eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Empirical Applications and Policies

# Measuring healthcare deprivation from self-reported data

**Author:** María Clemente Marcuello

**Supervisors**: María José Gutiérrez Huerta and Susan Orbe Mandaluniz

University of the Basque Country UPV/EHU - Master in Economics: Empirical Applications and Policies

2021/22

**Abstract**

The aim of this Master's thesis is to analyze households' healthcare deprivation and its determinants in Spain for the years 2014, 2017 and 2020. The analysis is divided in two main sections; the first one aims to determine the evolution and presence of healthcare deprivation in Spain by region and area using counting deprivation measures. The second aims to analyze the determinants of healthcare deprivation using zero inflated models from a multidimensional and unidimensional approach. We obtain that healthcare deprivation varies depending on the region and area; having the northern and rural areas less deprivation. And we conclude that income, area, number of retired and incapacitated members in the household and household size are variables that determine the level of deprivation of the households.

**Keywords:** healthcare deprivation, multidimensional, unidimensional, zero inflated

## Contents

# 1. INTRODUCTION

The right to health protection is recognized in the Article 43 of the Spanish Constitution and specified in the General Health Law (Law 14/1986), which establishes its public financing, universality and gratuity; its decentralization to the Autonomous Regions and integration into the National Health System (NHS). To achieve that, its access must be considered as basic right; i.e., everyone must have the right to healthcare under equal conditions.

However, we can say that this does not uphold. According to OECD (2019), Portugal, Latvia, Spain, and Estonia display the largest concentration of unmet health related needs for financial reasons among lower income groups and the largest degrees of inequality in Europe. Moreover, Amnesty International (2019) reported that austerity measures implemented during the Economic Crisis in 2008 by the Spanish government have worsened the health situation of the most economically vulnerable groups in society, due to the long waiting lists and economic problems to afford the medication needed.

These are the reasons why measuring healthcare deprivation as the lack of access to healthcare services due to financial problems is considered a relevant issue. This study focuses on determining the presence and determinants of healthcare deprivation in Spain. The analysis is developed from multidimensional perspective. The data used for the analysis was obtained from the surveys European Survey of Health in Spain (EHSS) and National Health Survey (NHS), in particular, from the section denominated *Unmet healthcare needs*. The total observations analyzed are 67,943.

Counting deprivation measures are used to compute scores for healthcare deprivation for the years 2014, 2017 and 2020, both for Spain and its Autonomous Regions. Also, deprivation measures have been calculated classifying households as located in rural and urban areas in Spain. Moreover, to establish the determinants of healthcare deprivation a regression analysis using zero inflated models has been developed in a multidimensional and unidimensional framework.

The report is organized as follows. Section 2 describes the data used. Section 3 provides information about the measurement for healthcare deprivation: methodology used and results found through graphs, as well as a discussion about them. Section 4 provides information about the methodology used to establish the determinants of healthcare deprivation, as well as the results found and discussion about them. And, lastly, Section 5 summarizes the main conclusions of the study.

## 2. DATA

The data used in this study has been obtained from the surveys European Survey of Health in Spain (EHSS) and National Health Survey (NHS), both compiled by the National Statistics Institute (INE) an independent administrative Autonomous Spanish institution. The data used is from the years 2014, 2017 and 2020 due their surveys starting in 2014 to collect information on the unmet need for health care. Specifically, the information needed for 2017 was obtained from the NHS and for 2014 and 2020 from the EHSS. Moreover, the type of sampling used is stratified three-stage sampling for the three years and it is significant at regional level (NUTS2[1]) (INE, 2020).

Both surveys are carried out every five years and are alternated every two and a half. This is because the Ministry of Health and the INE agreed to alternate them, so the information about the population’s health status would be more efficiently obtained with an adequate periodicity and to avoid duplicity in the information gathered. Moreover, they also agreed on adapting some of the questions of the EHSS2014 and EHSS2020, so the data required for the health indicators of the National Health System would be obtained through the questionnaires used to conduct the interviews of the EHSS. This collaboration was formalized in the agreements signed by both organizations in April 2014 and in July 2019 for EHSS2014 and EHSS2020, respectively (INE, 2020).

Moreover, the surveys provide different micro datafiles with the household information and the interviewed adult of reference in the household. The first contains information about the household's identification data and composition as well as the household's monetary income and housing characteristics. The second micro datafile, in addition to general data such as identification information, demographic and physical characteristics and economic activity of the individual, contains information on health status, accidents, activity restrictions, physical, sensory and cognitive limitations and limitations in daily activities, mental health, medical consultations, hospitalizations, emergencies, health

[1] Nomenclature of territorial units for statistics; hierarchical system that divides EU and UK territories to collect and develop European statistics.

insurance, medication consumption, preventive/healthy practices, unmet health care needs, physical activity, diet, consumption of harmful substances, support available and whether the individual takes care of another person with health problems.

Regarding the data collected in the EHSS2020, administrative issues delayed the start of the interviews, hence, the data collection was done from July 2019 to July 2020. This developed into further problems due to the difficulties to gather all the information scheduled because of the Covid-19 pandemic in 2020, which caused Spain to declare the State of Alarm and subsequent lockdown the 13 of March and 14 of March respectively. Therefore, during the 2020 lockdown the interviewers changed the method of collecting the information from a personal interview to a telephonic interview instead and consequently less information was gathered.

Nevertheless, the issue was addressed by the INE and the sample was corrected using corrected weights for the lack of answers, which are not homogeneously distributed throughout the reference weeks (INE, 2020).

Furthermore, the final dataset consists in 67,943 observations; 22,817 (2014), 23,073 (2017) and 22,053 (2020) and each of them with a statistical representation of 18,300,000; 17,800,000 and 18,800,000 households respectively computed using the *weights* associated to the observations. All 21 variables considered for the analysis can be found in Appendix 1 and Table 1, being the most relevant in Table 1:

