GRADO EN ECONOMÍA

CURSO ACADÉMICO 2019/2020

TRABAJO FIN DE GRADO

CONVERGENCIA ESPACIAL DE LA ESPERANZA DE VIDA: UN ANÁLISIS EMPÍRICO

SPATIAL CONVERGENCE OF LIFE EXPECTANCY: AN EMPIRICAL ANALYSIS

AUTORA:
IRENE GONZÁLEZ RODRÍGUEZ

DIRECTORES:
DAVID CANTARERO PRIETO
MARTA PASCUAL SÁEZ

8 DE JULIO DE 2020

INDEX

LIST OF TABLES

LIST OF FIGURES

Resumen

El principal objetivo de este trabajo es analizar la convergencia en salud entre los países de la Unión Europea y las provincias españolas, teniendo en cuenta las interacciones espaciales entre unos territorios y otros. El estado de salud se cuantifica a través de la esperanza de vida en cuatro de sus formas: la esperanza de vida al nacer, la esperanza de vida a los 65 años, la esperanza de vida ajustada por discapacidad y la esperanza de vida saludable basada en la salud percibida por uno mismo. El periodo analizado difiere de unos indicadores a otros, estando comprendido en su totalidad entre 1998-2018. Se estudian dos tipos de convergencia: la σ-convergencia, a través de la desviación estándar, y la β-convergencia absoluta, a través de la estimación de una regresión determinada. Adicionalmente, se realiza el Test de Moran con el fin de conocer el grado de dependencias espaciales existentes y se usan dos modelos de regresión espaciales (Modelo de Retardo Espacial y Modelo del Error Espacial) para incluir dichas dependencias en la estimación de la β-convergencia absoluta. Los resultados sugieren que existe un proceso de convergencia, tanto en σ como en β, entre los países de la Unión Europea y entre las provincias españolas. Por tanto, la dispersión se reduce y las tasas de crecimiento han sido mayores en los lugares que tenían menores valores de esperanza de vida al inicio del periodo analizado. Por otro lado, la β-convergencia es mayor en las mujeres entre los países europeos y en los hombres entre las provincias españolas, independientemente del modelo seleccionado. Sin embargo, el aumento de la dispersión en los últimos años hace que no haya una evidencia clara sobre su tendencia en el próximo periodo. Este estudio tiene una contribución doble. En primer lugar, aporta una visión actualizada de la convergencia en salud en los últimos años. En segundo lugar, contribuye a la incorporación de la econometría espacial en la investigación de la Economía de la Salud. Nuestro análisis se puede considerar para aplicar medidas de intervención públicas cuya finalidad sean reducir las desigualdades en salud de la población.

Palabras clave: econometría espacial de la salud, convergencia, esperanza de vida, Unión Europea.

Abstract

The main objective of this study is to analyse the health convergence between the European Union countries and between the Spanish provinces, considering the spatial interactions among several territories. Health status is quantified through life expectancy in four of its forms: life expectancy at birth, life expectancy at age 65, disability-adjusted life expectancy and healthy life expectancy based on self-perceived health. The period analysed differs from one indicator to another, being in its entirety between 1998-2018. Two types of convergence are studied: σ-convergence, through standard deviation, and absolute β-convergence, estimating a determined regression. Furthermore, the Moran Test is performed in order to know the degree of existing spatial dependencies and two models of spatial regression (Spatial Lag Model and Spatial Error Model) are used to include these dependencies in the absolute β-convergence estimation. The results suggest that there is a convergence process, both in σ and β, between the European Union countries and between the Spanish provinces. Therefore, dispersion is reduced and growth rates have been higher in the places where had lower life expectancy values at the beginning of the analyzed period. On the other hand, β-convergence is higher in women among European countries and in men among Spanish provinces, whether we consider the model without spatial dependencies or the models which include them. Nevertheless, the rise of the dispersion in recent years means that there is no clear evidence about its trend in the following period. This study has a double contribution. First of all, it provides an updated vision of the health convergence in the last years. Secondly, it contributes to the incorporation of spatial econometrics in Health Economics research. Our analysis can be considered in order to apply policy intervention measures to reduce health population inequalities.

Keywords: spatial health econometrics, convergence, life expectancy, European Union.

1 INTRODUCTION

The increasing life expectancy of the older population in many countries makes ageing a key issue in today's society. In this regard, one of the Governments objectives is to improve the quality and duration of population's lives. The state of people's health has traditionally been measured by life expectancy and mortality indicators. However, in recent years, other perspectives have been included and taken into account: quality of life, socio-economic environment, disease prevention, freedom from disability, etc. Despite the fact that people's health quality increases over time, disparities continue to exist in many ways. On one hand, inequalities may exist between different countries. For example, West European Union (EU) countries have longer life expectancies than the East ones. On the other hand, this kind of differences may also be among regions within the same country (Forster et al., 2018).

The health convergence analysis allows to approximate the evolution of those inequalities among different territories. Convergence is understood as *"the reduction or equalizing of disparities"* (Maynou et al., 2015). Although Solow (1956) advanced the convergence hypothesis in his growth theory, the most famous studies are given by Baumol (1986) and Barro and Sala-i-Martin (1990, 1992, 1995). They analyzed income approximation processes between regions or countries. Inside the literature of growth empirics, there are different perspectives for measuring convergence. A significant number of papers broadly categorized it into two parts: σ-convergence and β-convergence. The first one occurs when the dispersion between the studied regions decreases, while the second one exists when the areas that have lower values of the selected indicator at the beginning of the period have the final highest growth rates. Most convergence analysis have been based on income indicators. However, there aren't as many studies that analyze convergence regarding health inequalities due to the complexity of measuring people's health (Hembram and Haldar, 2020). According to the World Health Organization (WHO) Constitution (1946), having good health is *"a state of complete physical, social and mental well-being, and not merely the absence of disease or infirmity"* (Modranka and Suchecka, 2014). In this sense, the last half of the 20th century has been characterized as a few decades of world health convergence (Clark, 2011).

To deepen these convergence analyses, spatial econometrics have been used in recent years. The beginnings of spatial econometrics date back to the 1970s, when Jean Paelinck defined it as the part of applied econometrics that deals with estimation and specification problems arising from traditional econometrics (Anselin and Bera, 1988). In other words, spatial econometrics serves as a tool that corrects the spatial problems of standard econometric models. By using these models, we have the possibility to take into account the spatial dependence between observations, which often arises if data are collected on regions or countries located in space. In addition, these models are useful for assessing and quantifying the spillover effects in different geographical entities (Lesage, 2008).

In this context, the two most important limitations of not considering space are spatial autocorrelation and spatial heteroscedasticity (González, 2016). The first one occurs when the value of a certain variable in a specific spatial unit is related to its value in one or more other locations. Therefore, an observation associated with a given spatial location i, depends on others linked to locations j (where i is different from j) (Bohórquez and Ceballos, 2008). This concept is multidirectional, since a spatial unit may not be affected only by another neighboring region or country, but it may be affected by many others that surround it (Moreno et al., 2002). Spatial autocorrelation can be expressed also as spatial dependency. The main difference is that the first one refers simultaneously to a phenomenon and a statistical technique while the second one refers only to

a theoretical explanation (Vilalta, 2005). Nevertheless, in this study, as in the majority, both concepts will be used as synonyms. This dependency may be caused mainly by two reasons: by the existence of measurement errors and by the existence of spatial interaction, overflow effects and spatial hierarchies. In turn, spatial autocorrelation can be positive or negative. When the presence of a certain phenomenon in a specific spatial unit makes it easily spread to its neighboring localities favoring its grouping, spatial autocorrelation will be positive. Meanwhile, when the presence of this factor hinders its propagation to neighboring towns, regions or countries, spatial autocorrelation will be negative (González, 2016). Secondly, spatial heterogeneity refers to the relationships variations in space. Usually, two aspects of this concept can be distinguished: structural instability and heteroscedasticity. On one side, structural instability is understood as the absence of stability that exists in the space of the behavior of the study variable. This means that the functional form and the parameters of a regression may change from place to place and, therefore, they will not be homogeneous throughout the sample. On the other side, heteroscedasticity appears when specification errors occur, which give rise to the appearance of measurement errors (Moreno et al., 2002). The vast majority of problems caused by spatial heterogeneity can be solved by standard econometric techniques. For this reason, the spatial literature usually focuses on the analysis of spatial autocorrelation (so the same will be done in this study).

In recent years, a new branch of spatial econometrics has emerged that focuses on the health analysis: the Spatial Health Econometrics. It measures spatial relationships between agents at both the micro and macro levels, while testing health economic theories (Baltagi et al., 2018). It's a good tool to reduce the degree of uncertainty about the complexity of the issues addressed by Health Economics. In this regard, the following paragraph review the most recent literature on the use of spatial econometrics in health convergence among European countries.

Jaworska (2014) studies life expectancy convergence in the European Union NUTS II[1] regions between 2002 and 2012. The results show that regions with the lowest life expectancy in 2002 have experienced the greatest increases. However, it highlights that the process hasn't been the same for all regions, as there is higher speed of convergence in the more developed countries (South-West Europe). In addition, Stańczyk (2015) analyzes the same as Jaworska (2014) and draws a complementary conclusion: socioeconomic factors have a major impact on the public health state. Other authors quantify these inequalities by including more variables. This is the case of Mayou et al. (2015), who study the speed of convergence between the mortality rate and life expectancy at birth between 1995 and 2009. The results suggest that the dispersion between this two indicators isn't decreasing. Considering the reduction of dispersion as a measure of convergence, this article is the first one who expose a lack of health convergence in all the EU regions. Maynou and Saez (2016) complement the previous study by estimating the impact of the 2008 economic crisis on the health convergence process in the European Union. Their results show that there is indeed a recovery process among the different EU regions, but that there is also no reduction in the dispersion levels. On the other hand, there are several authors who carry out this type of analysis, but focus on the regions of a particular country. Eibich and Ziebarth (2014) estimate spatial health inequalities at the county level in Germany. They find large and significant spatial dependencies and groupings, and an East-West spatial health pattern. Zeren et al. (2016) study the health care service impartiality at the county level in Turkey between 2002 and 2010. They use the convergence hypothesis to test whether inequalities are reduced and draw a double con-

[1]NUTS level 2: includes the Spanish autonomous regions, French regions and overseas departments, Polish Voivodships, etc. This term was created by the European Statistics Office to have an equal statistical standard across the EU (Eurostat, 2020d).

clusion: there is province convergence and there is health care service spatial interaction between regions. Finally, Dearden et al. (2019) show the relationship between health and geography based on health inequalities in Great Britain in 1991, 2001 and 2011. Thanks to a growing positive spatial association, neighbouring areas increasingly share similarities and there is less distinction between areas characterised by good and bad health. It should be noted that one of the reasons for the existence of this health inequalities is the differences between regions in terms of factors determining the population's health status. The most prominent are the level of environmental pollution, the wealth of society and the access to health services. Modranka and Suchecka (2014) analyze these determinants using a Spatial Durbin Model (SMD). They suggest specific factors that generate positive impacts on the general health status of inhabitants of adjacent regions, such as the pharmacy locations or the number of people who go to the gym.

