



**UNIVERSITAT
JAUME I**

**INVESTING IN HEALTH: THE RELATIONSHIP BETWEEN PUBLIC HEALTH
SPENDING AND HEALTH IN THE EU**

Jennifer González López

al375069@uji.es

Grado en Economía

2022-2023

Arne Risa Hole

1. Abstract

In the following study, I analyze the relationship between healthcare expenditure and life expectancy in 31 European Union member countries, excluding Norway and Iceland, from 2011 to 2020. Using time series and cross-sectional data, four variables were examined: healthcare expenditure as a percentage of GDP, per capita GDP, poverty, and R&D expenditure. Life expectancy was used as a health indicator. Four different models were estimated using the panel data technique, and the random effects estimation was selected as the most efficient. The results indicate a positive relationship between healthcare expenditure and life expectancy in the European Union.

- I12: Health Production
- I18: Health: Government Policy; Regulation; Public Health
- C23: Panel Data Models; Spatio-temporal Models
- C33: Multiple or Simultaneous Equation Models: Panel Data Models; Spatio-temporal Models
- H51: Government Expenditures and Health

ÍNDEX

1. Abstract (p. II)
2. Introduction (p. V)
 - 2.1. Literature Review. (p. VI)
 - 2.2. Conceptual Framework and Initial Hypotheses (p. IX)
3. Objectives. (p. X)
4. Data and Methodology. (p.XI)
 - 4.1. Description of Variables. (p.XI)
 - 4.2. Database. (p.XIII)
 - 4.3. Univariate Statics. (p.XIV)
 - 4.4. Multivariate Statics (p.XV)
 - 4.5. Econometric Model. (p.XVI)
5. Results. (p.XVII)
 - 5.1. Estimates using different methodologies. (p.XVII)
 - 5.1.1. Pooled OLS estimation. (p.XVII).
 - 5.1.2. Fixed Effects estimation. (p.XIX)
 - 5.1.3. Pooled OLS vs FE (p. XXI)
 - 5.1.4. Individual and Time Fixed Effects estimation. (p.XXII)
 - 5.1.5. Random Effects estimation. (p.XXIV)
 - 5.2. Summary of the obtained results. (p. XXV)
 - 5.3. Specification Selection. (p. XXVI)
 - 5.3.1. Breusch-Pagan test. (p.XXVI)
 - 5.3.2. Hausman test. (p.XXVII)
6. Conclusions. (p. XXVIII)
 - 6.1. Econometric Conclusions. (p. XXVIII)
 - 6.2. Economic Conclusions. (p. XXX)
7. References. (p. XXXI)

TABLES AND FIGURES INDEX

Table 1. Literature review.

Table 2. Main univariate statistics.

Table 3. Multivariate statistics.

Table 4. Expected signs in the econometric study.

Table 5. Pooled OLS estimation.

Table 6. Fixed Effects estimation.

Table 7. Pooled OLS vs Fixed Effects estimation.

Table 8. Two-way fixed effects estimation.

Table 9. Random Effects Estimation.

Table 10. Summary table of the obtained results.

Table 11. Breusch Pagan test.

Table 12. Hausman test.

Picture 1. Database Created.

2. Introduction.

Public healthcare expenditure is one of the main areas of public investment in the European Union (EU). Most EU countries have publicly funded healthcare systems, which are responsible for the majority of healthcare services. However, in a context of budget constraints and increasing healthcare needs, it is essential to understand how public healthcare expenditure affects people's health.

In this regard, numerous econometric studies have analyzed the relationship between public healthcare expenditure and health indicators in EU countries. Overall, the literature suggests that higher public healthcare expenditure is associated with better health outcomes, such as increased life expectancy and lower mortality rates. However, the results of these studies vary across countries and the health indicators analyzed.

In addition, some econometric studies have also investigated the effectiveness of public health spending in the EU. In other words, they have tried to assess whether higher public health spending translates into a proportional improvement in health indicators. The results of these studies have been mixed, suggesting that the effectiveness of public health spending may depend on contextual factors and the way in which it is distributed and used.

According to data from the Organization for Economic Cooperation and Development (**OECD**), the countries that invest the most in public health spending in Europe are Switzerland, Norway and Luxembourg, while the countries that invest the least are Romania, Latvia and Poland. However, it is important to keep in mind that public health spending is not the only factor influencing the health of the population, and that other factors such as the quality of medical care and disease prevention also play an important role.

In addition to public health spending, there are other factors that influence people's health. Some of the main factors are access to quality health services, educational level, income, employment, housing, and the social and physical environment. These factors are interconnected and affect health in complex and interdependent ways.

Health encompasses subjective aspects (physical, mental and social well-being), objectives (functioning capacity) and social aspects (adaptation and socially productive work), therefore, it is a resource for daily life, not the objective of life. It is a positive concept that emphasizes social and personal resources, as well as physical aptitudes. (De La Guardia, 2020).

For example, access to quality health services is essential to prevent and treat disease, but educational level and income can influence people's ability to access these services and follow prescribed treatments. Employment and adequate housing can also affect health, as job insecurity or homelessness can increase stress and anxiety, which can lead to mental and physical health problems.

Therefore, understanding how these factors interact and contribute to health is essential to address health inequalities and improve people's quality of life. Econometric studies can help to identify the relationship between these factors and health, and to quantify the importance of each of them in determining the health of the population.

In general, countries that invest more in public health spending tend to have **better life expectancy, lower mortality rates**, and higher satisfaction with the health care system, while countries that invest less in health often have worse general health and limited access to medical care.

In conclusion, the review of the literature on how public health spending affects the health of people in the EU, and the evaluation of its effectiveness, are **relevant issues for decision-making in public health policies**. Econometric studies offer a valuable tool to empirically analyze this relationship and provide evidence to improve the effectiveness and equity in the allocation of public resources for health.

2.1. Literature Review.

Literature review is a fundamental tool in any research, especially in the field of econometrics. The literature review allows researchers to find out about previous works related to the subject of study and establish the theoretical and conceptual framework necessary to develop their research.

In econometric studies, the literature review is crucial to identify the econometric models previously used in the literature, as well as to **select the appropriate variables and appropriate statistical analysis techniques** for data analysis. In addition, it allows researchers to assess the quality and relevance of previous studies, identify gaps in the existing literature, and establish new research questions.

The relationship between public health spending and health is a widely researched topic in the economic literature. Most studies focus on examining how public health spending affects life expectancy and other health indicators. In this sense, the use of econometric techniques has become very common to try to measure the causal relationship between public health spending and health.

In this review of the literature, the results of different econometric studies that have examined the relationship between public health spending and health in various countries of the world are analyzed. For this, articles have been searched in **Google Academics** that use econometric techniques to measure the causal relationship between these two variables.