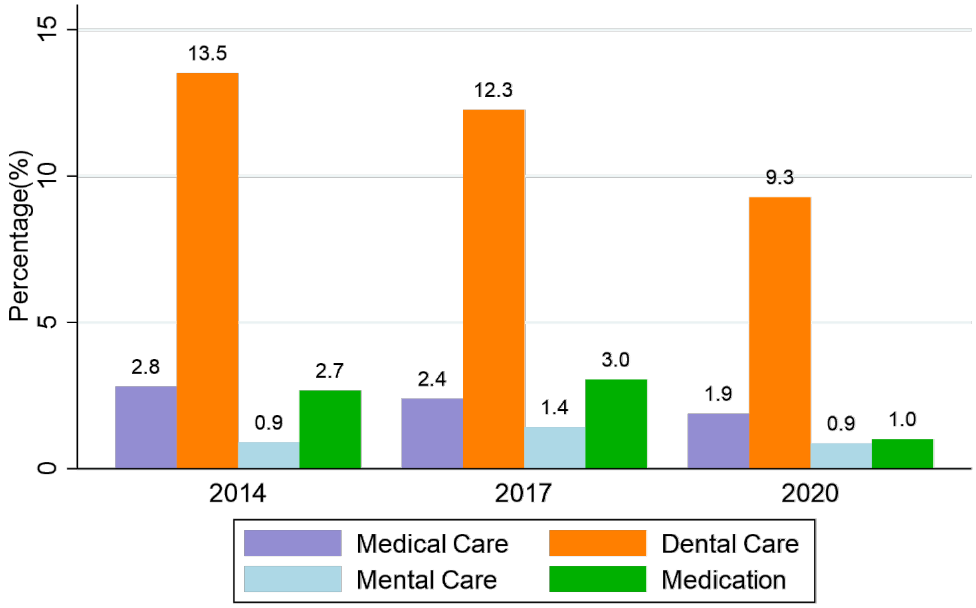
Table 1: Variables

| Variable name | Variable description | % or Mean (SD) |
|---|---|---|
| Dependent variables: | | |
| Dental care | =1 if they had financial problems to access dental care; =0 otherwise | 0.110 (0.313) |
| Mental care | =1 if they had financial problems to access mental care; =0 otherwise | 0.009 (0.098) |
| Medication | =1 if they had financial problems to access medication; =0 otherwise | 0.020 (0.141) |
| Medical care | =1 if they had financial problems to access medical care; =0 otherwise | 0.023 (0.156) |
| Deprivation counts of the households | =0 if household not deprived in any dimension | 87.75% |
| | =1 if household deprived in one dimension | 9.27% |
| | =2 if household deprived in two dimensions | 2.12% |
| | =3 if household deprived in three dimensions | 0.65% |
| | =4 if household deprived in all dimensions | 0.22% |
| Independent variables | | |
| Region | =1 if household is in Andalucía region | 12.30% |
| | =2 if household is in Aragón region | 4.29% |
| | =3 if household is in Asturias region | 3.95% |
| | =4 if household is in Balearic Islands region | 3.14% |
| | =5 if household is in Canarias region | 4.94% |
| | =6 if household is in Cantabria region | 3.74% |
| | =7 if household is in Castilla and Leon region | 5.48% |
| | =8 if household is in Castilla-La Stain region | 4.97% |
| | =9 if household is in Catalonia a region | 10.06% |
| | =10 if household is in Valencia a region | 8.08% |
| | =11 if household is in Extremadura region | 4.18% |
| | =12 if household is in Galicia region | 5.85% |
| | =13 if household is in Madrid region | 10.02% |
| | =14 if household is in Murcia region | 4.48% |
| | =15 if household is in Navarre region | 3.47% |
| | =16 if household is in Basque Country region | 5.86% |
| | =17 if household is in The Rioja region | 2.95% |
| | =18 if household is in Ceuta city | 0.99% |
| | =19 if household is in Melilla city | 1.25% |
| Household size | Household size (OECD-modified scale proposed by Haagenars *et al.* (1994)) | 1.657 (0.538) |
| Income | Income per capita (thousands of euros) | 0.763 (0.732) |
| Area | =1 household is located in an urban area; =0 otherwise | 0.772 (0.419) |
| Retired and incapacitated | Number of household members retired and incapacitated | 0.520 (0.824) |
| Self-assessed health perception (two levels) | =1 bad health; =0 otherwise | 0.319 (0.466) |
| Self-assessed health perception (five levels) | =1 very good | 19.32% |
| | =2 good | 48.78% |
| | =3 regular | 22.52% |
| | =4 bad | 7.2% |
| | =5 very bad | 2.18% |

*Source: Own elaboration, data from INE*

The variables, which are used to compute whether a household is healthcare deprived, are: *dental care*, *mental care*, *medication* and *medical care*. These are dichotomous and take value 1 for positive answers and 0 for the negative ones, so when a household has a value 1 associated to all of these variables it can be said that it has not been able to afford health care in those 4 dimensions. Also, the observations that had not these questions answered were deleted; due *dental care* not being answered there were deleted 17, due *mental care* 8, due *medication* 2 and due *medical care* 33. In the following Graph 1, a summary of the positive answers to these dimensions is gathered.

Graph 1: Summary of the answers to the variables: health care, medication, mental care and dental care




*Source: Own elaboration, data from INE*

As can be seen in Graph 1 above, the percentages corresponding to households that cannot allow themselves dental care is considerably higher than the percentages for the other dimensions even it has been reduced from 2014 to 2017 and again to 2020. On the other hand, the percentage of positive answers for the variables *medication* and *mental care* increase from 2014 to 207 to decrease again to 2020.

The income was divided in intervals in the NHS and EHHS. Therefore, it was chosen to use the mean of each interval and divide it by the *household* size so the income per capita of the households was obtained. This way we obtained the variable *income* measured in thousand euros.

The variable *area* was also obtained by recoding the original variable *ESTRATO* from the surveys detailed in Appendix 1. In this case, according to the Ministry of Transport, Mobility and Urban Agend, GD[2] of Housing and Land (2021), an area can be considered as urban when it has more a population higher than 10,000 inhabitants. Hence, in this study, we have applied that criteria to differentiate urban and rural households.

## 3. MEASURING MULTIDIMENSIONAL HEALTHCARE DEPRIVATION

Deprivation is a term that can be define as "a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs" (Townsend, 1987). Hence, in this study, we will define healthcare deprivation as the disadvantage that households hold due to their lack of access to any type of healthcare because of economic reasons. There have been previous studies that have connected healthcare deprivation with inequality (McGillivray et al (2009) and Benzeval and Judge (2001)), neighborhood (Bilger and Carrieri (2013) and Rój (2020)) or material deprivation (Vázquez et al. (2014)). Therefore, since healthcare deprivation can be associated with more than one variable or dimension, it must be analyzed from multidimensional approaches. This is a similar approach to one use to study poverty, which is also analyzed through various multidimensional measures due to depending on more than one variable. Various authors have studied multidimensional poverty and how to measure it using the inequality metrics approach, like Bourguignon and Chakravarty (1999, 2003) and Cowell (1988).

In this case, we follow the counting poverty measures proposed by Alkire and Foster (2011) and Chakravarty and D'Ambrosio (2006) to compute healthcare deprivation in Spain in 2014, 2017 and 2020. This procedure has been used previously to study energy poverty in Spain by Aristondo and Onaindia (2018a,b). Moreover, due to the fact that we are using different indexes to compute healthcare deprivation, we do a robustness check constructing the dominance curves to ensure comparability of the indexes for different thresholds and years following the method proposed by Lasso de la Vega (2010).

[2] General Directorate

## 3.1 METHODOLOGY

The dimensions $(k)$ we consider to compute the healthcare deprivation are in Table 2 and have been obtained from EHSS and EHN surveys from the INE. Moreover, as far as we know this would be the first time using these items as dimensions to measure healthcare deprivation.