The main objective of this study is to quantify the health convergence in the last years of the 20th century and first decades of the 21st. Reviewing the health convergence studies, it can be seen how some of them measure convergence between countries while others focus on regions within a single country. However, in this case, both types of research will be done. Firstly, health convergence between the European Union countries will be analyzed. Secondly, the same study will be done, but focused on the Spanish provinces. Health status is quantified through life expectancy in four of its forms: life expectancy at birth, life expectancy at age 65, disability-adjusted life expectancy and healthy life expectancy based on self-perceived health. Two types of convergence will be studied: σ and absolute β-convergence. As the data to be used refer to regions and countries, it's important to analyse whether there is any spatial relationship between them. To take into account these possible spatial dependencies, spatial econometrics is incorporated. In this sense, Moran's Test will be used to find out if there is spatial dependence in the observations. If there is, the Spatial Lag Model and the Spatial Error Model will be used to estimate spatial absolute β-convergence. Additionally, the analysis is separated between men and women to show gender differences.

This study will be structured as follows. After this introduction, the data to be used will be described. Next, the methodology of the two types of convergence studied plus that of spatial analysis will be showed. The main results are then presented and discussed to finally introduce the conclusions.

2 DATA DESCRIPTION

As previously discussed, this study focuses on life expectancy spatial convergence analysis. To quantify health status with various dimensions and to give a broad vision of the life expectancy concept, four variables have been chosen: life expectancy at birth, life expectancy at age 65, disability-adjusted life expectancy and healthy life expectancy based on self-perceived health. The study will be divided into two cases: the European Union countries and the Spanish provinces convergence analysis. For the first case, all 27 EU countries have been considered, in order to update the existing research on these countries. In the second case, the 50 Spanish provinces and the two autonomous cities (Ceuta and Melilla) will be analysed. Spain has been chosen for the detailed analysis because it's one of the countries with the highest life expectancy value in recent years. Provinces data will be used to take into account the spatial dependencies within the territory of this country. Furthermore, to make a more detailed research, a distinction will be made between men and women, in all cases. Table 2.1 shows the selected variables, its abbreviation, the time period and the data extraction source in each case. It should be noted that the source of the European spatial data is Eurostat (2020a) and that of the Spanish data is GADM (2020).

Table 2.1: Selected variables

Variable	**Abbreviation**	**Temporary period EU / Spain**	**Source EU / Spain**
Life expectancy at birth	LEB	1998-2018 / 1998-2018	World Bank Open Data / INE
Life expectancy at age 65	LE65	2005-2018 / 1998-2018	Eurostat / INE
Disability-adjusted life expectancy	DALE	2000-2015 / -	WHO Database / -
Healthy life expectancy based on self-perceived health	HLESPH	2006-2016 / -	Eurostat / -

Source: Authors' elaboration

Firstly, life expectancy at birth (LEB) will be analyzed in 1998-2018, in both cases. In the European analysis, data has been extracted from the World Bank Open Data (2020), who defines life expectancy at birth as the number of years a newborn baby would live if current mortality rates don't change throughout their lives. In the Spanish research, data has been extracted from the Spanish Statistical Office (INE) (2020b). Secondly, life expectancy at age 65 (LE65) will be studied between 2005-2018 for the European Union countries and between 1998-2018 for the Spanish provinces. In the first case, the data extraction source is Eurostat (2020c), who understands this indicator as the average number of years that a person has to live when he turns 65 if the rest of his life is subject to current mortality rates. In the second case, the data source is the INE (2020a). The third variable selected is disability-adjusted life expectancy (DALE). It's an indicator that measures the difference between current health status and an ideal health situation in which all individuals live to a high age, without disease and disabilities (WHO, 2020). It will be studied between 2000-2015 and the data source is the World Health Organization Database (2020). Finally, the healthy life expectancy based on self-perceived health (HLESPH) variable will be used, which combines self-perceived health data with mortality data. It's shows the number of years that a person expects to live in good self-perceived health (Eurostat, 2020b). The time period considered will be between 2006-2016 and the data source is Eurostat (2020b). These last two variables will only be incorporated in the European analysis [2].

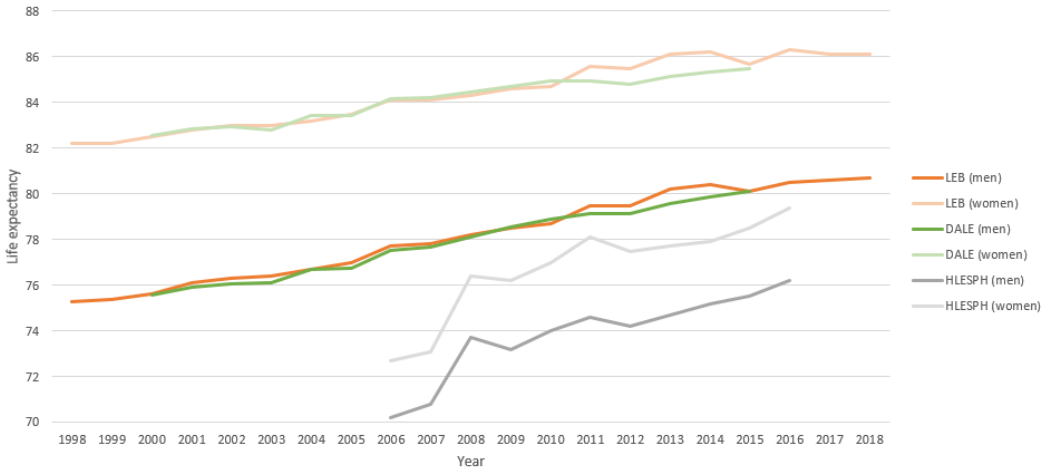
A small data pre-processing has been carried out in order to correct their inconsistencies and transform them intro simpler formats (Hernández et al., 2008). The main problem to correct in this study is the treatment of not available data (NaNs). The Roma-

[2]DALE and HLESPH are only incorporated in the European analysis due to the lack of Spanish provinces data.

nian and Croatian time series for healthy life expectancy based on self-perceied health variable had 1 (year 2006) and 4 (years 2006-2009) NaNs, respectively. To resolve this, an attempt has been made to capture the trend through a rolling window [3]. The accumulated growth rate for the following 7 years of the NaN has been calculated and the result has been extrapolated.

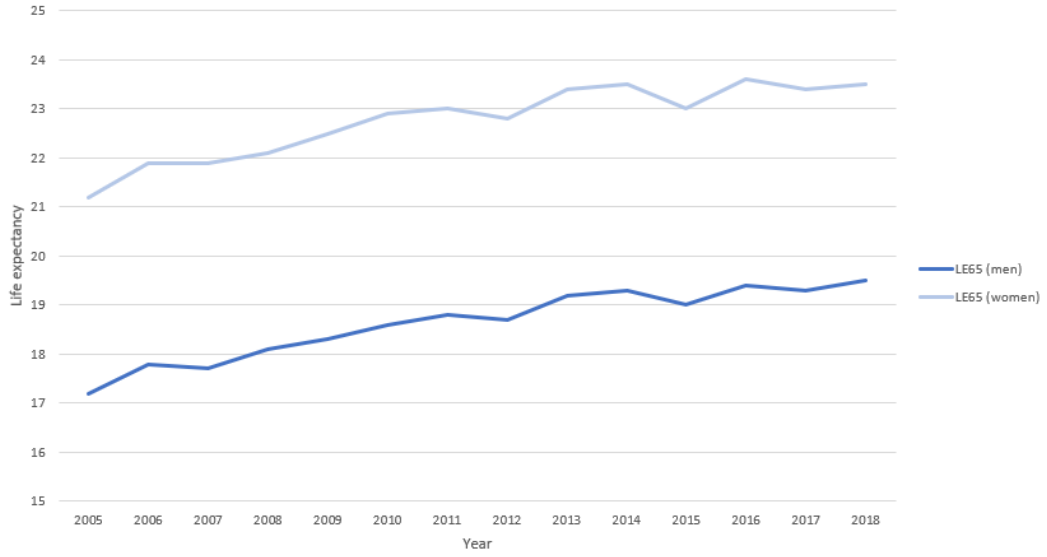
The evolution of each variable has been different depending on the country or province selected. Figures 2.1 and 2.2 show the Spanish four variables evolution [4]. It's exemplified with this country because it will be examined in depth in the second part of this study.

Figure 2.1: Evolution of Spanish life expectancy (LEB, DALE, HLESPH)



Source: Authors' elaboration from World Bank Open Data (2020), WHO (2020) and Eurostat (2020b) data

Figure 2.2: Evolution of Spanish life expectancy (LE65)



Source: Authors' elaboration from Eurostat data (2020c)

[3] Rolling window: window of consecutive observations (Li et al., 2014).
[4] For reasons of word limitation, only one example will be presented.

As we can see, all variables have an increasing trend. Moreover, in all cases, the final values are greater than the initial values and the curves for women are higher than for men, that is, women have a longer life expectancy. The evolution of life expectancy at birth and of disability-adjusted life expectancy are very similar. In the first years, there's practically a curve on top of the other one, but from 2010 they begin to diverge. Furthermore, these two variables growth is higher for men than for women. On the other hand, the healthy life expectancy based on self-perceived health curves have a much greater slope: they grow more than the rest in a shorter period of time. However, they have bigger variability, since they have more notable streaks. Growth for men has been practically the same as for women. Lastly, the life expectancy at age 65 evolution doesn't show great streaks or big differences between men and women.

Table 2.2 shows the main statistics of the previous variables: the minimum and maximum, the first and third quartiles, the median and the mean for the initial and final study year of each variable.