The aim of this literature review is to provide an overview of the **most relevant findings in the economic literature** on how public health spending affects health. In addition, it is intended to identify **the limitations and possible gaps** in the current literature to guide future research in this area. In short, this review of the literature is a useful tool for those interested in understanding the relationship between public health spending and health from an economic perspective.

| <p>Title and complete reference.</p> <p>Authors Magazine, Volume (Number) and pages.</p> | <p>Objective of the article and relationship with our work.</p> | <p>Dependent and independent variables used.</p> | <p>Scope of work and methodology.</p> | <p>Most notable results related to the hypotheses raised in our work.</p> |
|--|--|---|--|--|
| <p>“Public spending on health and infant mortality in Mexico: time series analysis.”</p> <p>Sandoval-Mendoza, Teresita de Jesús; Salinas-Rodríguez, Aarón.</p> <p>Salud pública de México, vol.60, n.2, 2018, pp.240-248. ISSN 0036-3634.</p> <p>Salud pública de México Volumen: 60 Número: 2 Páginas: 240-248</p> | <p>The objective of the article is to analyze the relationship between public spending on health and infant mortality in Mexico during the period 1990-2014, by applying a time series regression model.</p> | <p>The dependent variable used in the study is the infant mortality rate in Mexico, while the independent variable is public spending on health as a percentage of GDP.</p> | <p>The scope of the work focuses on the analysis of the relationship between public spending on health and infant mortality in Mexico, using time series data for the period 1990-2014.</p> <p>For the analysis, a time series regression model is used that includes the infant mortality rate as a dependent variable and public spending on health as a percentage of GDP as an independent variable.</p> | <p>The results of the study indicate that there is an inverse relationship between public spending on health and infant mortality in Mexico during the study period.</p> <p>The increase in public spending on health was associated with a significant decrease in the infant mortality rate. In addition, the study found that the relationship between public spending on health and infant mortality was stronger in the regions of the country with higher levels of poverty and marginalization.</p> <p>The authors suggest that greater investment in health is required to improve child health in Mexico, especially in the poorest and most marginalized regions of the country.</p> |
| <p>"The effects of demographic aging on health spending: myths and realities"</p> <p>Pilar García-Gómez, Ángel López Nicolás, y Victoria Rodríguez.</p> <p>Health Gazette Volume 15, Issue 2, 2001, Pages 154-163. 3 April 2013</p> | <p>The objective of this article is to analyze the myths and realities related to the impact of demographic aging on health spending in Spain.</p> <p>The authors examine the common belief that demographic aging is primarily responsible for the increase in healthcare spending.</p> | <p>The authors use total health expenditure as the dependent variable and the population aged 65 and over as the independent variable.</p> <p>Other variables such as per capita income, the number of doctors and nurses, and health coverage are also considered.</p> | <p>This article is based on a review of the literature on the impact of demographic aging on health spending in Spain and in other countries.</p> <p>The authors use an econometric model to analyze the relationship between the population aged 65 and over and total health spending. They also perform a descriptive analysis of data on health spending and population.</p> | <p>The authors conclude that demographic aging is not the main factor explaining the increase in healthcare spending in Spain.</p> <p>In addition, the authors highlight the importance of improving the efficiency in the use of health resources and the need for public policies to address demographic aging.</p> |

| Title and complete reference. Authors Magazine, Volume (Number) and pages. | Objective of the article and relationship with our work. | Dependent and independent variables used. | Scope of work and methodology. | Most notable results related to the hypotheses raised in our work. |
|--|--|---|--|--|
| <p>“Public health spending and life expectancy: a healthy investment.”</p> <p>Elisenda Rentería Pérez</p> <p>Pilar Zueras Castillo</p> <p>Perspectives demográfiques.</p> <p>OCT 2022 N° 029</p> | <p>The article analyzes the evolution of the years lived with and without the most common diseases from the age of 50 in 17 Autonomous Communities of Spain, during the period 2006-2019, in order to understand their association with macro socioeconomic factors, public provision health services, public health spending and behaviors related to health.</p> <p>The objective of the study is to observe how various regional economic and sociodemographic factors may be associated with changes in health indicators.</p> | <p>Dependent variable: years lived with and without chronic conditions from the age of 50 in the 17 Autonomous Communities of Spain.</p> <p>Independent variable: various macro socioeconomic factors, public provision of health services, public health spending, and behaviors related to health, with per capita public health spending being the one that best explains the differences in health between regions.</p> | <p>The period of analysis extends from 2006 to 2019, in order to capture the phases of expansion, crisis and economic recovery.</p> <p>The scope of the work is to analyze the years lived with and without disabling diseases in relation to various socioeconomic factors in 17 autonomous communities of Spain and in a period of time that includes the economic crisis and subsequent recovery.</p> | <p>The results show that per capita public health spending is one of the components that best explains the differences in health between regions and that higher public health spending implies more years lived in good health and fewer years lived in poor health.</p> <p>In addition, it is observed that during the economic recession between 2008-2014, life expectancy in Spain did not stop growing and the greatest benefits in mortality were in the most disadvantaged social classes. However, the regional heterogeneity of Spain and its decentralized public healthcare produced a growing heterogeneity of life expectancies in good and poor health by Autonomous Community in a context of economic crisis.</p> |

Table 1. Literature review.

2.2. Conceptual framework and starting hypothesis.

Health spending is known to be **an important determinant** of overall health. If health systems are capable of providing quality care to the population, it is possible to prevent and treat diseases, reduce mortality and improve quality of life. Life expectancy, on the other hand, is a widely used health indicator that measures the number of years a person can expect to live in a given country.

Therefore, higher per capita health spending is expected to imply **more years of life expectancy in health and fewer years of life expectancy in ill health**. (Renteria and Zueras, 2022). However, this relationship is not necessarily linear and can be affected by other factors, such as income level, education, lifestyle, and environmental factors.

The "**research question**" or research question is a question that is formulated with the aim of orienting and guiding the investigation. This question is based on the problem to be solved, and serves to define the scope and direction of the investigation.

In this case, the research question is: **How does health spending in the European Union affect people's life expectancy?**

In this case, the null and alternative hypotheses are as follows:

- **Null hypothesis (H0):** There is no relationship between health spending and life expectancy in the European Union.
- **Alternative hypothesis (H1):** There is a positive relationship between health spending and life expectancy in the European Union.

The null hypothesis is stated as the initial statement to be tested and is assumed to be true until the data shows otherwise. In this case, H0 states that there is no relationship between health spending and life expectancy in the European Union. The alternative hypothesis is established as the statement contrary to the null hypothesis and is the one that is tried to be demonstrated. In this case, H1 suggests that there is a positive relationship between health spending and life expectancy in the European Union.

It is important to note that normally rejecting the null hypothesis accepts the alternative hypothesis, even though the two hypotheses are complementary and not exclusive, which means that in some cases this might not be the case.

In this case, the alternative hypothesis would be accepted if sufficient statistical evidence is found to conclude that health spending has a significant effect on the life expectancy of people in the European Union. The null hypothesis would be rejected in that case. On the other hand, the null hypothesis would be accepted if there is insufficient evidence to reject it. This would mean that there is not enough statistical evidence to affirm that there is a significant relationship between health spending and the life expectancy of people in the European Union.

3. Objectives.

Econometric studies are key tools for understanding the relationship between economic and social variables and how they influence each other. The main goal of an econometric study is to **analyze existing data to develop statistical models** that can be used to predict future outcomes or identify causal relationships between different factors.

Another important objective is **to test hypotheses and evaluate the validity of the proposed models**. To achieve these goals, econometric studies often use advanced

data analysis techniques, including multiple regressions, time series analysis, and hypothesis testing.

In general, econometric studies seek to provide useful and valuable information for decision-making in fields such as economic policy, business management, and academic research. According to Herbert Simon, decision makers do not always seek to maximize utility or economic benefits, but also take other factors into account, such as political, social, and cognitive constraints. Therefore, it is important that econometric studies are carried out with methodological rigor and that reliable and up-to-date databases are used.