Table 2: Healthcare Deprivation k Dimensions

| Variables (INE) | Renamed | Description |
|---|---|---|
| **R108_1** | Medical care | Lack of health care due to financial problems in the last 12 months: health care |
| **R108_2** | Dental care | Lack of health care due to financial problems in the last 12 months: dental care |
| **R108_3** | Medication | Lack of health care due to financial problems in the last 12 months: medication prescribed |
| **R108_4** | Mental care | Lack of health care due to financial problems in the last 12 months: mental health care |

*Source: Own elaboration, data from INE*

We start by defining the deprivation vector of household $i$; $x_i = (x_{i1},\ x_{i2}, \dots, x_{ik})$ where $x_{ij} = 1$ when household $i$ is deprived in dimension $j \in (1, \dots k)$ whilst $x_{ij} = 0$ otherwise. Then, the deprivation score $(d_i)$ of household $i$ can be written as follows:

$$d_i = \sum_{j=1}^{k} w_j x_{ij},$$

where $w_j\ \ \forall\, j \in k$ represents the weights associate to dimension $k$. For the purpose of this study. All dimensions are considered equally important. This means that a non-deprived household in all of the dimensions would have a deprivation score of 0, whereas a household deprived in all of them would have a $k$ score.

The first step in counting poverty measures is to identify those households considered as poor or deprived. To determine this, we define a cut-off $m = \{1,2, \dots, k\}$, and we stablish that when household $i$ is deprived in at least $m$ dimensions, $d_i \geq m$, then it is considered as healthcare deprived, while it is considered as non-deprived when $d_i < m$.

Through this procedure, we identify whether a household is healthcare deprived or not and we obtain $q_m$, which is the total number of deprived households identified using the dimension cut-off $m$.

The second step consist of obtaining a single numerical deprivation value by aggregating the deprivation scores. A deprivation measure that represents the level of healthcare deprivation in the society for a given cut-off $m$, $P_m(\boldsymbol{d})$, where $\boldsymbol{d}$ is the vector of all the weighted deprivation counts of the households taken into consideration. Moreover, when $w_j = 1$, we obtain $D = \{0,1,..,k\}$ being $D$ the set of admissible scores for $d_i$ whereas for $w_j \neq 1$, we obtain the discrete set $D$ with $2^k$ elements maximum. Any well-behave deprivation measure, $P_m(\boldsymbol{d})$, as a deprivation measure, has to fulfill the following five properties.

1. *Poverty Focus* (**PF**): $P_m$ does not change if the deprivation score for a non-deprived person decreases.

2. *Dimensional Monotonicity* (**MON**): $\forall\, m \in (0,k]$, $P_m(\boldsymbol{d}') < P_m(\boldsymbol{d})$ if $d_i' < d_i$ for an individual $i$ with $d_i \geq m$, and $d_j' = d_j$ for all $j \neq i$,

3. *Distribution Sensitivity* (**DS**): $\forall\, m \in (0,k]$ and $h > 0$; $P_m(\boldsymbol{d}) - P_m(d_1, \dots, d_i - h, \dots, d_j, \dots, d_n) > P_m(\boldsymbol{d}) - P_m(d_1, \dots, d_i, \dots, d_j - h, \dots, d_n)$ if $(d_1, \dots, d_i - h, \dots, d_j, \dots, d_n), (d_1, \dots, d_i, \dots, d_j - h, \dots, d_n) \in D^n$ and $d_i > d_j \geq m$

4. *Symmetry* (**SYM**): $\forall\, m \in (0,k]$, $P_m(\boldsymbol{d}') = P_m(\boldsymbol{d})$ if $\boldsymbol{d}'$ is a permutation of $\boldsymbol{d}$

5. *Replication Invariance* (**RI**): $\forall\, m \in (0,k]$, $P_m(\boldsymbol{d}') = P_m(\boldsymbol{d})$ if $\boldsymbol{d}' = (\boldsymbol{d}, \dots, \boldsymbol{d})$

In this study, we use the three most used which are the headcount multidimensional ratio and $P_m^\alpha$ for $P_m^\alpha$ for $\alpha = 1, 2$. The first counting deprivation measure, $P_m(\boldsymbol{d})$, that we use is the multidimensional headcount ratio:

$$H_m = \frac{q_m}{n},$$

where $q_m$ is the number of healthcare deprived households and $n$ the households in the society. It measures the percentage of healthcare deprived with cut-off $m \in (0,k]$ in a society. It should be noted that this index violates *MON* because it does vary when a household becomes deprived in one more dimension.

Then the second counting poverty measure introduced is $P_m^\alpha$ proposed by Chakravarty and D'Ambrosio (2006):

$$P_m^\alpha = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{d_i(m)}{k}\right)^\alpha,$$

where $\alpha$ can be interpreted as the aversion of society against deprivation, because if α raises *ceteris paribus*, $P_m^\alpha$ becomes more sensitive to the most deprived. Whether this index satisfies all of the counting poverty properties or not depends on the value that $\alpha$, i.e., when $\alpha = 0$, we obtain $H_m$ which does not fulfill *MON.* Then, if we use $\alpha = 1$ we obtain the index $M_m$ proposed by Alkire and Foster (2011):

$$P_m^1 = M_m = \frac{1}{n \cdot k}\sum_{i=1}^{n} d_i(m),$$

which is the weighted sum of the deprivations among households. This index fulfils all the properties except *DS* because it does not have into account the inequality between the healthcare deprived considering all of them equally deprived.

Then, for $\alpha \geq 2$, $P_m^\alpha$ all the properties of the counting deprivation measures will be fulfilled. Hence, in this study we use $P_m^\alpha$ for different values of $\alpha$.

On the other hand, different counting deprivation measures and different cut-offs, $m$, for the same measure can lead to contradictory results. To overcome this problem, the proposal on counting dominance proposed by Lasso de la Vega (2019) is taken into consideration.

First, we have to draw the *dimension deprivation curves*, that in this case are denominated as *FD* curves associated with vector $\boldsymbol{d}$ (sum of the deprivation scores). This curve represents the multidimensional headcount ratio for any vector of deprivation counts and admissible dimension cut-offs, ranked in decreasing order.

$$FD(\boldsymbol{d}; p) = H_{k-p} \quad p \in [0, k].$$

The use of these dominance curves enables the results obtained for the different deprivation counting measures and cut-offs to be compared. This is because the *FD* dominance curves hold the proposition in Appendix 2 which indicates that even though $H_m$ does not fulfill *MON*, the ordering with respect to $H$ is equivalent to agreement over all counting measures satisfying *MON* (Lasso de la Vega, 2010). Consequently, if the FD curves of two vectors $\boldsymbol{d}$ do not intersect, then all the counting deprivation measures satisfying *MON* will lead to the same ranking. Moreover, if they do intersect the cut-off can be changed to more restrictive to establish dominance conditions.

## 3.2 RESULTS

First, using the variables *dental care, mental care, medication* and *medical care* as dimensions to measure healthcare deprivation, we compute the indexes $H_m$, $P_m^1$ and $P_m^2$. In Table 3, we compiled the results obtained for indexes $H_m$, $P_m^1$ and $P_m^2$ and the years 2014, 2017 and 2020 as well as their values for the different cut-offs, $m$.