Table 2.2: European Union countries summary statistics (men / women)

Main statistics (men / women)	**LEB**		**LE65**	
Year	1998	2018	2005	2018
Min.	63.50 / 73.70	69.80 / 78.40	12.30 / 15.90	14.10 / 18
1st Qu.	68.60 / 76.70	74.10 / 81.60	13.55 / 17.80	15.70 / 19.75
Median	73.52 / 79.40	78.73 / 83.40	16.30 / 19.40	18.20 / 21.60
Mean	71.69 / 78.82	77.20 / 82.94	15.49 / 19.13	17.43 / 20.97
3rd Qu.	74.95 / 80.70	79.75 / 84.40	16.85 / 20.30	19.10 / 21.95
Max.	76.90 / 82.60	81 / 86.10	17.70 / 22	19.70 / 23.80
Summary statistics (men / women)	**DALE**		**HLESPH**	
Year	2000	2015	2006	2016
Min.	64.38 / 74.27	68.12 / 77.98	57.70 / 62.7	62.20 / 67.70
1st Qu.	69.22 / 77.58	73.30 / 81.25	61.30 / 65.05	66.55 / 71.30
Median	74.02 / 79.71	78.31 / 83.45	70.20 / 72.70	73.60 / 75.60
Mean	72.28 / 79.21	76.59 / 82.45	67.60 / 70.95	71.59 / 75.15
3rd Qu.	75.20 / 80.92	79.39 / 83.86	72 / 75	75.50 / 78.85
Max.	77.26 / 82.67	80.68 / 85.46	75.60 / 79.40	78 / 81.30

Source: Authors' estimations from World Bank Open Data (2020), WHO (2020), Eurostat (2020b) and Eurostat (2020c) data

The most relevant and common characteristic at all levels is that the statistics, both in the different years and in the different variables, are higher for women than for men. This shows us in advance the existence of a gender life expectancy inequality. In the case of men, Sweden is the country that reaches the maximum in both the initial and final years for the life expectancy at birth and for disability-adjusted life expectancy. In the case of life expectancy at age 65, it's France who has the maximum values, for both men and women. The country that reaches the minimum value in all variables for men is Latvia, except for disability-adjusted life expectancy in 2015 and for healthy life expectancy based on self-perceived health in 2016, which is Lithuania. In another sense, France, Spain, Ireland and Malta are the countries that have the maximum values for women. Latvia, Romania, Bulgaria and Croatia have the lowest values. These results show that there could be a spatial distinction between Western and Eastern Europe countries.

At the same time, Table 2.3 shows the main statistics of the Spanish provinces sample.

Table 2.3: Spanish provinces summary statistics (men / women)

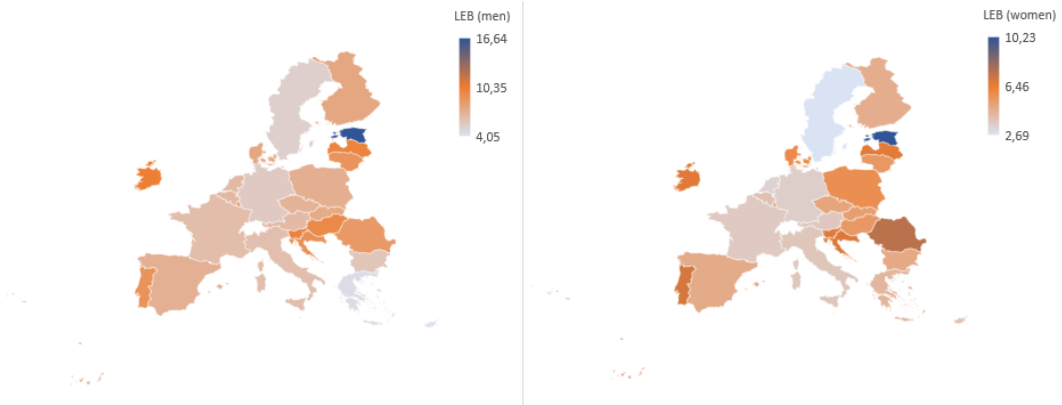
Summary statistics (men / women)	**LEB**		**LE65**	
Year	1998	2018	1998	2018
Min.	72.43 / 78.07	78.23 / 82.13	14.71 / 17.98	17.67 / 20.32
1st Qu.	74.83 / 81.58	79.69 / 85.14	15.74 / 19.51	18.66 / 22.31
Median	75.51 / 82.54	80.35 / 85.88	16.39 / 20.40	19.24 / 23.11
Mean	75.62 / 82.27	80.34 / 85.63	16.43 / 20.26	19.15 / 22.89
3rd Qu.	76.33 / 83.20	80.93 / 86.30	17.07 / 21.02	19.59 / 23.55
Max.	79.21 / 84.03	82.18 / 87.29	19.11 / 21.81	20.78 / 24.34

Source: Authors' estimations from INE (2020a) and INE (2020b) data

As in the previous case, the most important characteristic is that all values are bigger in women than in men. Comparing 1998 with 2018, the mean of both variables has increased more in men than in women. Moreover, there is a clear spatial distinction between the provinces that have the maximum and minimum values. On the one hand, the maximum values in both genders are reached in northern provinces. They are all from Castile and León (Soria, Salamanca and Segovia) except for the province of Álava, which occupies the first position in life expectancy at age 65 for women. On the other hand, the minimum values are occupied by places located in the south of Spain: the two autonomous cities (Ceuta and Melilla) and the province of Cádiz.

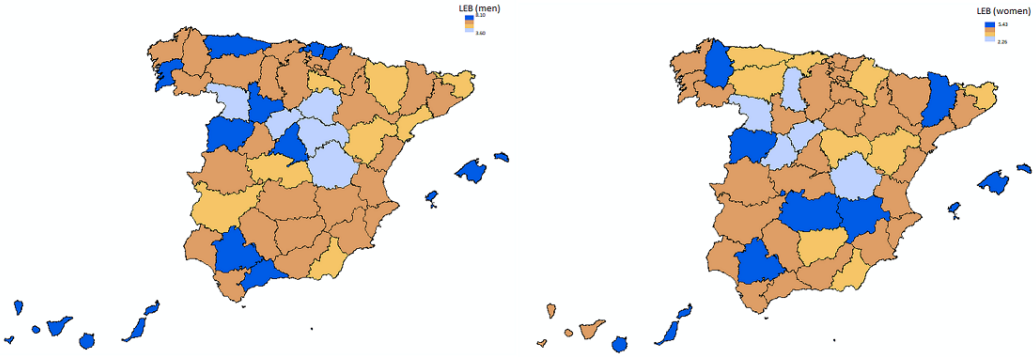
Finally, Figures 2.3 and 2.4 show, by way of example, the life expectancy at birth growth rate (measured in percentage) for men and women in the European Union countries and in the Spanish provinces. As can be seen, neighboring countries tend to have close values. Within the European Union, most Eastern countries have had higher growth rates than those in the West (excluding Portugal and Ireland). This may be a symptom of spatial dependency or autocorrelation.

Figure 2.3: Gap life expectancy at birth EU 1998-2018 (men / women)



Source: Authors' elaboration from World Bank Open Data (2020)

Figure 2.4: Gap life expectancy at birth Spain 1998-2018 (men / women)



Source: Authors' elaboration from INE (2020b) data

3 METHODOLOGY

The objective of this section is to describe the main tools that will be considered for the empirical analysis. First of all, the methodology corresponding to the σ-convergence and absolute β-convergence analysis will be presented because inside the literature of growth empirics this notion of convergence is broadly categorized into two parts: σ-convergence and β-convergence. The first type of convergence is estimated with the calculation of the standard deviation, while a particular regression model is estimated for the second type of convergence through Ordinary Least Squares (OLS). However, if the sample studied has spatial dependencies, the results of this model may not be valid (Moreno et al., 2002). That is why, secondly, the methodology referring to the spatial data analysis will be presented. The Moran Test will be carried out to determine whether or not there is spatial autocorrelation. If there is, the β-convergence will be re-estimated, but this time using two spatial regression models called Spatial Lag Model (SLM) and Spatial Error Model (SEM) through Maximum Likelihood (MV).

3.1 CONVERGENCE

3.1.1 Sigma-convergence

This type of convergence ocurrs when *"the dispersion, measured say by the standard deviation of the income per capita logarithm across a group of countries or regions, declines over time"* (Barro and Sala-i-Martin, 1990). In this study, the interest variable is life expectancy, so the Gross Domestic Product (GDP) per capita will be replaced for each selected definition of life expectancy. Therefore, we will analyze whether the health differences between countries or regions and genders have decreased over time.

According to the literature, there are several methods to examine the σ-convergence. The dispersion can vary absolutely, through the range, standard deviation, variance and mean deviation; or in a relative way, dividing the previous statistics among the mean. The two most widely used methods to quantify the decrease in dispersion between the different economies are the standard deviation and the coefficient of variation (Janssen et al., 2016). Following the methodology developed by Barro and Sala-i-Martin (1990), the standard deviation will be used. This statistic is defined as:

$$\boldsymbol{\sigma}_t = \sqrt{\frac{\sum_{i=1}^{N} \left(\log\left(\boldsymbol{y}_{i,t}\right) - \boldsymbol{u}_t\right)^2}{N}} \tag{3.1}$$

where y is the interest variable, i is the studied individual, t is the year analyzed, N is the total number of individuals in the sample and u_t is the sample mean of the logarithm of $y_{i,t}$.

Therefore, this type of convergence will occur when the life expectancy dispersion among all the individuals in the sample is reduced, that is, if:

$$\sigma_{t+1} < \sigma_t \tag{3.2}$$

where σ_t is the standard deviation of the logarithm of $y_{i,t}$ for all i.

3.1.2 Beta-convergence

This kind of convergence occurs when the partial correlation between GDP growth per capita over time and its initial level value is negative. It's understood as the situation where *"poor economies tend to grow faster than rich ones"* (Sala-i-Martin, 1996). The β-convergence is a necessary but not sufficient condition for σ-convergence (Gächter and Theurl, 2011).

In turn, this β-convergence can be divided into two: the absolute and the conditioned. Absolute β-convergence occurs when all studied economies converge towards the same steady state. However, conditional β-convergence occurs when each country converges towards its own steady state, rather than a common steady state (economies would converge in the long term, but not towards the same levels) (Sala-i-Martin, 1996). This study will focus on the analysis of absolute β-convergence due to space and data limitations, so the life expectancy growth rate will be returned over the Napierian logarithm of the variable value in the first year analyzed.