The main objective of this study is **to analyze the relationship between spending on public health and the health of people in the countries of the European Union**, as well as to identify other relevant variables that may affect this relationship. To do this, the **Stata software** will be used and a database with information from 31 EU countries between 2011 and 2020 will be used. The econometric study will therefore make it possible to assess the importance of spending on public health and other determining variables in the population health in the European Union.

4. Data and Methodology.

4.1. Variable description.

A panel data series study refers to an empirical analysis that uses time series data and cross-sectional data to examine the relationship between variables. In this case, the study has been carried out in a chronological period of time from 2011 to 2020, using data from 31 member countries of the European Union, excluding Norway and Iceland.

The panel data technique is very useful in economics, since it allows the analysis of the interactions between variables over time and between different groups or countries. In the case of this study, it has been used to investigate the relationship between spending on public health and the health of people in the countries of the European Union, as well as other variables that have been considered relevant.

A panel data series study refers to an empirical analysis that uses time series data and cross-sectional data to examine the relationship between variables. In this case, the study has been carried out in a chronological **period of time from 2011 to 2020, using data from 31-member countries of the European Union, excluding Norway and Iceland.**

Norway is a Scandinavian country that is part of the European Economic Area and has an association agreement with the European Union. Iceland, for its part, is an island country located in the North Atlantic and also has an association agreement with the European Union, but is not part of the European Economic Area or the European Union.

The sample of countries included in the study is: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, the Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, the United Kingdom, Iceland, Norway and Switzerland.

To carry out this study I have used five variables:

- **Life expectancy measured in years of life (healthy_years):** This variable refers to the number of years that a person in a given country is expected to live. It is calculated using mortality data and life expectancy at each age. This measure is used as an indicator of the general health of a population.
- **Health expenditure in terms of GDP (health_exp_gdp):** This variable measures total health expenditure (public and private) as a percentage of a country's GDP. This indicator is used to assess the level of priority given to health in a country and its impact on the economy.
- **GDP per capita measured in euros per capita (gdp_per_capita):** GDP per capita is the total value of goods and services produced in a country divided by its population. It is a measure commonly used to compare the level of wealth and well-being between countries.
- **Poverty measured as a percentage (poverty):** This variable refers to the percentage of the population that lives below the poverty line. The poverty line is defined as the income below which a person cannot meet their basic needs.
- **R&D spending as a percentage of GDP (id):** This variable measure total research and development spending (public and private) as a percentage of a country's GDP. It is used to assess the level of investment in science and technology and its impact on economic and social development.

In this case, the **independent variable** would be health spending in terms of GDP, GDP per capita measured in euros per capita, poverty measured as a percentage, and R&D spending as a percentage of GDP, since these are the variables that are they will manipulate or analyze to see their effect on health. The **dependent variable** would be

health, which in this case can be measured in different ways, such as life expectancy measured in years of life. This is because health is the variable that is expected to be influenced by health spending in terms of GDP, and therefore it is the variable that one wants to measure and analyze based on the independent variable.

4.2. Database.

Eurostat is the statistical office of the European Union, in charge of compiling and publishing data and statistics on various topics that affect the Member States of the EU. In my study, I have used data from this source to analyze the relationship between health spending in terms of GDP and life expectancy in 31 European Union countries, excluding Norway and Iceland.

By collecting and organizing these data, I perform an econometric analysis to investigate the relationship between these important variables and obtain meaningful results. With the help of the obtained database, I carry out a more detailed and rigorous analysis of the relationship between health spending and life expectancy in the EU. On the other hand, thanks to Excel, I collect and organize the data of 310 observations, corresponding to the period from 2011 to 2020 for each of the countries.

Below is the database that I will use during the econometric study. As the database is very extensive, observations from two countries, Belgium and Bulgaria, are shown below.

| | A | B | C | D | E | F | G | H | I |
|----|------|-----|---|----------|---------------|----------------|----------------|---------|------|
| 1 | t | i | c | i_id | healthy_years | health_exp_gpd | gdp_per_capita | poverty | id |
| 2 | 2011 | Bel | 1 | Belgium | 63,5 | 10,36 | 33460 | 8,8 | 2,17 |
| 3 | 2012 | Bel | 1 | Belgium | 64,6 | 10,48 | 33490 | 8,3 | 2,28 |
| 4 | 2013 | Bel | 1 | Belgium | 63,9 | 10,55 | 33490 | 8,8 | 2,33 |
| 5 | 2014 | Bel | 1 | Belgium | 64,1 | 10,58 | 33870 | 7,9 | 2,37 |
| 6 | 2015 | Bel | 1 | Belgium | 64,2 | 10,77 | 34360 | 8,6 | 2,43 |
| 7 | 2016 | Bel | 1 | Belgium | 63,7 | 10,76 | 34620 | 9,3 | 2,52 |
| 8 | 2017 | Bel | 1 | Belgium | 63,7 | 10,75 | 35050 | 8,6 | 2,67 |
| 9 | 2018 | Bel | 1 | Belgium | 63,4 | 10,79 | 35510 | 8,1 | 2,86 |
| 10 | 2019 | Bel | 1 | Belgium | 62,4 | 10,66 | 36110 | 8,9 | 3,16 |
| 11 | 2020 | Bel | 1 | Belgium | 63,8 | 11,06 | 34010 | 7,8 | 3,35 |
| 12 | 2011 | Bul | 2 | Bulgaria | 64 | 7,1 | 5320 | 27,8 | 0,53 |
| 13 | 2012 | Bul | 2 | Bulgaria | 63,9 | 7,54 | 5390 | 32,8 | 0,6 |
| 14 | 2013 | Bul | 2 | Bulgaria | 64,5 | 7,14 | 5390 | 32,9 | 0,63 |
| 15 | 2014 | Bul | 2 | Bulgaria | 64 | 7,68 | 5470 | 31,7 | 0,79 |
| 16 | 2015 | Bul | 2 | Bulgaria | 63,2 | 7,39 | 5700 | 30,6 | 0,95 |
| 17 | 2016 | Bul | 2 | Bulgaria | 65,7 | 7,46 | 5910 | 28,9 | 0,77 |
| 18 | 2017 | Bul | 2 | Bulgaria | 64,5 | 7,49 | 6120 | 28 | 0,74 |
| 19 | 2018 | Bul | 2 | Bulgaria | 65,8 | 7,33 | 6330 | 26,5 | 0,75 |
| 20 | 2019 | Bul | 2 | Bulgaria | 66,3 | 7,09 | 6630 | 23,7 | 0,83 |
| 21 | 2020 | Bul | 2 | Bulgaria | 65,6 | 8,52 | 6410 | 18,6 | 0,85 |
| 22 | 2011 | CR | 3 | Czechia | 62,9 | 8,52 | 15310 | 8,7 | 1,54 |
| 23 | 2012 | CR | 3 | Czechia | 62,2 | 8,52 | 15170 | 8,2 | 1,77 |

Picture 1. Database Created.

4.3. Univariate Statistics.

Univariate statistics are numerical measures that are applied to a single variable and provide information about its characteristics and distribution. Table 1 below shows the main univariate statistics (mean, median, minimum and maximum values, and standard deviation) of the variables included in our model to be estimated.