*Table 3: Healthcare deprivation for the years 2014, 2017 and 2020*

| | $\boldsymbol{m}$ | 2014 | 2017 | 2020 |
|---|---|---|---|---|
| $\boldsymbol{H_m}$ | 1 | 15.239 | 13.744 | 10.067 |
| | 2 | 3.568 | 3.770 | 2.239 |
| | 3 | 0.908 | 1.245 | 0.576 |
| | 4 | 0.177 | 0.364 | 0.148 |
| $\boldsymbol{P_m^1}$ | 1 | 0.04973 | 0.04781 | 0.01293 |
| | 2 | 0.02055 | 0.02287 | 0.00804 |
| | 3 | 0.00725 | 0.01025 | 0.00469 |
| | 4 | 0.00177 | 0.00364 | 0.00148 |
| $\boldsymbol{P_m^2}$ | 1 | 0.01983 | 0.02114 | 0.01294 |
| | 2 | 0.01253 | 0.01491 | 0.00804 |
| | 3 | 0.00588 | 0.00860 | 0.00388 |
| | 4 | 0.00177 | 0.00364 | 0.00148 |

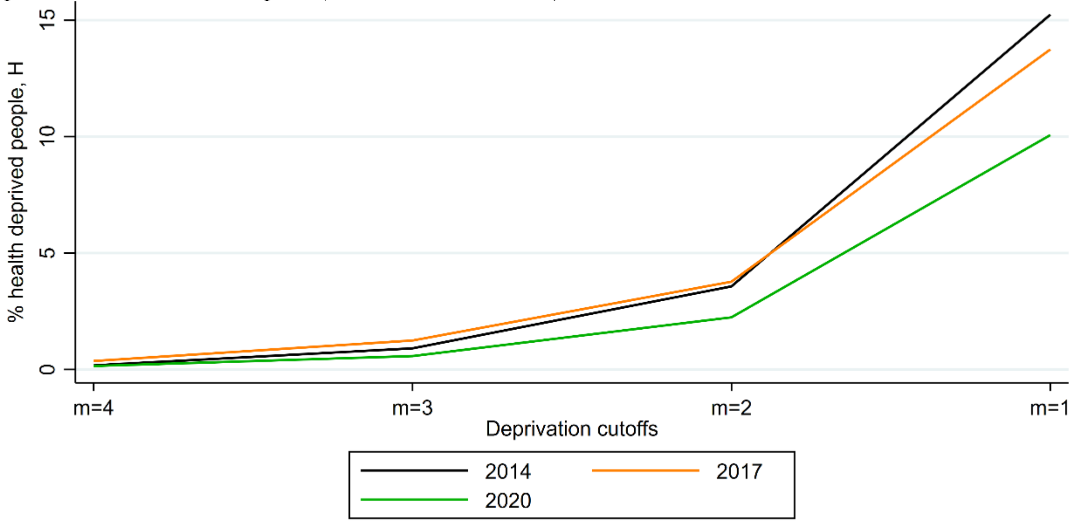
*Source: Own elaboration, data from INE*

As can be observed, the values for $m = 4$ are the same for the three measures. This confirms that the indexes have been computed correctly due to:

$$\frac{q_m}{n} = \frac{1}{n}\sum_{i=1}^{4}\left(\frac{d_i(4)}{4}\right)^1 = \frac{1}{n}\sum_{i=1}^{4}\left(\frac{d_i(4)}{4}\right)^2 \Rightarrow H_4 = P_4^1 = P_4^2.$$

Then, we present the FD curves of $H_m$ for the different years in a decreasing order of $m$. As can be seen in Graph 2, the curves for 2014 and 2017 intersect which implies that the cut-off must be changed to a more restrictive one if we want to compare both years. Hence, to proceed with the analysis that includes comparisons among years 2014, 2017 y 2020 we set the cut-off $m = 2$, so the dominance conditions can be established unambiguously and the values obtained compared between waves because they have the same ranking for any cut-off or wave.

*Graph 2: FD curves $H_m$ in Spain (2014, 2017 and 2020)*




*Source: Own elaboration, data from INE*

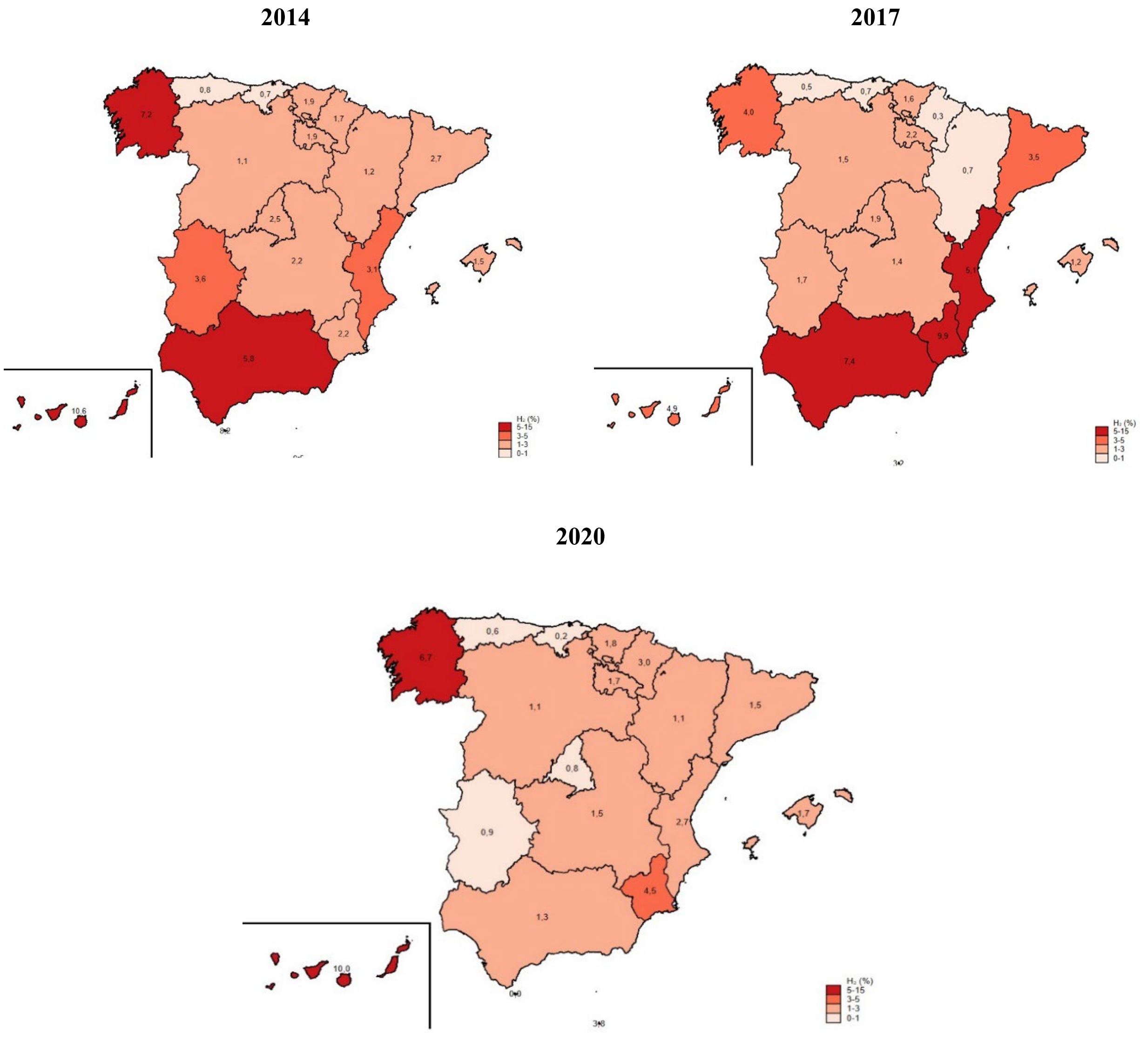
So, after establishing the cut-off as $m = 2$, we can compare the results obtained with the different deprivation measures and cut-offs. It can be said that there was an increase in healthcare deprivation in 2017 from 2014. This is due to the fact that the values obtained are higher for all three cutoffs analyzed in 2017 than in 2014. In Graph 2, it can be observed how the curve form 2017 is consistently above the curve from 2014, which indicated presence of higher percentages of deprived households in 2017 than in 2014.