Following the methodology used by Sala-i-Martin (1996), the following regression will be estimated with cross section data:

$$\ln\left(y_{i,t}/y_{i,t_0}\right) = a + b\ln\left(y_{i,t_0}\right) + u_{it} \tag{3.3}$$

where the dependent variable is the life expectancy growth rate, ln $(y_{i,0})$ is the logarithm of the variable in the first year, i are all the countries/regions in the sample, t_0 is the first year of study and t is the last one. The intercept a is assumed to be constant for all economies.

Based on this equation, there will be absolute β-convergence when the parameter b meets two requirements: be negative and statistically significant. Additionally, the annual rate of convergence to steady state is calculated as the negative value of the Napierian logarithm of 1 plus the coefficient b divided by the number of years in the period studied (T) (Jaworska, 2014), this is:

$$\beta = -\frac{\ln(1+b)}{T} \tag{3.4}$$

3.2 SPATIAL ANALYSIS

The classic convergence analysis between different places doesn't take into account the spatial relationships that can have between them. If we carry out a study of this caliber on geographically close areas, it's important to consider the possibility of spatial interactions (Stańczyk, 2015). For this, two spatial models are included as an analysis tool: the Spatial Lag Model and the Spatial Error Model. This allows us to take into account various spatial relationships between neighboring countries or regions (Jaworska, 2014).

3.2.1 Spatial weight matrix

The spatial weight matrix W is a non-stochastic square matrix. Its elements, w_{ij}, represent the intensity of the interdependence that exists between each pair of regions i and j. There are many ways to define matrix weights, as there is no single accepted definition. The only condition that must be met is that said weights must be non-negative and finite (Ullah et al., 1988). The most common and simplest way is through the use of binary notation, that is, w_{ij} will take the value 1 if the regions i and j are physically neighboring (they share a border) and 0 when they aren't. It should be noted that, generally, the matrix W is transformed to carry out standardization. Therefore, each element w_{ij} is divided by the total addition of the row to which it belongs. So, the addition of each row of the new standardized matrix W will be equal to unity (Moreno et al., 2002).

This matrix is defined as:

$$\begin{bmatrix} \mathrm{w_{11}w_{12}w_{13}\dots w_{1j}} \\ \mathrm{w_{21}w_{22}w_{23}\dots w_{2j}} \\ \dots \\ \mathrm{w_{i1}w_{i2}w_{i3}\cdots w_{ij}} \end{bmatrix} \tag{3.5}$$

where w_{ij} are the different weights between each pair of regions i and j.

3.2.2 Moran's I spatial autocorrelation contrast

Inside the literature of growth empirics, the most widely used test to measure spatial autocorrelation comes from a statistic developed by Moran (1958) (Anselin, 2001). The Moran's I statistic is used to detect the existence of spatial dependence in the samples studied. The null hypothesis of this contrast is the spatial non-autocorrelation, in other words, that the variable is randomly distributed throughout the territory (Moreno et al., 2002). Based on the methodology used by Moreno et al. (2002), this statistic is interpreted similarly to a covariance. It's based on the cross products of the differences with respect to the mean. It's formally defined as:

$$\mathrm{I} = \frac{\mathrm{N}}{\mathrm{S_0}} \cdot \frac{\Sigma_{\mathrm{ij}}^{\mathrm{N}} \mathrm{w_{ij}} \left(\mathrm{x_i} - \overline{\mathrm{x}}\right)\left(\mathrm{x_j} - \overline{\mathrm{x}}\right)}{\mathrm{S_i^N} \left(\mathrm{x_i} - \overline{\mathrm{x}}\right)^2} \quad \mathrm{i} \neq \mathrm{j} \tag{3.6}$$

where N is the sample size, S_0 is the scale factor equal to the addition of the weights, w_{ij} are the weights of the weight matrix W, x_i is the value of the interest variable x in the region i and $\overline{\mathrm{x}}$ is the sample mean of the variable x. If the weight matrix is standardized, N will be equal to S_0.

Furthermore, the interpretation of the statistic depends on its expected value, which is defined as:

$$E(I) = -\frac{1}{N-1} \tag{3.7}$$

If $I > E(I)$ there is positive spatial autocorrelation and if $I < E(I)$ there is negative spatial autocorrelation (Jaworska, 2014). However, if the I values are close to zero, there will be an absence of spatial pattern, so the observations will be distributed randomly in space (Jackson et al., 2009).

3.2.3 Spatial models

Based on literature, there are several regression models that include spatial effects through dependent and / or independent variables or through the error term. To consider both cases, two different models will be used in this study.

First of all, it's assumed that it's the endogenous variable that is spatially correlated. In this sense, the Spatial Lag Model (SLM) is applied, defined as:

$$y = \rho W y + X\beta + u \tag{3.8}$$

where y is the vector of the interest variable growth rate (life expectancy) over the given time period, p is the autoregressive parameter that picks up the intensity of the interdependencies between the sample observations, Wy is the spatial delay of the variable y, X is the vector of the observations on the interest variable in logarithms in the initial year and u is the white noise error term.

Secondly, the Spatial Error Model (or spatial error dependency) (SEM) is used. This model considers that spatial autocorrelation is present in the error term and it's formally defined as:

$$y = X\beta + \varepsilon \tag{3.9}$$

$$\varepsilon = \lambda W\varepsilon + \xi \tag{3.10}$$

where y and X are defined like the previous case, W is the spatial weight matrix, ξ is an error vector and λ is an autoregressive parameter in the error dependence model.

In addition, the conditions for absolute β-convergence are the same as in the model estimated by OLS: the estimation coefficient of X must have a negative sign and be significant (Jaworska, 2014). It should be noted that the specification by OLS is no longer valid for spatial models, so the description is made using the Maximum Likelihood method, in accordance with the chosen model (González, 2016).

4 RESULTS

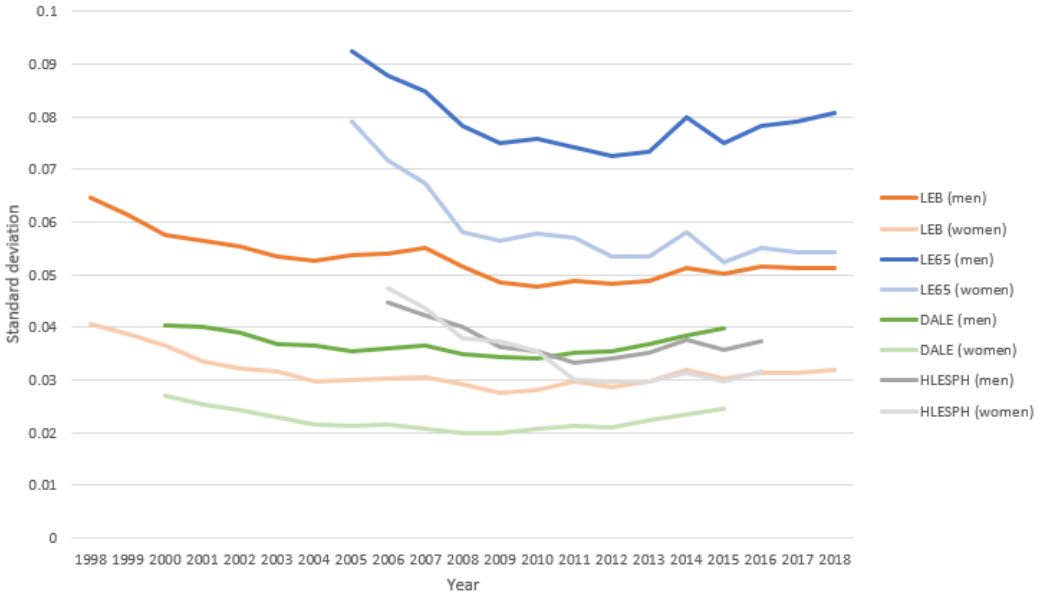
The purpose of this section is to present the convergence analysis results for men and women in the four variables selected for the 27 European Union countries and for the 50 Spanish provinces (plus the two autonomous cities). First of all, results of σ-convergence and absolute β-convergence will be presented. Then, Moran's Test results will be shown. If this contrast accepts the null hypothesis of no autocorrelation, no further calculations will be necessary and the previous results will be completely valid. However, if the test rejects the null hypothesis, it will be necessary to estimate the spatial models to take into account spatial dependencies (Moreno et al., 2002).

4.1 ANALYSIS OF THE EUROPEAN CASE

4.1.1 Sigma-convergence

Figure 4.1 shows the evolution of the standard deviation in all cases. The most striking feature is that, in all variables, the women's curve is similar to that of men, but it's underneath. This means that the dispersion among women is smaller than the existing dispersion among men.

Figure 4.1: EU σ-convergence



Source: Authors' estimations

Firstly, the life expectancy at birth curves are presented. Their initial values are greater than the final ones, for both men (0.065 versus 0.051) and women (0.041 versus 0.032). Those values decrease until 2004, then they go up a little bit but they go down again until 2009-2010. From here the trend is increasing, despite having small jumps. When the standard deviation decreases, it means that the dispersion between the life expectancy at birth of all men and women in the sample is reduced and, therefore, that there is a σ-convergence process. The opposite happens when the curve grows. Therefore, in this case, we could say that the convergence and divergence processes alternate: until 2004 there is convergence, then it diverges, then it converges again, etc.

Secondly, the life expectancy curves at age 65 are the ones with the most variability. They have two clear trends and a peak: the standard deviation decreases until 2012-2013 and increases from 2015. Between them there is a peak. Therefore, there is a

σ-convergence process until the beginning of the 2010s and, from 2015, a divergence process. In this case, the initial value is much higher at the end, for both men (0.093 versus 0.080) and women (0.079 versus 0.054). It's the variable that in the initial years has the sharpest decrease.

Third, the trend of the disability-adjusted life expectancy curves are similar to those of the life expectancy at birth variable, although the streaks are somewhat more pronounced. The general trend is that dispersion decreases until 2010 for men and 2008 for women, and thereafter increases. However, there is a peak in 2007 for men and in 2006 for women. In the case of men, sigma's initial value is practically the same as the final one (approximately 0.04). Nonetheless, in the case of women, it occurs as in the previous variables, the initial value (0.027) is higher than the final one (0.024).