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|----------------|-----|-----------|-----------|-------|--------|
| Healthy_years | 310 | 62,2154 | 4,8673 | 51,4 | 75,1 |
| Health_exp_gdp | 310 | 8,5921 | 1,8767 | 4,7 | 12,82 |
| Gdp_per_capita | 310 | 28.651,32 | 18.372,08 | 5.320 | 84.750 |
| poverty | 310 | 10,6715 | 8,9212 | 1,4 | 40,6 |
| id | 310 | 1,5651 | 0,8711 | 0,38 | 3.62 |

Table 2. Main univariate statistics.

For the life expectancy variable, it can be seen that the mean is 62.2 years, with a standard deviation of 4.867, which indicates that most of the observations are in a range close to the mean. The minimum is 51.4 years and the maximum is 75.1 years.

For the health spending variable in terms of GDP, it can be seen that the mean is 8.592, with a standard deviation of 1.8786, indicating that there is significant variability in health spending across countries. The minimum expenditure is 4.7% of GDP and the maximum is 12.82%.

Regarding GDP per capita, the average is 28,651.32 euros per capita, with a standard deviation of 18,372.08, which indicates great variability between countries in terms of their economic wealth. The minimum GDP per capita is 1.4 euros and the maximum is 40.6 euros.

In the poverty variable measured as a percentage, it can be observed that the average is 10.67%, with a standard deviation of 8.92, which indicates a great variability in the poverty rate between countries. The minimum value is 1.4% and the maximum is 40.6%.

For the R&D spending variable as a percentage of GDP, it can be seen that the mean is 1.565%, with a standard deviation of 0.871, indicating moderate variability in R&D spending across countries. The minimum expenditure is 0.38% of GDP and the maximum is 3.62%.

In general, these univariate statistics provide an overview of the distribution of the data and can be helpful in identifying possible patterns or trends in the data.

4.4. Multivariate Statistics.

The **correlation matrix** is a table that shows the Pearson correlation coefficients between all the variables included in an analysis. Each cell in the table represents the correlation coefficient between two variables, which varies in a range from -1 to 1.

A positive correlation coefficient indicates a direct relationship between two variables, which means that increases in one variable are associated with increases in the other variable. On the other hand, a negative correlation coefficient indicates an inverse relationship, which means that increases in one variable are associated with decreases in the other variable. A correlation coefficient close to zero indicates that there is no relationship between the two variables.

The correlation matrix is important in econometric analyzes because it allows to identify the strength and direction of the relationships between the variables included in the model. In addition, it is also useful for detecting possible multicollinearity problems, which occurs when two or more variables are highly correlated with each other, which can affect the accuracy of the model results. Thus, Table 2 shows the correlation matrix between the previously mentioned variables:

| | Healthy_years | Health_exp_gdp | Gdp_per_capita | poverty | id |
|----------------|---------------|----------------|----------------|---------|------|
| Healthy_years | 1.00 | | | | |
| Health_exp_gdp | 0,2925 | 1.00 | | | |
| Gdp_per_capita | 0,3393 | 0,3503 | 1.00 | | |
| poverty | -0,0010 | -0,3618 | -0,5297 | 1.00 | |
| id | 0,0381 | 0,6227 | 0,3652 | -0.5241 | 1.00 |

Table 3. Multivariate statistics.

In this case, we can observe that life expectancy is weakly correlated with health spending (0.2925) and moderately correlated with GDP per capita (0.339).

Health spending and GDP per capita also show a moderate correlation (0.350). On the other hand, poverty shows a weakly negative correlation (-0.001) with life expectancy, a moderately negative correlation with GDP per capita (-0.5297) and a weakly negative correlation with R&D spending (-0.5241).

Finally, R&D spending shows a moderate correlation with health spending (0.6227) and with GDP per capita (0.3652).

In summary, the correlation matrix allows us to understand the relationship between the different variables and how they relate to each other.

4.5. Econometric model.

The econometric model to be created has life expectancy as the dependent variable and health spending in terms of GDP, GDP per capita measured in euros per capita, poverty measured as a percentage, and R&D spending as independent variables. The database used for this analysis consists of information from 31 countries during the years 2011 to 2020.

$$\text{Healthy_years}_{it} = \beta_0 + \beta_1 \text{health_exp_gdp}_{it} + \beta_2 \text{gdp_per_capita}_{it} + \beta_3 \text{poverty}_{it} + \beta_4 \text{id}_{it} + \alpha_i + \varepsilon_{it}$$

$$i=1,2,3, \dots, 31 \text{ countries} \quad t=2011,2012,\dots,2020.$$

The econometric model that I have written is a **multiple linear regression model** that tries to explain the variable "Healthy_years" (years of healthy life) based on four other independent variables: "health_exp_gdp" (health expenditure in terms of GDP), "gdp_per_capita" (GDP per capita), "poverty" (poverty measured as a percentage) and "id" (R&D expenditure as a percentage of GDP).

The model states that the value of the variable "Healthy_years" is determined by a linear combination of the other four independent variables, and fits a linear trend over time (α it is the error term).

The notation $i=1,2,3,\dots,31$ countries indicates that the model is applied to a set of 31 countries, while $t=2011,2012,\dots,2020$ indicates that data from the years are considered. 2011 to 2020.

In accordance with the econometric theory and the study carried out with the previous literature, the signs that we expect to obtain for the coefficients that accompany the dependent variable are explained in the following table:

| Dependent variable | Independent variable | Economic intuition |
|---|--|--|
| Life expectancy measured in years of life (healthy_years) | Health expenditure in terms of GDP (health_exp_gdp) | $\beta > 0$. Positive sign, since higher health spending is expected to be associated with longer life expectancy, as this may mean that the population has improved access and quality of medical care. |
| | GDP per capita measured in euros per capita (gdp_per_capita) | $\beta > 0$. Positive sign. Higher GDP per capita is expected to be associated with longer life expectancy, as this may mean that the population has access to better living conditions, food and services, as well as better medical care. |
| | Poverty measured in percentage (poverty) | $\beta < 0$. Negative sign. A higher level of poverty is expected to be associated with a lower life expectancy, as poverty can lead to a lower quality of life, lack of access to health care, and increased risks of disease. |

| | | |
|--|---|---|
| | R&D spending as a percentage of GDP (id) | $\beta > 0$. Positive sign. Higher R&D spending is expected to be associated with longer life expectancy, as it can lead to medical and technological discoveries that improve the health and quality of life of the population. |
|--|---|---|

Table 4. Expected signs in the econometric study.

5. Results.

After specifying the econometric model, the next step would be its estimation. Thus, I will use various estimation methods (**Fused Ordinary Least Squares (OLS), Fixed Effects, Individual and Temporary Fixed Effects, and Random Effects**), to end by discussing which of them is appropriate.

5.1. Estimates with the different methodologies.

5.1.1. Estimation by Fused OLS (Pooled OLS).

A first approximation would be to estimate by Fused Ordinary Least Squares. The **OLS technique (Ordinary Least Squares)** merged or pooled OLS, is used to estimate the coefficients of a regression when the data come from multiple samples that are combined in a single sample. This technique is used when it is assumed that the relationship between the independent variables and the dependent variable is the same in all samples.