For the year 2020, there is an improvement in healthcare deprivation for all the cut-offs $m \geq 2$. The percentages of healthcare deprived households have been considerably reduced obtaining lower values than those obtained in 2014. This means that healthcare deprivation in 2020 is lower than in 2014.

### 3.2.1 REGIONAL ANALISYS

After computing of the deprivation measures and establishing the cut-off, we focus the analysis on computing the healthcare deprivation value for each region of Spain. The three measures proposed in the methodology have been used to compute the different values for the regions. However, we only show Figure 1 which are the maps corresponding to the values obtained for $H_2$ for 2014, 2017 and 2020, respectively. Results for $P_2^1$ and $P_2^2$ similar are available in Appendix 2.

*Figure 1: Healthcare deprivation by region in 2014, 2017, 2020 ($H_2$)*




*Source: Own elaboration, data from INE*

Maps in Figure 1 allow us to observe the changes in healthcare deprivation over the years and by region in Spain. It can be observed that the regions of Canarias and Galicia have consistently higher values of healthcare deprivation compared to the other regions. Also, Andalucía in 2014 and 2017 had a high value of deprivation, however, in 2020 this value has been considerably reduced. Also, in Murcia's case, we can see how the percentage of deprived people increased considerably in 2017 and it has not been reduced for 2020 as much as it increased. This makes Murcia the third region with highest percentage of deprived people in 2020.

On the other hand, regions like Asturias, Cantabria and Aragón have not had deprivation values higher than 1.2 in the three years for $H_2$. This indicates that those are the regions in Spain with lowest percentage of deprived people. Lastly, it can be said that in 2020 most of the regions reduced the healthcare deprivation values after the 2017 increase.

### 3.2.2 AREA ANALYSIS

In this section, the differences in healthcare deprivation are analyzed for households in *urban* and *rural* areas. *Urban* areas are those that have at least 10.000 inhabitants per municipality and *rural* areas are those that have less than 10.000 inhabitants per municipality. This criterion is used by the Ministry of Transport, Mobility and Urban Agenda in the report of "Áreas Urbanas en España, 2021" to determine which areas can be considered as urban.

The three indexes detailed in the Section 3.1, $H_m$, $P_m^1$ and $P_m^2$ have been computed taking into consideration an identification cut-off $m = 2$ and the results obtained are compiled in Table 3.

*Table 3: Healthcare deprivation by area (2014, 2017 and 2020)*

| | ***area*** | **2014** | **2017** | **2020** |
|---|---|---|---|---|
| $\boldsymbol{H_2}$ | Rural | 2.707 | 2.651 | 1.800 |
| | Urban | 3.796 | 4.070 | 2.350 |
| $\boldsymbol{P_2^1}$ | Rural | 0.01636 | 0.01625 | 0.01045 |
| | Urban | 0.02166 | 0.02465 | 0.01365 |
| $\boldsymbol{P_2^2}$ | Rural | 0.01057 | 0.01060 | 0.00643 |
| | Urban | 0.01385 | 0.01606 | 0.00845 |

*Source: Own elaboration, data from INE*

Values of healthcare deprivation are significantly higher for the urban areas than for the rural areas. So, despite the initial idea that better access to healthcare is available in urban areas, these deprivation values show the existence of worse access to healthcare for deprived people living in them. This is similar to the conclusions obtained by Padeiro (2017) regarding the access of elderly people to pharmacies in metropolitan areas, because despite the suggestion of good coverage at the metropolitan scale, accessibility measures proved the opposite.

On the other hand, areas considered as urban have deprivation values higher in 2017 than in 2014. Then, for 2020 the values decrease even more than the initial ones in 2014. So, it can be said the healthcare deprivation has been reduced from 2014 to 2020. These results coincide with the previously obtained when the indexes were computed by cut-off and by region. Lastly, rural areas observe a reduction in healthcare deprivation for the three indexes in 2020 from 2017. Also, the changes between 2014 and 2017 show no great differences.

## 4. DETERMINANTS OF HEALTHCARE DEPRIVATION

In this section, we establish the determinants of healthcare deprivation in Spain by using regression techniques. In order to do this, healthcare deprivation determinants are analyzed from multidimensional approach, where all the dimensions are studied jointly, as well as from unidimensional approach, where each dimension is studied separately. The problem found in the data is the presence of an elevated number of zeros in the dependent variables *deprivation counts of the households, dental care, mental care*, *medication* and *medical care*. This characteristic of the data set may compromise the results obtained if it is not taken into account.

Zero inflated models developed by Lambert (1992) and later extended to ordinal outcomes with the zero inflated ordered logit (ZIOL) model by Anderson and Kelley (2008) are proposed as estimation method. This is because they model more accurately the distribution of ordinal outcomes when the dependent variable exhibits zero inflation. Harris and Zhao (2007) have used these to estimate tobacco consumption and Lin and Tsai (2013) to model health survey data.

These models also address the problem of the presence of two types of zeros that are present in our sample and are the reason why are considered as inflated. On the one hand, we have those zeros representing the households that have not been able to afford healthcare services; and on the other hand, those representing the households that have not needed to use them.

## 4.1 METHODOLOGY

Taking into account that the dependent variables are quantitative and ordered, the estimation model that should be used is an ordered logit. However, given the zero-inflation problem evaluated, the extension of the zero inflated model's framework to ordinal outcomes with the ZIOL model proposed by Anderson and Kelley (2008) is used to regress *deprivation counts of the households* $(d_i)$ and the unidimensional variables: *dental care, mental care, medication* and *medical care*.