Finally, the case of the healthy life expectancy based on self-perceived health variable is the only one in which the curve of men meets that of women. The women's dispersion in the initial year (0.048) is greater than that of men (0.045). Despite the fact that in the final year the standard deviation in both cases is less than in the initial year, the sigma value is higher in men (0.037) than in women (0.032). The men's dispersion decreases until 2011 and, thereafter, it increases until 2016, with a small drop in 2015. The women's curve also decreases until 2011 but it remains constant for the following two years. Then it follows the same evolution as in men: it grows until the last year but with a slight drop in 2015.

4.1.2 Beta-convergence

As stated above, β-convergence occurs when the partial correlation between the growth of life expectancy over time and its initial level is negative. In this case, we focus on the absolute β-convergence analysis, which will occur if the countries converge towards the same steady state (Sala-i-Martin, 1996).

Table 4.1 shows the results of returning the growth rate of the interest variable over its value in the initial year.

Table 4.1: EU estimated results of OLS (men / women)

OLS (men / women)	**LEB**		**LE65**	
	Estimate	t value	Estimate	t value
a (intercept)	1.28*** / 1.50***	4.84 / 4.85	0.24 . / 0.72***	1.73 / 4.02
b	- 0.28*** / -0.33***	-4.55 / -4.68	-0.04 / -0.21**	-0.83 / -3.47
converg. speed (annual)	1.65% / 2.02%		- / 1.87%	
R-squared	0.45 / 0.47		0.03 / 0.33	
Adjusted R-squared	0.43 / 0.46		-0.01 / 0.30	
OLS (men / women)	**DALE**		**HLESPH**	
	Estimate	t-value	Estimate	t-value
a (intercept)	0.59* / 0.99**	2.50 / 3.64	1.13*** / 1.71***	6.82 / 7.33
b	-0.12* / -0.22**	-2.48 / -3.50	-0.25*** / -0.39***	-6.45 / -7.06
converg. speed (annual)	0.88% / 1.63%		2.92% / 4.89%	
R-squared	0.17 / 0.33		0.62 / 0.67	
Adjusted R-squared	0.13 / 0.30		0.61 / 0.65	

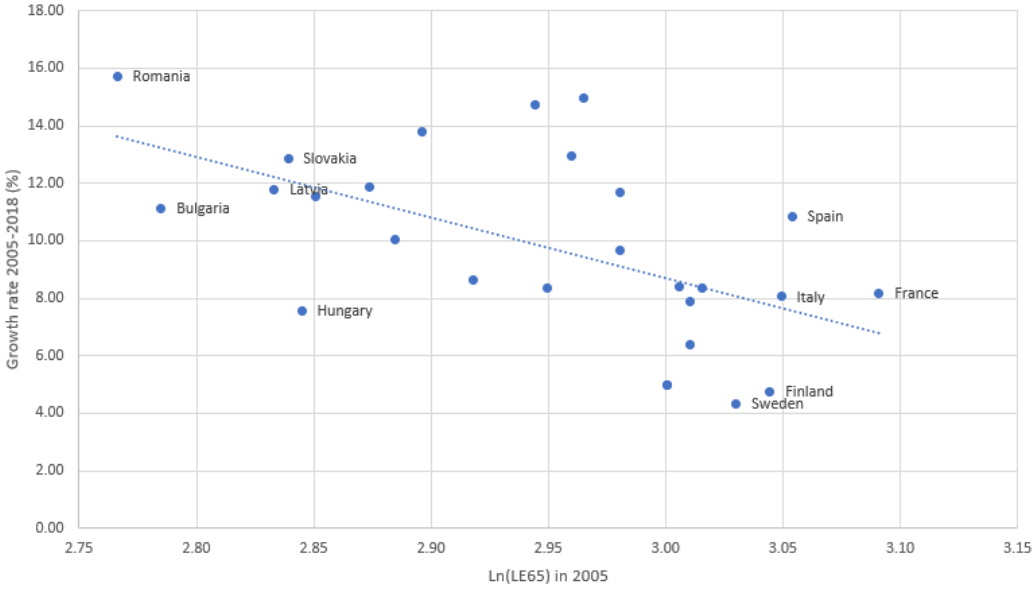
***Significant at level 0.001; **Significant at level 0.01; *Significant at level 0.05; . Significant at level 0.1

Source: Authors' estimations

In the first case, the b interest coefficient is negative and significant for men and women (-0.28*** and -0.33*** respectively). So, there is an absolute β-convergence process for males and females between 1998 and 2018 in life expectancy at birth. It can be affirmed that the countries that initially had the lowest life expectancy at birth are those that have experienced higher growth rates and that the 27 EU countries converge towards the same steady state. However, the speed of convergence (annual) is higher for women (2.02%) than for men (1.65%).

Secondly, the b interest coefficient of the regression for life expectancy at age 65 is negative and significant for women (-0.21**) but not for men (-0.04). That means that the 27 EU countries converge towards the same steady state but only for women. There is absolute β-convergence for women but not for men between 2005 and 2018. In other words, the life expectancy at age 65 growth rate and its value at the initial moment (2005) have a negative relationship for women, as shown in Figure 4.2 [5] (negative trend). The speed of annual convergence is slightly smaller than life expectancy at birth (1.87%).

Figure 4.2: EU LE65 scatter plot (women)



Source: Authors' estimations

As it can be seen in the previous figure, the countries that have had the highest growth rates are those that had a lower life expectancy value in the initial year (2005). Romania is the country with the highest growth rate and the lowest life expectancy at age 65 value in 2005. However, countries like Sweden, Finland and France have had lower growth rates and their initial values were also lower.

Then, the interest variable in the following regression is the disability-adjusted life expectancy. This case is similar to the first one, the b coefficients for men and for women are negative and significant (-0.12* and -0.22**, respectively), but lower than those of the first regression. Therefore, there is an absolute β-convergence process, but in a lower rate than for life expectancy at birth between 2000-2015. The speed of annual convergence is 0.88% for men and 1.63% for women. So, the countries that initially had the lowest disability-adjusted life expectancy are those that have experienced higher growth rates.

[5] All points in the graph correspond to one country. For ease of viewing, only the names of the countries with extreme values have been written.

Finally, the b coefficients for the healthy life expectancy based on self-perceived health are also significant and negative, so there is absolute β-convergence between 2006 and 2016. However, it should be noted that it's the variable with the highest speed of convergence: 2.92 % for men and 4.89 % for women.

It's worth noting that, in all cases where absolute β-convergence exists, the necessary (but not sufficient) condition for the previous analyzed convergence (σ) to exist is fulfilled (Gächter and Theurl, 2011).

As previously anticipated, this traditional β-convergence approach might not be valid if there is spatial dependency. So, Moran's I statistic will be used to identify spatial dependence. Its null hypothesis is spatial no autocorrelation (Moreno et al., 2002). Table 4.2 shows this results for both sexes in the four different variables. This suggests that the null hypothesis can be rejected in all cases and evidences the presence of spatial dependence in the analyzed sample. Additionally, the autocorrelation is positive because the I statistic in all cases is higher than - $\frac{1}{N-1}$ (-0.038). Therefore, the presence of a certain phenomenon (in this case, the different types of life expectancy) in a specific spatial unit makes it easily spread to its neighboring towns, so favoring its grouping. However, it can be seen how the value of the I statistic is lower in the final year studied than in the initial year and how it's higher for men than for women, in most variables. This indicates that, in this case, positive autocorrelation decreases over time and that men have a bigger tendency to cluster than women.

Table 4.2: EU Moran's I statistic results (men / women)

Moran's I statistic (men / women)	**LEB**		**LE65**	
Year	1998	2018	2005	2018
I	0.64 / 0.66	0.62 / 0.56	0.62 / 0.60	0.57 / 0.55
p-value	0.00026 / 0.00020	0.00042 / 0.0012	0.00049 / 0.00058	0.0013 / 0.0014
Moran's I statistic (men / women)	**DALE**		**HLESPH**	
Year	2000	2015	2006	2016
I	0.64 / 0.62	0.64 / 0.58	0.55 / 0.44	0.55 / 0.38
p-value	0.00025 / 0.00037	0.00023 / 0.00089	0.001872 / 0.01003	0.001762 / 0.02354

Source: Authors' estimations

Because of the previous results, spatial dependence between countries must be taken into account. This means that the absolute β-convergence must be estimated through the Spatial Lag Model and the Spatial Error Model for the same variables and the same previous years. The conditions for absolute β-convergence to exist remain the same: b must be negative and significant (Jaworska, 2014).

Table 4.3 shows that there have been no major changes. All the b interest coefficients that were previously negative and significant continue to be. So that in all cases in which there was absolute β-convergence it continues to be, even taking into account spatial dependencies. In this sense, there is still not enough evidence to justify an absolute β-convergence in life expectancy for men at age 65. Depending on the model considered, the speeds of convergence increase or decrease. The ultimate reason for this is that the first model considers that correlation is present in the endogenous variable (SLM) while the second includes autocorrelation in the error term (SEM) (Moreno et al., 2002). The most important thing to note is that the speed of annual convergence is bigger for women than for men, in any of the life expectancy variable studied. Regardless of the model selected, the fastest convergence is among women in HLESPH (6.26% and 4.86%, respectively).

In summary, the inclusion of spatial models allows for the consideration of spatial dependencies between observations. In this case, the results are quite similar to those previously estimated: all variables spatially β-converge, except men in LE65.