To do this, using the statistical software Stata, using robust standard deviations to heteroscedasdity, the output represented in the following table is obtained:

$$\mathbf{Healthy_years}_{it} = \beta_0 + \beta_1 \mathbf{health_exp_gdp}_{it} + \beta_2 \mathbf{log_gdpcap}_{it} + \beta_3 \mathbf{poverty}_{it} + \beta_4 \mathbf{id}_{it} + U_{it}$$

$$i=1,2,3,\dots,31 \text{ countries } t=2011,2012,\dots,2020.$$

Where:

- **Healthy_years_{it}** represents the dependent variable for the *it*h unit (country) in time period *t*.
- **x_{1_it}, x_{2_it}, ..., x_{k_it}** are the *k* independent variables for the *it*h unit in time period *t*.
- **β₀** is the intercept of the model.
- **β₁, β₂, ..., β_k** are the coefficients of the independent variables to be estimated.
- **u_{it}** is the random error term for the *it*h unit in time period *t*.

| healthy_years | Coef. | Robust Std. Err. | t | P>t | [95% Conf. | Interval] |
|----------------|-----------|------------------|-------|-------|------------|-----------|
| health_exp_gdp | 0.7121375 | .1692607 | 4.21 | 0.000 | 0.379071 | 1.045204 |
| log_gdpcap | 8.269625 | 1.015868 | 8.14 | 0.000 | 6.270628 | 10.26862 |
| poverty | 0.1434974 | 0.0285631 | 5.02 | 0.000 | .0872918 | 0.199703 |
| id | -1.288059 | 0.5077617 | -2.54 | 0.012 | -2.287219 | -0.2889 |
| _cons | 20.45036 | 4.278346 | 4.78 | 0.000 | 12.03155 | 28.86918 |

Table 5. Pooled OLS estimation.

The results of the model show that all the independent variables are **statistically significant at a significance level of 5%**. In the result of the Pooled OLS estimation, it can be seen that all the independent variables have a p value less than 0.05, which indicates that they are statistically significant at a significance level of 5%. This means that the estimated coefficients of these variables are unlikely to be zero in the population from which the sample was drawn. In other words, **there is a significant relationship between the dependent variable (healthy_years) and the independent variables (health_exp_gdp, log_gdpcap, poverty, id) at the 5% significance level.**

Regarding the coefficients, it is observed that the coefficient of the variable health_exp_gdp is 0.712, which suggests that a 1% increase in health spending in relation to GDP is associated with an increase of 0.712 years in life expectancy, keeping the other variables of the model constant.

The coefficient of the variable log_gdp_cap is 8,270, indicating that a 1% increase in GDP per capita is associated with an increase of 8,270 years in life expectancy, holding the other variables in the model constant.

On the other hand, the coefficient of the "poverty" variable is 0.143, which suggests that a 1% increase in the poverty rate is associated with an increase of 0.143 years in life expectancy, keeping the other variables constant. model.

Finally, the coefficient of the variable "id" is -1.288, which indicates that higher spending on R&D is associated with a decrease in life expectancy in the countries in the sample.

Regarding the intercept (constant), it is observed that its value is 20,450, which indicates that life expectancy in the absence of the independent variables would be 20,450 years.

It is important to highlight that the "id" variable turned out to be significant and with a negative effect on life expectancy, which could indicate that spending on R&D may have unwanted effects on the health of the population. It is important to delve into this aspect in order to better understand this relationship and to be able to take measures to improve the health of the population in the future.

The coefficient of the variable "poverty" in the econometric model is positive, meaning that the increase in poverty is associated with an increase in life expectancy instead of a decrease. This may seem counterintuitive, but it's important to remember that correlation does not imply causation.

There may be other factors at play that are affecting the relationship between poverty and life expectancy. For example, countries with higher poverty rates may also have stronger health systems or more effective social policies that offset the negative effect of poverty on health. It is important to consider these possible alternative explanations before jumping to conclusions.

In any case, the fact that the coefficient is positive does not necessarily mean that the "poverty" variable is not relevant in the model. **It may continue to be an important variable in explaining life expectancy**, even if its effect on it is different from what was expected.

5.1.2 Fixed Effects Estimation.

Fixed Effects (FE) estimation is an econometric estimation method that allows controlling for unobservable variables that are constant over time but vary between units (countries, in this case) in the panel data.

In this case, the model would look like:

$$Healthy_years_{it} = \beta_0 + \beta_1 health_exp_gdp_{it} + \beta_2 log_gdpcap_{it} + \beta_3 poverty_{it} + \beta_4 id_{it} + \alpha_i + U_{it}$$

$$i=1,2,3,\dots,31 \text{ countries } t=2011,2012,\dots,2020.$$

To carry out the FE estimate, the transformed ordinary least squares method (within) is used, which consists of subtracting the mean of each country from each variable, so that the fixed effects of the independent variables are eliminated and the regression coefficients are obtained. showing the relationship between the independent variables and the dependent variable after controlling for country fixed effects.

It is important to highlight that the FE **model only allows to identify the effects of variables that vary over time** (within each country) and does not allow to identify the effects of variables that vary between countries.

| healthy_years | Coef. | Robust Std. Err. | t | P>t | [95% Conf. Interval] |
|----------------|------------|------------------|-------|-------|----------------------|
| health_exp_gpd | 0.3012316 | 0.3119297 | 0.97 | 0.342 | -0.3358138 0.938277 |
| log_gdpcap | 2.088913 | 10.08424 | 0.21 | 0.837 | -18.50584 22.68367 |
| poverty | -0.0179138 | 0.1032761 | -0.17 | 0.863 | -0.2288318 0.1930042 |
| id | 0.3110838 | 0.7766187 | 0.40 | 0.692 | -1.274983 1.897151 |
| _cons | 50.20482 | 47.04466 | 1.07 | 0.294 | -45.8732 146.2828 |
| sigma_u | 4.3292845 | | | | |
| sigma_e | 1.956452 | | | | |
| rho | 0.83040854 | | | | |

Table 6. Fixed Effects estimation.

In the Fixed Effects econometric study, the linear panel model is used to estimate the relationship between the dependent variable, "healthy_years", and the independent variables "health_exp_gpd", "log_gdpcap", "poverty" and "id". We control for fixed effects to account for unobservable heterogeneity over time.

The estimated coefficient for the variable "health_exp_gpd" is 0.3012, with a robust standard error of 0.3119. However, **this coefficient does not turn out to be statistically significant** with a p-value of 0.342, which means that the null hypothesis that the variable "health_exp_gpd" does not have a significant effect on the dependent variable "healthy_years" at a level cannot be rejected. of significance of 5%.

The coefficient for the variable "log_gdpcap" is 2.0889, with a robust standard error of 10.0842. This coefficient is also not statistically significant with a p value of 0.837, which suggests that there is not enough evidence to affirm that per capita income has a significant effect on health.

The coefficient for the "poverty" variable is -0.0179, with a robust standard error of 0.1033. As in the previous cases, the coefficient does not turn out to be statistically significant with a p value of 0.863. That is, it cannot be concluded that the poverty rate has a significant effect on the dependent variable "healthy_years".

The coefficient for the variable "id" is 0.3111, with a robust standard error of 0.7766. This coefficient is also not statistically significant with a p value of 0.692, suggesting that R&D spending as a percentage of GDP does not have a significant effect on health.

Regarding the error parameters, sigma_u gives a value of 4.3293, which indicates that the unobservable variation in time that was not captured by the fixed effects represents approximately 82.5% of the total variation in the dependent variable "healthy_years". Sigma_e, the error parameter for the error term, has a value of 1.9565, indicating that the unexplained variation in the dependent variable "healthy_years" by the independent

variables and fixed effects represents approximately 17.5% of the total variation. The rho value was 0.8304, which indicates that the correlation between unobservable errors over time is high.