Lets define the susceptibility of household $i$ of being a user of healthcare services as $I_i = 1$ and not being a user as $I_i = 0$ This variable depends on a latent variable, $I_i^*$, such that: $I_i = 1$ for $I_i^* > 0$ and $I_i = 0$ for $I_i^* \leq 0$. This latent variable represents the utility derived by the households of being users of healthcare services, which can be determined by:

$$I_i^* = \gamma z_i + \varepsilon_i.$$

Where in our case $z_i$ is the instrumental variable, $\gamma$ is the coefficient and $\varepsilon_i$ is the error term. So, the probability of being a user is determined by a logit model used by the ZIOL inflation equation:

$$Pr(I_i = 1|z_i) = Pr(I_i^* > 0|z_i) = F(\gamma z_i),$$

where $F(\cdot)$ is the logistic distribution function: $F(\varepsilon) = e^{\varepsilon}/(1 + e^{\varepsilon})$.

Let $\widetilde{d_i}$ be the score of deprivation by a user of healthcare services. Then, an ordered logit model is used to model outcome intensity levels $\widetilde{d_i}$ conditioned on $I_i = 1$. Their corresponding probabilities being:

$$Pr\left(\widetilde{d_i} = h|I_i = 1, x_i\right) = F(\kappa_h - \beta x_i) - F(\kappa_{h-1} - \beta x_i), \quad (1)$$

where $\mathrm{x_i}$ is a vector of covariates, $\beta$ is a vector of coefficients and cutpoints $\kappa_h$ are boundary parameters of the scores, both to be estimated.

The observed response variable is defined as $d_i = I_i\widetilde{d_\iota}$. Hence, a zero outcome will occur when $I_i = 1$ and $\widetilde{d_\iota} = 0$ (household $i$ can afford healthcare services) or when $I_i = 0$ (household $i$ is a nonuser of healthcare services and is considered as an excess of zero). So, to observe a positive $d_i$, it is necessary that $I_i = 1$ and $\widetilde{d_\iota} > 0$. Therefore, the distribution of the variable $d_i$ can be expressed as following.

$$Pr(d_i) = \begin{cases} Pr(d_i = 0|z_i, x_i) \\ Pr(d_i = h|z_i, x_i) \quad h = 1, 2, 3{,}4 \end{cases}$$
$$= \begin{cases} Pr(I_i = 0|z_i) + Pr(I_i = 1|z_i)Pr\left(\widetilde{d_\iota} = 0|I_i = 1, x_i\right) \\ Pr(I_i = 1|z_i)\, Pr\left(\widetilde{d_\iota} = h|I_i = 1, x_i\right) \end{cases} \qquad (2).$$

We can observe how the probability of zero is inflated because it is the result of the sum of the probability of not being healthcare deprived from the ordered logit model and the probability of being a non-healthcare user from the logit model. Then, combining equations (1) and (2), the model estimates the following distribution of the observed variable $d_i$:

$$Pr(d_i) = \begin{cases} Pr(d_i = 0|z_i, x_i) \\ Pr(d_i = h|z_i, x_i) \quad h = 1, 2, 3 \\ Pr(d_i = 4|z_i, x_i) \end{cases}$$
$$= \begin{cases} \{1 - F(\gamma z_i)\} + F(\gamma z_i)F(\kappa_0 - \beta x_i) \\ F(\gamma z_i)\{F(\kappa_h - \beta x_i) - F(\kappa_{h-1} - \beta x_i)\} \quad h = 1, 2, 3. \\ F(\gamma z_i)\{1 - F(\kappa_3 - \beta x_i)\} \end{cases}$$

The parameters ($\gamma, \beta$ and $\kappa_h$) can be consistently and efficiently estimated using maximum likelihood criteria. Being the log-likelihood function as follows:

$$ln\, L = \sum_{i=1}^{N} w_i \sum_{h=0}^{\hat{h}} J(d_i = h)\, ln\{Pr(d_i = h|z_i, x_i)\},$$

where *N* is the number of observations, $\hat{h}$ is the maximum potential value of the score, $w_i$ the weight associated to each of the observations in the sample and $J(d_i = h) = \begin{cases} 1 & if\ d_i = h \\ 0 & otherwise \end{cases}$.

In applying this method to our case, we use as instrumental variable *Self-assessed health perception with two categories* and if we find any discrepancies in the postestimations tests, we use *Self-assessed health perception with five categories*. This is because the endogenous variable may need a more complex differentiation in the categories of the *self-assessed health* variable. Moreover, the covariates used in estimating (1) are: *number* *of retired and incapacitated* members in the household, *area*, *household size, income* per household member and *region*.

In order to check the regressions' results with a Hausman specification test (Hausman, 1978) for significant differences between ZIOL and ordered logit estimates. This test checks the consistency of the estimators. Also, the Akaike information criterion (AIC) (Akaike, 1974) and Bayesian information criterion (BIC) (Stone, 1979) are used to determine whether the ZIOL model is a better fit than an ordered logit model. These last two criteria indicate that a model is a better fit when it presents lower AIC and BIC.

The variables *dental care, mental care, medication* and *medical care* are regressed using the same method and set of covariates in the estimations as well as checking the results with the Hausman test and AIC and BIC. It should be noted as well that the levels coded of the ordinal response variables will be 0 and 1.

## 4.2 RESULTS

The regression analysis is divided in two subsections. First one, the endogenous variable *deprivation counts of the households* is analyzed. Second, each dimension considered to measure healthcare deprivation (*dental care, mental care, medication* and *medical care)* is considered as an independent variable. This differentiation is used to provide information about the multidimensional and unidimensional approach.

### 4.2.1 MULTIDIMENSIONAL ANALISYS; $d_i$

Table 4 presents the estimation results for the multidimensional regression using as dependent variable, *deprivation counts of the households* considering the four dimensions jointly, $d_i$. The regression includes as regressors *number of retired and incapacitated* members in the household, *area*, *household size, income* per household member and the *regions* through dummy variables. It is estimated for 2014, 2017 and 2020 waves. For an easier interpretation, the odd ratios are obtained after exponentiating the coefficients from the ZIOL model.