Table 4.3: EU estimated results of ML (men / women)

SLM (men / women)	**LEB**		**LE65**	
	Estimate	z value	Estimate	z value
a (intercept)	1.29*** / 1.57***	4.26 / 4.76	0.34* / 0.81***	2.54 / 4.59
b	-0.28*** / -0.35***	-4.06 / -4.63	-0.07 / -0.24***	-1.48 / -4.03
converg. speed (annual)	1.64% / 2.14%		- / 2.08%	
Log likelihood	70.85 / 82.12		58.54 / 60.96	
SLM (men / women)	**DALE**		**HLESPH**	
	Estimate	z value	Estimate	z value
a (intercept)	0.54*** / 1.11***	2.20 / 4.25	1.31*** / 2.06***	6.75 / 9.04
b	-0.11*** / -0.24***	-1.98 / -4.09	-0.29*** / -0.46***	-6.53 / -8.82
converg. speed (annual)	0.80 % / 1.85%		3.45% / 6.26%	
Log likelihood	76.08 / 88.86		72.71 / 68.31	
SEM (men / women)	**LEB**		**LE65**	
	Estimate	z value	Estimate	z value
a (intercept)	1.29*** / 1.50***	4.63 / 4.78	0.22 / 0.71***	1.41 / 3.56
b	- 0.28*** / -0.33***	-4.35 / -4.61	-0.03 / -0.20**	-0.56 / -3.03
converg. speed (annual)	1.68% / 2.02%		- / 1.76%	
Log likelihood	71.14 / 82.09		57.54 / 61.03	
SEM (men / women)	**DALE**		**HLESPH**	
	Estimate	z-value	Estimate	z-value
a (intercept)	0.63* / 0.97***	2.42 / 4.03	1.14*** / 1.70***	6.65 / 8.11
b	-0.13* / -0.21***	-2.19 / -3.86	-0.26*** / -0.39***	-6.30 / -7.82
converg. speed (annual)	0.95% / 1.60%		3.01% / 4.86%	
Log likelihood	75.04 / 85.54		72.09 / 65.64	

***Significant at level 0.001; **Significant at level 0.01; *Significant at level 0.05

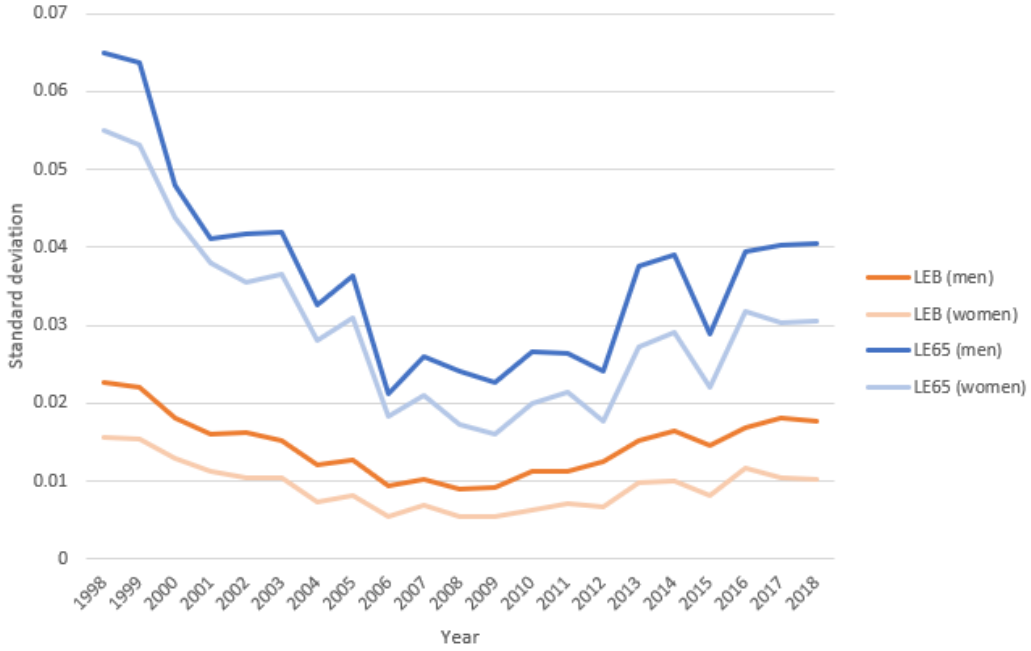
Source: Authors' estimations

4.2 ANALYSIS OF THE SPANISH CASE

4.2.1 Sigma-convergence

The evolution of σ, that is to say, of the standard deviation, of the selected variables between 1998-2018 for the Spanish provinces is presented graphically below. It can be differentiated in both cases between the dispersion curve of men and women. As in the European Union countries analysis, it can be seen how the evolution of the curves is very similar between one gender and the other, but the women's dispersion is always smaller than that of men.

Figure 4.3: Spanish σ-convergence



Source: Authors' estimations

The variability of life expectancy at birth curves is much smaller than that of life expectancy at age 65. The σ evolution in the first case is much more stable. Despite having small streaks, the variability isn't as great as for the second variable. In fact, the initial and final values are very close. In 1998 the dispersion value is 0.023 for men and 0.016 for women. In 2018 these values are 0.018 and 0.010, respectively. Looking at the general trends, they suggest that the standard deviation fell to 2008 for men and 2009 for women. From these years and until 2018, it increases. Therefore, it could be said that there is a σ-convergence process in approximately the first half of the sample studied, while in the second there is a σ-divergence process.

On the other hand, the variability of the life expectancy at age 65 curves is much greater than that of life expectancy at birth. It can be seen how, in the case of the first variable, the σ value is much greater in 1998 than in 2018. In the initial year this value is 0.065 for men and 0.055 for women, while in the final year the values are 0.040 and 0.031, respectively. As said before, when dispersion decreases there is a σ-convergence process, while when σ increases there is a divergence process. In this case, the numerous streaks invite us to think that there have been constant changes in these processes. However, the general trend is a decrease in the standard deviation until 2009. From there, if the streaks are ignored, the σ value grows on both curves. Therefore, it could be said that for life expectancy at age 65, ignoring streaks, there is a σ-convergence process from 1998 to 2009 and a σ-divergence process from that year to 2018 for both sexes.

4.2.2 Beta-convergence

After having analyzed the σ-convergence in the Spanish provinces, we proceed to study the absolute β-convergence between the same years as before (1998-2018).

First of all, results of the traditional absolute β-convergence analysis are presented. Table 4.4 shows the estimation results of the model by OLS. The growth rate between 1998 and 2018 has been estimated on the Napierian logarithm of the interest variable in the initial year.

The requirements for absolute β-convergence are still the same, the value of the b interest coefficient must be significant and negative (see 3.2.3). As it can be seen, the values of b for life expectancy at birth are -0.45*** for men and -0.25*** for women. In the case of life expectancy at age 65, these values are -0.58*** and -0.24***, respectively. In all four situations they are significant. Therefore, according to the estimation by OLS, there is an absolute β-convergence process in both variables for both men and women. This means that the provinces converge towards their own steady state (Sala-i-Martin, 1996). However, it should be noted that the annual convergence rate is much higher (more than double) for men. In life expectancy at birth it's 2.7% while for life expectancy at age 65 it reaches a value of 4.28%, compared to women, who have a speed of convergence of 1.46% and 1.37%, respectively.

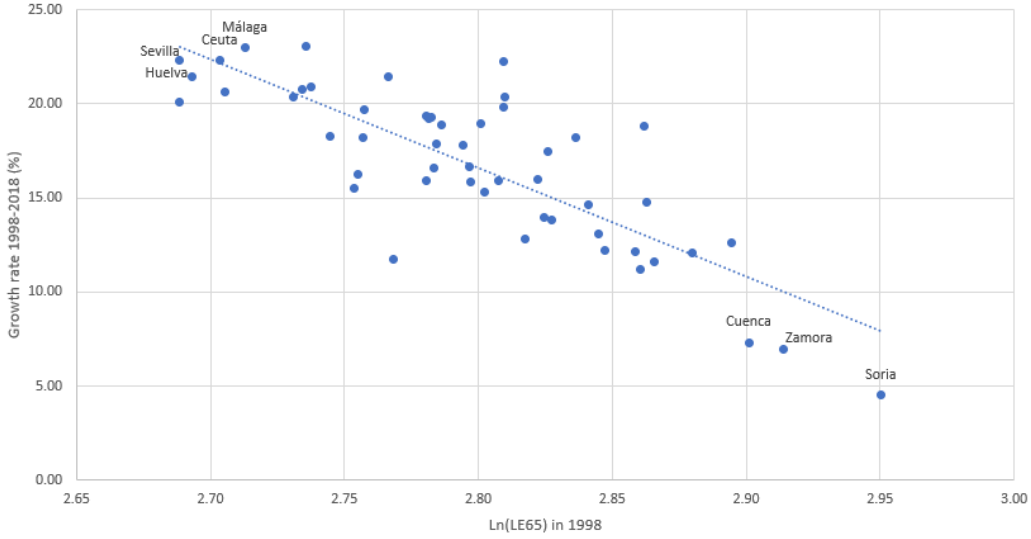
Table 4.4: Spain's estimated results of OLS (men / women)

OLS (men / women)	**LEB**		**LE65**	
	Estimate	t value	Estimate	t value
a (intercept)	2.00*** / 1.15***	8.37 / 5.40	1.78*** / 0.85***	10.99 / 4.99
b	-0.45*** / -0.25***	-8.10 / -5.21	-0.58***/ -0.24***	-9.95 / -4.22
converg. speed (annual)	2.97% / 1.46%		4.28% / 1.37%	
R-squared	0.57 / 0.35		0.66 / 0.26	
Adjusted R-squared	0.60 / 0.40		0.66 / 0.25	

***Significant at level 0.001; **Significant at level 0.01; *Significant at level 0.05
Source: Authors' estimations

The highest convergence speed is found by men of life expectancy at age 65 (4.28%). Figure 4.4 [6] shows the negative slope of this regression. The provinces that initially had lower life expectancy values are those that have had a higher growth rate, and vice versa. Most of them are provinces in southern Spain, such as Seville, Málaga and Huelva. On the other hand, the province that had the highest life expectancy value in 1998 (Soria), is the one with the lowest growth rate. All of the above is evidence in favor of absolute β-convergence.

Figure 4.4: Spanish LE65 scatter plot (men)



Source: Authors' estimations

[6] All points in the graph correspond to one province. For ease of viewing, only the names of the provinces with extreme values have been written.

Then, it's need to know whether the sample data is subject to spatial dependency or not. As said before, the traditional estimation results of absolute β-convergence may not be valid if there is spatial autocorrelation in the sample (Moreno et al., 2002). Table 4.5 shows these results. Both variables reject the null hypothesis of no autocorrelation for the initial and final year. There is enough evidence to justify the presence of spatial dependence between the different Spanish provinces. Men have higher values of statistic I than women in all cases except in 1998 LE65. Moreover, since all the statistics are higher than - $\frac{1}{N-1}$ (-0.020), it can be affirm that there is positive spatial autocorrelation. This means that the presence of high life expectancy rates in a specific province favors the high rates in its neighboring regions.

