) The distribution of HLE related to female data in 2016 conditioned to data of 2010.