*Table 4: ZIOL regression odds ratios. Multidimensional case: dependent variable. $d_i$*

| **Variables** | $d_i$ | | |
|---|---|---|---|
| | **2014** | **2017** | **2020** |
| Retired and incapacitated | 1.189***<br>(0.061) | 1.341***<br>(0.077) | 1.343***<br>(0.063) |

| | | | |
|---|---|---|---|
| Area | 1.427*** | 1.304*** | 1.266*** |
| | (0.143) | (0.128) | (0.103) |
| Household size | 0.813** | 0.702*** | 0.862** |
| | (0.077) | (0.068) | (0.064) |
| Income | 0.432*** | 0.493*** | 0.329*** |
| | (0.063) | (0.040) | (0.021) |
| Region ("Madrid" as reference) | | | |
| Andalucía | 1.997*** | 2.435*** | 0.618*** |
| | (0.275) | (0.403) | (0.085) |
| Aragón | 0.415*** | 0.422*** | 0.646** |
| | (0.091) | (0.0.074) | (0.141) |
| Asturias | 0.335*** | 0.583*** | 1.215 |
| | (0.079) | (0.104) | (0.195) |
| Baleares Islands | 0.726 | 0.204*** | 1.082 |
| | (0.150) | (0.051) | (0.278) |
| Canarias | 4.109*** | 1.117 | 5.090*** |
| | (0.639) | (0.260) | (0.719) |
| Cantabria | 0.139*** | 0.103*** | 0.095*** |
| | (0.044) | (0.033) | (0.037) |
| Castilla y León | 0.522*** | 0.272*** | 0.800 |
| | (0.100) | (0.058) | (0.144) |
| Castilla La Mancha | 0.941 | 0.744** | 1.464** |
| | (0.169) | (0.121) | (0.220) |
| Cataluña | 1.190 | 1.180 | 1.378** |
| | (0.159) | (0.194) | (0.185) |
| Valencia | 1.252** | 1.636*** | 1.794*** |
| | (0.174) | (0.213) | (0.233) |
| Extremadura | 1.014 | 0.235*** | 0.408*** |
| | (0.235) | (0.055) | (0.090) |
| Galicia | 1.583 | 0.506*** | 2.712*** |
| | (0.807) | (0.105) | (0.405) |
| Murcia | 1.029 | 4.505*** | 2.867*** |
| | (0.174) | (0.840) | (0.430) |
| Navarra | 0.461*** | 0.409*** | 3.052*** |
| | (0.108) | (0.083) | (0.528) |
| Basque Country | 0.564*** | 1.138 | 0.923 |
| | (0.102) | (0.143) | (0.154) |
| La Rioja | 0.716 | 0.568** | 1.351 |
| | (0.160) | (0.134) | (0.251) |
| Ceuta | 3.203*** | 0.014*** | 0.136*** |
| | (1.361) | (0.014) | (0.071) |
| Melilla | 2.706** | 1.499*** | 1.870** |
| | (1.045) | (0.260) | (0.550) |
| Constant | 0.336*** | 0.284*** | 0.123*** |
| | (0.076) | (0.030) | (0.020) |
| $\kappa_0$ | -0.321 | -0.559 | 0.726 |
| | (0.535) | (0.311) | (0.183) |
| $\kappa_1$ | 1.910 | 1.516 | 2.566 |
| | (0.336) | (0.233) | (0.176) |
| $\kappa_2$ | 3.426 | 2.785 | 3.997 |
| | (0.319) | (0.230) | (0.193) |
| $\kappa_3$ | 5.099 | 4.072 | 5.374 |
| | (0.348) | (0.251) | (0.262) |
| Households (sample) | 22,817 | 23,073 | 22,053 |

| | | | |
|---|---|---|---|
| Statistical representation | 183e+05 | 178e+05 | 188e+05 |
| Hausman test | 105.10*** | 56.18*** | 41.69** |
| AIC | 1.87e+07 | 1.74e+07 | 1.35e+07 |
| AIC (no inflation) | 1.90e+07 | 1.77e+07 | 1.38e+07 |
| BIC | 1.87e+07 | 1.74e+07 | 1.35e+07 |
| BIC (no inflation) | 1.90e+07 | 1.77e+07 | 1.38e+07 |

Note: standard errors are reported in parentheses.
Data is from INE; NHS and EHSS.
*, **, *** represent 10%, 5%, and 1% levels of statistical significance respectively.

*Source: Own elaboration, data from INE*

Therefore, we can conclude that variables *retired and incapacitated, area, household size,* and *income* are all statistically significative. It can be observed that, *ceteris paribus,* an increase of one member who is incapacitated or retired in the household will increase a healthcare service user's odds of being healthcare deprived by a factor of 1.189 in 2014, 1.341 in 2017 and 1.343 in 2020. Also, if the household is located in an urban area, *ceteris paribus*, it will increase a healthcare service user's odds of being healthcare deprived by a factor of 1.427 from those who are located in a rural area in 2014. Whereas, a rise in income per household member of 1,000 euros will reduce the odds of a healthcare service user of being deprived by a factor of 0.432 in 2014, and also reduces them in the years 2017 and 2020.

Regarding the regions, Andalucía, Aragón, Cantabria, Valencia, Navarra, Ceuta and Melilla present significant differences for the three years with Madrid (reference group). These differences indicate whether they are better or worse off in terms of healthcare deprivation than Madrid during 2014, 2017 or 2020 depending on the odds ratio value. On the one hand, Aragón and Cantabria present positive differences (odds ratios lower than 1) for the three years. This means that they are better off in terms of healthcare deprivation than Madrid. On the other hand, Valencia and Melilla present odds ratios higher than 1 (negative differences) for the three years estimations which means that they are always worse off than Madrid in terms of healthcare deprivation.

Moreover, Andalucía presents significant worse values than Madrid for 2014 and 2017 in terms of healthcare deprivation. Nonetheless, it improved its health deprivation and achieved better deprivation values than Madrid in 2020. Lastly, for the case of Navarra during the years 2014 and 2017, the region is better off in terms of healthcare deprivation, whilst it is worse for 2020.

To check whether the zero inflated model is appropriate we provide the Hausman test[3] (results compiled in Table 4). We can conclude that the estimator of the ZIOL model is consistent and asymptotically efficient when the *self-assessed health perception* is used as instrument. The AIC and BIC error criteria shown in Table 4 also evidence that the ZIOL model better fits the data than the ordered logit one.

### 4.2.2 UNIDIMENSIONAL ANALISYS

For a more complete analysis, in this subsection, we proceed with a unidimensional analysis of the healthcare deprivation analyzing the four dimensions (*dental care, mental care, medication* and *medical care*) independently for the waves 2014, 2017 and 2020. The regressions include as regressors *number of retired and incapacitated* members in the household, *area*, *household size, income* per household member and the *regions* through dummy variables. As mentioned above, odd ratios are presented for an easier interpretation. Also, while initially *self-assessed health perception with two categories* was used as instrument, the *self-assessed health perception with five categories* was finally used as instrument for the estimations of *medical care.* This is due to the need of a more complex and complete instrument.

The estimation results can be found in Table 5 of the Appendix 3. We highlight that *income* is statistically significant for all the independent variables and for all years. For *dental care* the variables *retired and incapacitated and area* are also relevant for the three years. In the case of *mental care*, the relevant variables for all three years are *retired and incapacitated* and *income*, nevertheless, the variable *area* is relevant just for the last two waves, 2017 and 2020. However, for the cases *medication* and *medical care* as endogenous variables, only the variable *income* is found to be significative for the three years.

Lastly, we can see that, in the case of *dental care,* Aragón and Cantabria present significant differences for the three years with Madrid (reference group) and these differences are positive (odds ratio lower than 1). This means that they are better off in terms of access to *dental care* services than Madrid for the three years estimations. Navarra and La Rioja also present significant better values in access to *dental care* than Madrid for the years 2014 and 2017. However, both regions worsened its access and

[3] Also called the Durbin-Wu-Hausman test

obtained worse values than Madrid in 2020. Also, Andalucía presents significant differences for the three years, which were negative for 2014 and 2017 and positive in 2020.