Table 4.5: Spain Moran's I statistic results (men / women)

Moran's I statistic (men / women)	**LEB**		**LE65**	
Year	1998	2018	1998	2018
I	0.44 / 0.40	0.47 / 0.41	0.49 / 0.62	0.55 / 0.51
p-value	7.092e-07 / 3.318e-06	1.648e-07 / 2.94e-06	5.366e-08 / 1.267e-11	1.292e-09 / 8.367e-09

Source: Authors' estimations

Finally, it's necessary to take into account the spatial dependencies affecting the sample. For this reason, the Spatial Lag Model and the Spatial Error Model will be estimated using the Maximum Likelihood method. The interest variable growth rate between 1998 and 2018 has been returned on the logarithm value of that variable in the initial year (1998). These results are reflected in Table 4.6. As mentioned above, absolute β-convergence occurs when b is negative and significant (Jaworska, 2014).

As can be seen, for both men and women in both variables the requirements for convergence are met. Therefore, even taking into account spatial dependencies, there is an absolute β-convergence process. Convergence rates continue to be higher for men than for women. In this sense, when considering spatial autocorrelation, it can be seen how the gap between men and women is somewhat greater than if it weren't considered (except for the LE65 variable in SLM). In this case, the speed of convergence also increase or decrease depending on the model selected. Whether SLM or SEM is chosen, the fastest convergence is possessed by men in the LE65 variable (4.17% and 5.34%, respectively).

Table 4.6: Spain's estimated results of ML (men / women)

SLM (men / women)	**LEB**		**LE65**	
	Estimate	z value	Estimate	z value
a (intercept)	1.92*** / 1.02***	8.02 / 4.31	1.76*** / 0.88***	11.37 / 5.18
b	-0.43*** / -0.22***	-7.71 /-4.11	-0.57*** / 0.25***	-10.19 / -4.40
converg. speed (annual)	2.79% / 1.26%		4.17% / 1.44%	
Log likelihood	184.67 / 199.56		120.49 / 132.97	
SEM (men / women)	**LEB**		**LE65**	
	Estimate	z value	Estimate	z value
a (intercept)	2.09*** / 1.14***	8.39 / 5.55	2.00*** / 0.85***	10.76 / 6.46
b	-0.47***/ -0.25***	-8.13 / -5.35	-0.66*** / -0.24***	-9.83 / -5.47
converg. speed (annual)	3.16 % / 1.43%		5.34% / 1.37%	
Log likelihood	184.20 / 199.02		120.53 / 133.65	

***Significant at level 0.001; **Significant at level 0.01; *Significant at level 0.05

Source: Authors' estimations

5 DISCUSSION

This section reviews the main results of the health inequalities analysis between genders, countries and regions.

The first inequality presented is that life expectancy is much higher for women than for men, regardless of the life expectancy concept considered (see Table 2.2 and Table 2.3). This dissimilarity is mostly explained by the different risk behaviours and habits of the working age population. Although it's true that, over time, societies are much more egalitarian. Women have entered the labour market and can carry out the same activities and occupations as men, which may favour gender convergence in life expectancy (Garcia and Grande, 2018).

In the European Union context, the results show, in general terms, that the convergence process has taken place in the period 1998-2018, both in σ and in absolute β.

First of all, the σ analysis propose that stages of σ-convergence have alternated with stages of σ-divergence between EU countries (see Figure 4.1). It could be said that life expectancy has converged at the end of the 20th century, while the first two decades of the 21st century show an opposite trend. Despite this behaviour in recent years, dispersion has decreased in all variables. Both in the initial and final year, the dispersion is greater for men than for women (except for the HLESPH variable in 1998). However, it has decreased more in women than in men (except for the LEB). The indicator that has suffered the biggest dispersion drop is LE65 (0.013 for men and 0.025 for women).

Secondly, the spatial results (see Table 4.2) suggest that spatial dependence should be taken into account. Moran's Test shows that there is a positive spatial autocorrelation, that is, increases in life expectancy in one country positively influence increases in its neighbouring countries. Therefore, taking into account both the results of the analysis without considering the spatial dependence and those of the study including it, the absolute β-convergence occurs in all variables, except for men in LE65 (see Table 4.3). This means that, for the most part, the countries with the lowest life expectancy in 1998 have had the highest growth rates. In addition, each country is approaching its own steady state. Women's convergence rates are higher than for men in all variables. Women have the fastest convergence rate with the HLESPH indicator (4.89% for OLS estimation and 6.26% and 4.86% taking into account spatial dependence, SLM and SEM, respectively).

Thirdly, results show that there could be a spatial distinction between Western and Eastern Europe countries (see Figure 2.3 and Figure 4.2). Initially, Eastern countries had lower life expectancy values. However, there has been a convergence process between all this countries, as the Eastern ones have had higher growth rates than the Western ones.

It can therefore be said that health inequalities between countries within the EU have been reduced between 1998 and 2018. One possible justification is that it has been relatively long since a country last joined the EU (Croatia in 2013). As countries have been members longer, they tend to have higher economic growth (Dávila, 2011). Countries that were "poorer" than the existing members at the time of their entry have had enough time to grow and become as similar as possible to the older countries.

Summarizing, it's confirmed that there is life expectancy convergence among EU countries, but that women are converging to a bigger extent than men. Authors such as Jaworska (2014) and Stánczyk (2015) guarantee these results for the first years of the 21st century.

On the other side, the analysis results of the Spanish provinces also show that both convergence processes have taken place between 1998-2018: in σ and in absolute β. As in the previous occasion, several issues of interest will be highlighted.

First of all, life expectancy, like the in the European Union countries, is constantly changing between σ-convergence and σ-divergence processes (see Figure 4.3). Even so, two different stages could be distinguished. The first phase is from 1998 until the first years of the economic crisis in 2008, and is characterized by a sharp dispersion reduction. The second phase starts from these years and ends in 2018, and is characterized by an increase in standard deviation. However, in general terms, dispersion has been reduced in the two variables studied, but to a greater extent in LE65. The σ value is higher for men, both in the initial and final year. Despite the fact that the gender gap is still present (women have a longer life expectancy than men) and that the falls in dispersion in both genders have been very similar, it has been bigger in women for LEB and in men for LE65.

Secondly, Moran's Test shows that there is a positive spatial autocorrelation in the sample (see Table 4.5). The provinces have spatial interactions in life expectancy terms. Spatial positive autocorrelation suggests that increased life expectancy in one province has led to positive effects in its neighbouring provinces. The LE65 variable has a higher positive spatial autocorrelation than the LEB variable. Therefore, we have to take into account these issues when estimating absolute β-convergence.

Thirdly, the β-convergence results are different for Spanish provinces than for European countries. Taking into account the results of the estimation by OLS and those of MV, there has been a convergence process in both sexes (see Table 4.4 and Table 4.6). This indicates that the life expectancy growth rate in Spanish provinces has been higher in those that initially had smaller values. However, this β-convergence has occurred to a bigger extent in men (while for the EU it occurred more in women). Men have the highest convergence speed in LE65: 4.28% in the first regression and 4.17% (SLM) and 5.34% (SEM) in the spatial models.

Fourth, both the data analysis and the scatter plot (see Figure 2.4 and Figure 4.4) suggest that there are life expectancy inequalities in geographical in Spain. The northern provinces have higher life expectancy values than the southern ones. Nevertheless, convergence has been present at the end of the 20th century and the beginning of the 21st century. The dispersion between provinces has tended to be reduced, so these inequalities have also diminished. Despite the fact that since the economic crisis the dispersion between provinces has increased, β-convergence results show that in 1998-2018 the provinces with the lowest life expectancy have had the highest growth. Finally, the life expectancy gender gap has also narrowed over the last two decades in Spain. The speed of men's convergence is much greater than that of women. Men's life expectancy has increased considerably between 1998 and 2018, especially from the age of 65 onwards.

In summary, the convergence process in σ and spatial absolute β that the Spanish provinces suffer between 1998-2018 mean that the existing inequalities (geographical and gender) are reduced. However, the trend of recent years shows that perhaps the dispersion between provinces will move away again.

6 CONCLUSIONS

This study analyses the σ, absolute β and spatial absolute β-convergence of four different life expectancy definitions between men and women for the European Union countries and for Spanish provinces during several periods, all of them between 1998-2018.

The results show that there has been a convergence process in all cases, which could mean that health inequalities are lower in 2018 than at the end of the 20th century. Firstly, as the dispersion in the initial years studied is bigger than in the final years, there has been a σ-convergence process in all analyses. Nevertheless, this process hasn't remained constant throughout the analysed period. The trend in the last years studied seems to indicate that the dispersion is increasing again. Therefore, perhaps during the next few years health differences between some territories and others will increase again. Secondly, as the observations were spatially dependent, spatial interactions had to be taken into account for the calculation of β-convergence. Lastly, as the countries and provinces with lower life expectancy values at the beginning of the study are those with higher growth rates, there has been a β-convergence process. The speed of this process has been higher in women for European countries and higher in men for Spanish provinces. In this context, gender inequalities in terms of life expectancy have been reduced in recent decades.

One of the main limitations that have been presented is that the life expectancy convergence analysis serves to know whether life expectancies between some territories tend to get closer. However, it doesn't really show people's quality of life, since increased life expectancy doesn't necessarily translate into improved living conditions. Although two variables (disability-adjusted life expectancy and healthy life based on self-perceived health) have been included in this study for this purpose, they haven't been sufficient to reflect the population's standard of living, as other factors may influence it.

In this sense, two lines are proposed for future research. In the first place, and based on the previous paragraph, it's proposed to estimate conditional β-convergence, through the use of variables that measure living conditions in other areas, such as the economic or environmental (Gispert et al., 2007). Secondly, this same convergence analysis is invited to perform, but using predicted life expectancy data. In this way, it would be possible to have an estimate on whether the dispersion will really continue to increase, which could lead to divergence processes.

Finally, this study has a double contribution. First of all, it provides an updated vision of the health convergence in recent years, gender-differentiated. Secondly, it contributes to the incorporation of spatial econometrics in Health Economics research. Our analysis can be considered in order to apply policy intervention measures to reduce health population inequalities.