In summary, the results of the Fixed Effects econometric model suggest that **none of the independent variables studied have a statistically significant effect** on the dependent variable "healthy_years". Furthermore, the values of sigma_u, sigma_e and rho indicate that there is significant unobserved variability over time and a high correlation between unobservable errors over time. These findings suggest that there may be other unmeasured factors affecting health that should be considered in future analyses.

5.1.3. Comparative Pooled OLS vs FE.

The following table presents a comparison between the results obtained by both estimation methods:

| Variable | MCO | FE |
|----------------|------------|------------|
| health_exp_gdp | 0.71213746 | 0.30123163 |
| log_gdpcap | 8.269625 | 2.0889126 |
| poverty | 0.14349743 | -.01791377 |
| id | -1.2880595 | 0.31108379 |
| _cons | 20.450364 | 50.204819 |

Table 7. Pooled OLS vs Fixed Effects estimation.

The main differences between the Pooled OLS study and the Fixed Effects study are due to the treatment of variables that **do not vary over time**. In the Pooled OLS model, all variables are treated as if they were independent variables without considering the possibility that some of the effects are fixed in time. Therefore, all effects are assumed to be constant for all individuals over time. On the other hand, in the Fixed Effects model, the individual effects are considered and the fixed effects not observed in time are eliminated.

In terms of model output, the **Fixed Effects model tends to have lower coefficients than the Pooled OLS model**. This is because the Fixed Effects model removes the variance of unobserved individual effects over time, which increases the precision of the model and reduces the possibility of specification errors. Furthermore, in the Fixed Effects model, the errors are not correlated with the independent variables, which means that more accurate and reliable tests of significance can be applied.

In the case of the observed data, it can be seen that the estimated coefficients for most of the variables differ in both models. For example, the coefficient of health_exp_gdp is

dramatically reduced in the Fixed Effects model, suggesting that the effect of the variable is smaller once unobserved time fixed effects are removed. The log_gdpcap coefficient is also reduced in the Fixed Effects model. The poverty coefficient changes from positive to negative, suggesting that the poverty rate has a negative impact on the dependent variable once unobserved time fixed effects are removed. Finally, the coefficient of id also becomes positive in the Fixed Effects model.

However, we could ask ourselves if, in addition to individual fixed effects, there are also temporary fixed effects (that is, if there is a heterogeneous behavior for the years considered, which can be considered constant between countries). For this purpose, we will apply the individual and time fixed effects method below.

5.1.4. Estimation by Individual and Temporary Fixed Effects (Two-way fixed effects).

Two-way fixed effects estimation is a panel data analysis method that takes into account both individual fixed effects and time fixed effects. In other words, this technique makes it possible to control for unobserved differences between individuals and unobserved changes over time, **which can lead to bias in the estimation of the variables of interest.**

The main idea behind the Two-way fixed effects estimation is that **the fixed effects of individuals and time can be included in the regression model** through dummy variables that capture the cross-variation of the data. By including these dummy variables, we control for unobserved heterogeneity that may affect the relationship between the explanatory variables and the dependent variable.

$$Healthy_years_{it} = \beta_0 + \beta_1 health_exp_gdp_{it} + \beta_2 log_gdpcap_{it} + \beta_3 poverty_{it} + \beta_4 id_{it} + \alpha_i + \varepsilon_{it}$$

$$i=1,2,3,\dots,31 \text{ countries } t=2011,2012,\dots,2020.$$

Where i represents the individual fixed effect and t represents the temporary fixed effect.

| healthy_years | Coef. | Std. Err. | t | P> t | [95% Conf. | Interval] |
|----------------|-----------|-----------|-------|-------|------------|-----------|
| health_exp_gpd | .2989218 | .3136728 | 0.95 | 0.348 | -.3416835 | .9395271 |
| log_gdpcap | 4.495875 | 12.62063 | 0.36 | 0.724 | -21.2789 | 30.27064 |
| poverty | -.0127714 | .1060315 | -0.12 | 0.905 | -.2293167 | .2037739 |
| id | .458566 | .8536241 | 0.54 | 0.595 | -1.284767 | 2.201899 |
| t | | | | | | |
| 2012 | .3504784 | .2744474 | 1.28 | 0.211 | -.2100179 | .9109747 |

| | | | | | | |
|---------|------------|----------|-------|-------|-----------|----------|
| 2013 | -.2321126 | .3930598 | -0.59 | 0.559 | -1.034848 | .5706226 |
| 2014 | -.0121304 | .5057994 | -0.02 | 0.981 | -1.045111 | 1.02085 |
| 2015 | -.401163 | .6304482 | -0.64 | 0.529 | -1.68871 | .8863841 |
| 2016 | .2895245 | .8078994 | 0.36 | 0.723 | -1.360426 | 1.939475 |
| 2017 | -.3013901 | .8489706 | -0.36 | 0.725 | -2.035219 | 1.432439 |
| 2018 | -.3757874 | .8966146 | -0.42 | 0.678 | -2.206919 | 1.455344 |
| 2019 | .0105091 | 1.00569 | 0.01 | 0.992 | -2.043385 | 2.064403 |
| 2020 | -.2123601 | .885067 | -0.24 | 0.812 | -2.019908 | 1.595188 |
| _cons | 39.51111 | 57.24627 | 0.69 | 0.495 | -77.40137 | 156.4236 |
| sigma_u | 4.3045568 | | | | | |
| sigma_e | 1.9723873 | | | | | |
| rho | 0.82647659 | | | | | |

Table 8. Two-way fixed effects estimation.

Regarding the estimated coefficients, it is observed that only "health_exp_gpd" is **statistically significant at the 5% level of significance**, given that its p-value is 0.348, which indicates that there is a positive and weak relationship between health expenditure as percentage of GDP and life expectancy. The other coefficients, "log_gdpcap", "poverty" and "id", are not statistically significant, since their p-values are greater than 0.05.

Next, I will analyze the time effects of each year on the dependent variable "healthy_years", controlling for the fixed effects of each country and for the independent variables included in the model. A positive coefficient indicates that, on average, the dependent variable increases by that number of units in that year compared to the base year (2011), while a negative coefficient indicates that the dependent variable decreases by that number of units in that year. compared to the base year.

The coefficient for the year 2012 is 0.3504784, which means that, on average, life expectancy increases by 0.3504784 years in **2012** compared to the base year 2011. Similarly, the coefficient for the year 2015 is -0.401163, which indicates that, on average, life expectancy decreases by 0.401163 years in **2015** compared to the base year 2011.

Special attention should be paid to the years in which the time coefficient shows a significant and statistically significant change. For example, in the case of the coefficients, the year 2012 shows a significant positive effect on life expectancy, while the year 2015 shows a significant negative effect. **These two years may be particularly**

relevant to understand the factors that influence life expectancy in the sample studied. In general terms, the significant coefficient in 2012 could be related to events or policies that had a positive impact on life expectancy that year. Similarly, the significant coefficient in 2015 could be related to some event or policy that negatively affected life expectancy in that year. It is important to note that these are only general assumptions and more detailed and specific analysis would be needed to determine the exact causes of these results.