7 APPENDIX: R CODE

7.1 SIGMA-CONVERGENCE

The starting year and the sample size are adapted according to each type of analysis (Europe / Spain) and to each variable.

```
df <- read.table(file = "datos.txt", header = TRUE)

# Convert df into a time series

timeserie <- function(w){
  t <- ts(w, start = c(1998, 1), frequency = 1)
  return(t)
}

transformadas <- data.frame(lapply(df, timeserie))
traspuesta <- data.frame(t(transformadas))

# Average of each year for the 27 EU countries or for the Spanish provinces

media <- data.frame(apply(df, 1, mean))

# Passing averages to time series and transposing

a <- data.frame(ts(media, start = c(1998, 1), frequency = 1))
at <- data.frame(t(a))

# Function that calculates the sigma-convergence

sigmaconv <- function(d, m, N){
  sig = ((sum((log(d) - log(m))^2)/N))^0.5
  return(sig)
}

# Calculation for all countries or provinces

datosig <- c()
for (i in traspuesta){
  sigmaconv <- function(d, m, N){
    sig = ((sum((log(d) - log(m))^2)/N))^0.5
    return(sig)
  }
  datosig <- c(datosig, sigmaconv(i, at, 27))
}

datafinal <- data.frame(datosig)
datafinal <- ts(datafinal, start = c(1998, 1), frequency = 1)
```

7.2 BETA-CONVERGENCE

The logarithm is adapted to the initial year of the sample according to each variable.

```
datos <- read.table("datos.txt", header = T)
reg <- lm(datos$Tasacrecimiento ~ log(datos$X1998))
summary(reg)
```

7.3 SPATIAL ANALYSIS

7.3.1 Analysis of the European case

The logarithm and the Moran's Test are adapted to the initial and/or final year of the sample according to each variable.

```
library(spdep)
library(maptools)
library(RColorBrewer)
library(classInt)
library(rgdal)
library(RColorBrewer)
library(classInt)
library(spatialreg)

# 1. Read shapes files

rm(list = ls())
paises <- rgdal::readOGR("NUTS_RG_60M_2016_4326_LEVL_0.shp")
paises1 <- paises [-c(1, 3, 7, 11, 24, 26, 30, 31, 36, 37), ]
plot(paises1)

# 2. Read csv files

datos <- read.table("datos.txt", header = T)
datos$Cod <- paises1@data$PAIS
LlarC <- as.matrix(datos[,2])

ejemplo1 <- as.data.frame(LlarC)
names(ejemplo1) <- "ejemplo1"
row.names(ejemplo1) <- row.names(paises1)
paises.data <- SpatialPolygonsDataFrame(paises1, ejemplo1)

# 3. Build W

pr.nb <- poly2nb(paises.data, queen = TRUE)
wqueen <- nb2listw(pr.nb, style = "W", zero.policy = TRUE)

# 4. Moran Test

moran <- moran.test(paises.data$ejemplo1, wqueen, randomisation = TRUE,
alternative = "two.sided", na.action = na.exclude, zero.policy = T)
```

```
# 5. Spatial models

slm <- lagsarlm(Tasacrecimiento ~ log(X1998), data = datos, wqueen, zero.policy = T)
summary(slm)

sem <- errorsarlm(Tasacrecimiento ~ log(X1998), data = datos, wqueen, zero.policy = T)
summary(sem)
```

7.3.2 Analysis of the Spanish case

The logarithm and the Moran's Test are adapted to the initial and/or final year of the sample according to each variable.

```
library(spdep)
library(maptools)
library(RColorBrewer)
library(classInt)
library(rgdal)
library(RColorBrewer)
library(classInt)
library(spatialreg)

# 1. Read shapes files

rm(list = ls())
provincias <- rgdal::readOGR("gadm36_ESP_2.shp")

# 2. Read csv files

datos <- read.table("datos.txt", header = T)
datos$Cod <- provincias@data$NAME_2
LlarC <- as.matrix(datos[, 2])

ejemplo1 <- as.data.frame(LlarC)
names(ejemplo1) <- "ejemplo1"
row.names(ejemplo1) <- row.names(provincias)
provincias.data <- SpatialPolygonsDataFrame(provincias, ejemplo1)

# 3. Build W

pr.nb <- poly2nb(provincias.data, queen = TRUE)
wqueen <- nb2listw(pr.nb, style = "W", zero.policy = TRUE)

# 4. Moran Test

moran <- moran.test(provincias.data$ejemplo1, wqueen, randomisation = TRUE,
alternative = "two.sided", na.action = na.exclude, zero.policy = T)
```

```
# 5. Spatial models

slm <- lagsarlm(Tasacrecimiento ~ log(X1998), data = datos, wqueen, zero.policy = T)
summary(slm)

sem <- errorsarlm(Tasacrecimiento ~ log(X1998), data = datos, wqueen, zero.policy = T)
summary(sem)
```

8 REFERENCES

ANSELIN, Luc; BERA, Anil K. Introduction to spatial econometrics. Handbook of applied economic statistics, 1998, vol. 237.

ANSELIN, Luc. Spatial econometrics. A companion to theoretical econometrics, 2001, vol. 310330.

BALTAGI, Badi H.; MOSCONE, Francesco; SANTOS, Rita. Spatial health econometrics. Health Econometrics, 2018, p. 305.

BARRO, Robert J.; SALA-I-MARTIN, Xavier. Economic growth and convergence across the United States. National Bureau of Economic Research, 1990.

BOHÓRQUEZ, Ingrid Acevedo; CEBALLOS, Hermilson Velásquez. Algunos conceptos de la econometría espacial y el análisis exploratorio de datos espaciales. Ecos de Economía: A Latin American Journal of Applied Economics, 2008, vol. 12, no 27, p. 9-2.

CLARK, Rob. World health inequality: convergence, divergence, and development. Social Science Medicine, 2011, vol. 72, no 4, p. 617-624.

DÁVILA, Juan José Toribio. Áreas monetarias óptimas y la experiencia europea: algunas reflexiones. ICE, Revista de Economía, 2011, no 863.

EUROSTAT BACKGROUND 2020a. Available in:
https://ec.europa.eu/eurostat/web/nuts/background

EUROSTAT 2020b, Healthy life expectancy based on self-perceived health. Available in:
https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_silc_17&lang=en

EUROSTAT 2020c, Life expectancy at age 65, by sex. Available in:
https://ec.europa.eu/eurostat/databrowser/view/tps00026/default/table?lang=en

EURSOTAT 2020d, Regional Policy, Glossary. Available in:
https://ec.europa.eu/regional_policy/en/policy/what/glossary/n/nuts/

FORSTER, Timon; KENTIKELENIS, Alexander; BAMBRA, Clare. Health inequalities in Europe: setting the stage for progressive policy action. 2018.

GÄCHTER, Martin; THEURL, Engelbert. Health status convergence at the local level: empirical evidence from Austria. International journal for equity in health, 2011, vol. 10, no 1, p. 34.

GADM maps and data 2020. Available in: https://gadm.org/

GARCÍA, JM González; GRANDE, Rafael. The changing sex differences in life expectancy in Spain (1980-2012): decomposition by age and cause. Gaceta sanitaria, 2018, vol. 32, no 2, p. 151-157.

GISPERT, Rosa, et al. Differences in disability-free life expectancy by gender and autonomous regions in Spain. Revista española de salud pública, 2007, vol. 81, no 2, p. 155-165.

GONZÁLEZ, Miguel Ángel Mendoza. Econometría Espacial: Aplicaciones con R. 2016.

HERNÁNDEZ, Claudia, et al. Preprocesamiento de datos estructurados. Revista vínculos, 2008, vol. 4, no 2, p. 27-48.

HEMBRAM, Sulekha; HALDAR, Sushil Kr. Is India experiencing health convergence? An empirical analysis. Economic Change and Restructuring, 2020, p. 1-28.

INE 2020a, Esperanza de Vida a los 65 años por provincia, según sexo. Available in: https://www.ine.es/jaxiT3/Tabla.htm?t=1486&L=0

INE 2020b, Esperanza de Vida al Nacimiento por provincia, según sexo. Available in: https://www.ine.es/jaxiT3/Tabla.htm?t=1485

JACKSON, Monica C., et al. Comparison of tests for spatial heterogeneity on data with global clustering patterns and outliers. International journal of health geographics, 2009, vol. 8, no 1, p. 55.

JANSSEN, Fanny, et al. Sigma and beta convergence in regional mortality: A case study of the Netherlands. Demographic Research, 2016, vol. 35, p. 81-116.

JAWORSKA, Renata. Health inequalities across the European Union regions: A beta-convergence approach. Comparative Economic Research, 2014, vol. 17, no 4, p. 71-86.

LESAGE, James P. An introduction to spatial econometrics. Revue d'économie industrielle, 2008, no 123, p. 19-44.

LI, Lei, et al. Rolling window time series prediction using MapReduce. En Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014). IEEE, 2014. p. 757-764.

MAYNOU, Laia, et al. Health inequalities in the European Union: an empirical analysis of the dynamics of regional differences. The European Journal of Health Economics, 2015, vol. 16, no 5, p. 543-559.

MAYNOU, Laia; SAEZ, Marc. Economic crisis and health inequalities: evidence from the European Union. International journal for equity in health, 2016, vol. 15, no 1, p. 135.

MODRANKA, Emilia; SUCHECKA, Jadwiga. The determinants of population health spatial disparities. Comparative Economic Research, 2014, vol. 17, no 4, p. 173-185.

MORENO SERRANO, Rosina, et al. Econometría espacial: nuevas técnicas para el análisis regional. Una aplicación a las regiones europeas. 2002.

SALA-I-MARTIN, Xavier X. The classical approach to convergence analysis. The economic journal, 1996, p. 1019-1036.

STAŃCZYK, Renata. Convergence of Health Status in the European Union: A Spatial Econometric Approach. 2015.

ULLAH, A.; GILES, D. E. A.; MARCEL DEKKER, N. Y. Spatial Dependence in linear Regression Models with an Introduction to Spatial Econometrics. 1998.

VILALTA, Carlos Javier, et al. Cómo enseñar autocorrelación espacial. Economía, Sociedad y Territorio, 2005, no 18, p. 323-333.

WORLD BANK OPEN DATA 2020, Life expectancy at birth, total (years). Available in: https://data.worldbank.org/indicator/SP.DYN.LE00.IN

WORLD HEALTH ORGANIZATION 2020, Disability-adjusted life expectancy. Available in: https://gateway.euro.who.int/en/indicators/hfa_67-1080-disability-adjusted-life-expectancy-world-health-report/.

ZEREN, Fatma; ÖZCAN, Burcu; MENTEŞE, Emin Yahya. Health care convergence analysis in Turkey on the province level: Spatial quantile method. Procedia economics and finance, 2016, vol. 38, p. 90-97.