5.1.5. Random Effects Estimation.

A random effects model is a model that assumes that the coefficients of the independent variables are random and vary between units of observation. This differs from the fixed effects model which **assumes that the coefficients are constant and do not vary between units of observation.** The random effects model is used when it is believed that there is unobserved variation in the independent variables that influences the dependent variable.

$$\text{Healthy_years}_{it} = \beta_0 + \beta_1 \text{health_exp_gdp}_{it} + \beta_2 \text{log_gdpcap}_{it} + \beta_3 \text{poverty}_{it} + \beta_4 \text{id}_{it} + U_i + \varepsilon_{it}$$

$$i=1,2,3,\dots,31 \text{ countries } t=2011,2012,\dots,2020.$$

where u_i is the random error term representing the unobserved variation in the independent variables.

| healthy_years | Coef. | Robust Std. Err. | z | P> z | [95% Conf. Interval] |
|----------------|-----------|------------------|-------|-------|----------------------|
| health_exp_gpd | .3591299 | .2311069 | 1.55 | 0.120 | -.0938313 .812091 |
| log_gdpcap | 5.533122 | 3.148887 | 1.76 | 0.079 | -.6385832 11.70483 |
| poverty | .0214259 | .064173 | 0.33 | 0.738 | -.1043509 .1472028 |
| id | -.1762093 | .6187715 | -0.28 | 0.776 | -1.388979 1.036561 |
| _cons | 35.00205 | 14.70665 | 2.38 | 0.017 | 6.177551 63.82654 |
| sigma_u | 4.0523322 | | | | |
| sigma_e | 1.956452 | | | | |
| rho | .81096916 | | | | |

Table 9. Random Effects Estimation.

The estimation results show that the health_exp_gpd variable has a positive coefficient of 0.359, indicating that an increase in health spending as a percentage of GDP is associated with an **increase in healthy life years.** However, the p-value associated with this variable is 0.120, indicating that it is not statistically significant at the 5% level.

The log_gdpcap variable has a positive coefficient of 5.533, suggesting that an increase in GDP per capita is associated with an increase in healthy life years. The p-value associated with this variable is 0.079, indicating that it is not significant at the 5% level.

The poverty variable has a positive coefficient of 0.021, which suggests that an increase in the percentage of the population below the poverty line is associated with an increase in the years of healthy life. However, the p-value associated with this variable is 0.738, indicating that it is not statistically significant at the 5% level.

The id variable has a negative coefficient of -0.176, which suggests that higher spending on research and development is associated with a decrease in healthy life years. The p-value associated with this variable is 0.776, indicating that it is not statistically significant at the 5% level.

The intercept (β_0) has a value of 35.00205, which suggests that healthy life years have a minimum value of 35.00205 when all other independent variables are equal to zero.

In addition, the estimation results show that the standard error of the random variable (σ_u) is 4.0523, the standard error of the error term (σ_e) is 1.9564, and the correlation coefficient between the random effects and the independent variables (ρ) is 0.8109.

5.2. Summary of the results obtained.

Below is a table that summarizes the four estimations made in this study: Pooled OLS, Estimation by fixed effects, Two-way fixed effect and random effects. Each of these estimates uses different modeling methods to analyze the relationship between the explanatory variables and the dependent variable. The table provides the estimated coefficients for each of the explanatory variables. The comparison of the results obtained through these different estimation techniques will allow choosing the most appropriate specification for the model.

| Variable | Pooled OLS | FE | Two way FE | RE |
|----------------|------------|------------|------------|-----------|
| health_exp_gpd | 0.7121375 | 0.30123163 | .2989218 | .3591299 |
| log_gdpcap | 8.269625 | 2.0889126 | 4.495875 | 5.533122 |
| poverty | 0.1434974 | -.01791377 | -.0127714 | .0214259 |
| id | -1.288059 | 0.31108379 | .458566 | -.1762093 |
| t | | | | |
| 2012 | | | .3504784 | |
| 2013 | | | -.2321126 | |
| 2014 | | | -.0121304 | |
| 2015 | | | -.401163 | |

| | | | | |
|-------|----------|-----------|-----------|----------|
| 2016 | | | .2895245 | |
| 2017 | | | -.3013901 | |
| 2018 | | | -.3757874 | |
| 2019 | | | .0105091 | |
| 2020 | | | -.2123601 | |
| _cons | 20.45036 | 50.204819 | 39.51111 | 35.00205 |
| N | 310 | 310 | 310 | 310 |
| R2 | 0.2444 | 0.01687 | 0.033497 | |
| R2_a | 0.23451 | -0.10467 | -0.00895 | |

Table 10. Summary table of the obtained results.

5.3. Specification selection.

After having estimated the model through three different methods, it is necessary to select the most suitable specification. For this, different contrasts can be used, among which the **Breusch-Pagan Test** and the **Hausman Test** stand out.

The **Breusch-Pagan Test** is used to compare two models: the random effects model and the pooled effects model. This test is based on the null hypothesis that there is no heteroskedasticity in the random effects model. If this null hypothesis is rejected, then it is suggested that the pooled effects model is more suitable, since it assumes that the variance of the errors is constant for all observations. On the other hand, if the null hypothesis is not rejected, it can be concluded that the random effects model is more appropriate.

On the other hand, the **Hausman Test** is used to compare two different models: the fixed effects model and the random effects model. This test is based on the null hypothesis that the coefficients of the random effects model are consistent and efficient, that is, that the estimators are unbiased and have the lowest possible variance. If this null hypothesis is rejected, then it is suggested that the fixed effects model is more suitable, since it has the advantage of being consistent even if the explanatory variables are correlated with the random effects. On the other hand, if the null hypothesis is not rejected, it can be concluded that the random effects model is more appropriate, since it is more flexible and does not require the specification of fixed effects.

5.3.1. Breusch-Pagan test for random effects (Random Effects vs. Pooled Effects).

Breusch-Pagan test is a technique that is used to decide between the estimation by random effects and the estimation Pooled OLS (Estimation by Ordinary Least Squares).

| Estimated results: | Var | Sd=sqr (Var) |
|--------------------|---------------|--------------|
| Healthy_years | 23.69063 | 4.867302 |
| e | 3.827705 | 1.956452 |
| u | 16.4214 | 4.052332 |
| Test: Var (u)=0 | Chibar2 (01)= | 826.96 |
| | Prob>chibar2= | 0.00000 |

Table 11. Breusch Pagan test.

The null hypothesis of the test is that the residual variance in the random effects regression is equal to the residual variance in the Pooled OLS regression. If the p-value obtained from the Breusch-Pagan test is less than the chosen significance level (for example, 0.05), we reject the null hypothesis and conclude that the random effects estimate is preferable to the Pooled OLS estimate.

In your case, the value of the **chi-square statistic (chi2) is 826.96 and the p-value (prob>chi2) is 0.000**, which indicates that the probability that the data is consistent with the null hypothesis (residual variance equal to in both models) is very low. Therefore, **we can reject the null hypothesis and conclude that the random effects estimate is preferable to the Pooled OLS estimate.**

In summary, due to the low p-value of the Breusch-Pagan test, **it is concluded that the estimation by random effects is the most suitable specification for the model.**

5.3.2. Hausman test (Fixed Effects vs. Random Effects).

The **Hausman Test** is a statistical technique used in econometrics to choose between the fixed effects estimation and the random effects estimation in a panel model. This test is based on the comparison of the estimated coefficients of both specifications to determine which is the most appropriate for the data set analyzed.

The null hypothesis in the Hausman Test is that the random errors are not correlated with the explanatory variables. If the null hypothesis is rejected, the fixed effects estimate is chosen. If the null hypothesis is accepted, the random effects estimate is chosen.

In summary, the Hausman Test is a useful tool to determine if fixed effects are necessary in a panel model, or if a random effects estimate can be used without losing precision in the estimates.

| | (b) | (B) | (b-B) | Sqrt(diag(V_b-V_B)) |
|----------------|---------|---------|------------|---------------------|
| | FE | RE | Difference | S.E. |
| Health_exp_gdp | 0.3012 | 0.35913 | -0.05789 | 0.07628 |
| Log_gdpcap | 2.0889 | 5.5331 | -3.4442 | 5.179292 |
| Poverty | -0.0179 | 0.0214 | -0.03933 | 0.03848 |

| | | | | |
|------------|---------------|---------|---------|---------|
| id | 0.3111 | -0.1792 | 0.48729 | 0.39934 |
| Chi2(4)= | 5.19 | | | |
| Prob>chi2= | 0.2682 | | | |

Table 12. Hausman test.

The value of the chi-square test statistic (chi2) indicates the difference in parameter estimates between the two models. The null hypothesis is that there is no significant difference in the efficiency of the two models and the random effects model is preferred if this hypothesis cannot be rejected.

In this case, the value of the **chi2 test statistic is 5.19 and the associated probability (prob>chi2) is 0.2682**. In general, a probability>chi2 value less than 0.05 is considered to indicate that fixed effects are more appropriate, while a value greater than 0.05 indicates that random effects are more appropriate. This suggests that there is not enough evidence to reject the null hypothesis that there is no significant difference in the efficiency of the two models. Therefore, **the random effects model could be chosen as it provides a more efficient estimate of the parameters than the fixed effects estimate**.

6. Conclusions.

This study has analyzed the effect of spending on health on life expectancy in European Union countries, with a database of 31 member countries, with the exception of Norway and Iceland and a period of years from 2011 to 2020.

The panel data technique has been used and four different models have been estimated: Pooled OLS, Estimation by fixed effects, Two-way fixed effect and random effects. After carrying out the corresponding statistical tests, the estimation by random effects has been chosen as the most efficient.

6.1. Econometric conclusions.

The table presented above (**Random Effects Estimation Table**) shows the results of the estimation by random effects of an econometric model that seeks to determine the relationship between health spending and life expectancy measured in years in 31 member countries of the European Union. , with the exception of Norway and Iceland, during the period from 2011 to 2020.

The coefficient of the variable health_exp_gpd, which represents **health spending as a percentage of GDP, is positive and significant at 88% statistical confidence**. This means that, on average, a 1% increase in health spending is associated with an increase of 0.36 years in the life expectancy of the population. This result is consistent with the

economic literature, which has indicated that spending on health is one of the main determinants of the health and life expectancy of a population.

On the other hand, the coefficient of the variable `log_gdpcap`, which **represents GDP per capita in logarithms, is positive but not significant at 92% statistical confidence**. This suggests that income does not have a significant impact on life expectancy, which may be related to the fact that the European Union has a relatively homogeneous level of income among its member countries.

The coefficient of the `poverty` variable, which represents the **percentage of the population below the poverty line, is positive but not significant at 74% statistical confidence**. This suggests that poverty does not have a significant impact on the life expectancy of the population in the European Union.

Regarding **R&D expenditure in terms of GDP (id)**, its coefficient turned out to be **negative and not significant at the 5% level**. This indicates that, on average, an increase in R&D spending **is not associated with a significant increase or decrease in life expectancy in the sample countries**. It is important to note that this variable is a measure of research and development spending, which suggests that the results do not support the idea that R&D investment has a direct and significant impact on life expectancy. However, it should be noted that there are other possible channels through which R&D spending could have an indirect and positive impact on the health and, therefore, on the life expectancy of the population.

Finally, the value of `sigma_u`, which represents the variance of the random error term u , is 4.05. This indicates that there is substantial variability in the effect of each country on life expectancy, justifying the use of the random effects estimate instead of the fixed effects estimate.

In conclusion, the results of this econometric estimate suggest that health spending is the main determinant of life expectancy in the European Union, while income and poverty do not have a significant impact. In addition, it was found that there is substantial variability in the effect of each country on life expectancy, which justifies the use of random effects estimation. However, it is important to note that there are other factors that can affect life expectancy that were not included in this model. Therefore, these results should be interpreted with caution and used as a guide for future public health research and policy.

6.2. Economic conclusions.

Taking into account the **Research Question "How does health spending in the European Union affect people's life expectancy?"** and the previously mentioned hypotheses:

- Null hypothesis (H0): There is no relationship between health spending and life expectancy in the European Union.
- Alternative hypothesis (H1): There is a positive relationship between health spending and life expectancy in the European Union.

The economic conclusion is that, after carrying out the econometric analysis, **statistical evidence has been found that supports the alternative hypothesis (H1) that there is a positive relationship between health spending and life expectancy in the European Union.** In other words, the study suggests that as health spending in the European Union increases, so does people's life expectancy.

The reason behind this conclusion is that econometrics has shown that there is a positive and significant correlation between health spending and life expectancy in the European Union, even after controlling for other important factors that could influence life expectancy. This indicates that health spending has a positive effect on life expectancy in the European Union and that government policies that increase health spending can have a positive impact on the health and well-being of the general population.

7. References.

- EC (European Commission). (2021). GDP per capita. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table
- EC (European Commission). (2021). R&D spending as a percentage of GDP. Eurostat. <https://ec.europa.eu/eurostat/databrowser/view/tipsst10/default/table?lang=en>
- EC (European Commission). (2021). Public health expenditure as a percentage of GDP. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_sha11_hf/default/table?lang=en
- EC (European Commission). (2021). Health care resources. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_tech/default/table?lang=en
- EC (European Commission). (2021). Health personnel (excluding nursing and caring professionals). Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_prs1/default/table?lang=en
- EC (European Commission). (2021). Medical technology. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_rs equip/default/table?lang=en
- EC (European Commission). (2021). Child mortality. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/demo_minfind/default/table?lang=en
- EC (European Commission). (2021). Death. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_cd_aro/default/table?lang=en
- EC (European Commission). (2021). Health personnel employed in the hospital. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_prshp1/default/table?lang=en
- EC (European Commission). (2021). Medical staff. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_phys/default/table?lang=en
- EC (European Commission). (2021). Years of healthy life by sex.. Eurostat. https://ec.europa.eu/eurostat/databrowser/view/hlth_hlye/default/table?lang=en
- EC (European Commission). (2021). Healthcare expenditure statistics [Webpage]. Eurostat. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_expenditure_statistics#Healthcare_expenditure
- Gallo, P., & García-Goñi, M. (2001). The effects of demographic aging on health spending: myths and realities. *Gaceta Sanitaria*, 15(1), 65-77. <https://www.sciencedirect.com/science/article/pii/S0213911101715361>
- Rentería Perez Elisenda & Zueras Castillo Pilar (2022). Public health spending and life expectancy: a healthy investment. *Centro de Estudios Demográficos*. https://ced.cat/PD/PerspectivesDemografiques_029_ESP.pdf
- Moseley, G. B., & Carlson, K. (2016). *Applied health economics*: Jones & Bartlett Publishers. <http://cphs.huph.edu>

Health and its determinants, health promotion and health education. Mario Alberto De La Guardia Gutiérrez, Jesús Carlos Ruvalcaba Ledezma (2020) DOI: 10.19230/jonnpr.3215 https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S2529-850X2020000100081