Moreover, when considering *medication* as the independent variable we observe that Canarias, Valencia and Galicia present significant negative differences (odds ratios higher than 1) with Madrid for the three years. This means that for the three years, these regions exhibit worse access to *medication* due to financial problems by households than Madrid. It should be added that for the case of Navarra in 2014, the region presented significant negative differences with Madrid but for 2020 it was able to improve access to *medication*, and these differences turned positive (odds ratio lower than 1).

While using *medical care* as endogenous, the regression results provide significant information about Canarias, Galicia and Melilla. These regions present significant negative differences with Madrid for the three years analyzed, which means that access to *medical care* is worse in the three of them than in Madrid. There are no regions that present positive significant differences for any of three years.

To check whether the zero inflated model is appropriate for the estimations we provide the Hausman test in Table 5 of the Appendix 4. The results of Hausman test for *mental care* in 2014 evidences that the ordered logit model fit is more appropriate than ZIOL model given the instrument used, which should be studied further. Some other instruments must be used to check whether the obtained results are due to the instrument selected before changing the model specification and the estimator. Another candidate as instrument could be the 12-Item General Health Questionnaire (GHQ-12) also included in the surveys.

For the other estimation we can conclude that the estimator of the ZIOL model is consistent and asymptotically efficient when the *self-assessed health perception* is used as instrument. This conclusion is supported by the AIC and BIC error criteria, found in Table 5 of the Appendix 4, which show evidence that the ZIOL model better fits the data than the ordered logit one as well.

## 5. CONCLUSIONS

The purpose of this study is to inform about the healthcare deprivation in Spain and to establish the determinants of the healthcare deprivation for the years 2014, 2017 and 2020 using the data obtained from the surveys European Survey of Health in Spain (EHSS) and National Health Survey (NHS), both from the National Statistics Institute (INE).

We conducted the first part of the analysis by computing the selected deprivation measures $H_m$, $P_m^1$ and $P_m^2$ for Spain in 2014, 2017 and 2020. These measures are calculated also for the different regions in Spain and differentiating between rural and urban areas. According to the results analyzed throughout we can conclude that in in Spain, there is healthcare deprivation. Moreover, there was an increase in healthcare deprivation in 2017 form 2014, whereas it decreased in 2020. Also, there are regions that are consistently better in terms of healthcare deprivation (Aragón, Asturias and Cantabria) whilst there are others that are consistently worse (Canarias and Galicia). Lastly, we can also conclude that urban areas present worse deprivation values than rural areas despite the initial idea that better access to healthcare is available in urban areas.

Then, we conducted the second part of the analysis by using the zero inflated ordered logit models to establish which the determinants of healthcare deprivation are. This part is divided in a multidimensional and unidimensional analysis. In the multidimensional analysis the sum of the dimensions in which a household $i$ is deprived is used as the endogenous variable, and the variables *retired and incapacitated, area, household size,* and *income* were found statistically significant for the three years. In the unidimensional analysis, the endogenous variable considered and regressed are the four dimensions: *dental care, mental care, medication* and *medical care*. According to the results obtained *retired and incapacitated, area* and *income* were relevant for the three years when we used *dental care* as endogenous; *retired and incapacitated* and *income* were found relevant when *mental care* was used as endogenous; and when *medication* or *medical* care were used as dependent variables, *income* was the only variable found relevant. Finally, regarding the regions in Spain, we can say that Aragón and Cantabria are consistently better or similar in terms of healthcare deprivation than Madrid and, also, that Valencia and Murcia are the regions that are consistently worse off or similar than Madrid in terms of healthcare deprivation.

It has been found that for the estimations of *deprivation counts of the households* for 2017 and of *medical care* for the three years, the variable used as instrument should be *self-assessed health perception with five categories* because it provides more information about the self-assessed health and it is needed for the estimations to be consistent. For all the other regressions conducted, no problems were found with the instruments used and, therefore, the estimations obtained can be considered as consistent.

The results obtained with Hausman test show that the estimator of the ZIOL model is consistent and asymptotically efficient when the *self-assessed health perception (two and five categories)* is used as instrument. However, there is an exception which is for the estimation of *mental care* in 2014, since the test evidences that the ordered logit model fit is more appropriate than ZIOL model. Therefore, further research about *mental care* will be interesting, especially a deep regional analysis of the 12-Item General Health Questionnaire (GHQ-12).

## REFERENCES


Akaike, H. (1974). A new look at the statistical model identification . *IEEE transactions on automatic control, 19(6)*, pp. 716-723.

Aldi Hagenaars, Klaas de Vos & M. Asghar Zaidi. (1994). *Poverty Statistics in the Late 1980s: Research Based on Micro-data,*. Luxembourg: Office for Official Publications of the European Communities.

Amnesty International. (2019). *Wrong Prescription: The Impact of Austerity Measures on the Right to Health in Spain.* London: Amnesty International Ltd.

Bagozzi, B.E., Hill Jr, D.W., Moore, W.H. & Mukherjee, B. (2015). Modeling two types of peace: The zero-inflated ordered probit (ZiOP) model in conflict research. *Journal of Conflict Resolution, 59(4)*, pp. 728-752.

Bourguignon, F., & Chakravarty, S. R. (1999). A family of multidimensional poverty measures. . In *Advances in econometrics, income distribution and scientific methodology* (pp. 331-344). Physica-Verlag HD.

Bourguignon, F., & Chakravarty, S. R. (2019). The measurement of multidimensional poverty. . In *In Poverty, social exclusion and stochastic dominance* (pp. 83-107). Springer, Singapore.

Cowell, F. (1988). Poverty Measures, Inequality and Decomposability. In *Welfare and Efficiency in Public Economics, Bös, D., Rose, M., Seidl, C. (eds)* (pp. 149-166). Berlin: Springer, Berlin, Heidelberg.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pp. 1251-1271.

Health Ministry. (2022). *Informe Anual del Sistema Nacional de Salud 2020-21.* Madrid.

INE. (2020). *Metodologia. Encuesta Europea de Salud en España 2020.* Retrieved from INE: www.ine.es/metodologia/t15/t153042020.pdf

Justyna Rój. (2020). Inequality in the Distribution of Healthcare Human. *Sustainability*, 12(5), pp. 2043.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. . *Technometrics,* 34(1)*,* pp. 1-14.

Lasso de la Vega, C. (2010). Chapter 7 Counting poverty orderings and deprivation curves, Bishop, J.A. (Ed.). In C. L. Vega, *Studies in Applied Welfare Analysis: Papers from the Third ECINEQ Meeting (Research on Economic Inequality, Vol. 18)* (pp. 153-172). Emerald Group Publishing Limited, Bingley.

Lin, T. H. (2013). Modeling health survey data with excessive zero and K responses. *Statistics in Medicine,* 32(9), pp. 1572-1583.

Maite Blázquez, Elena Cottini and Ainhoa Herrarte. (2014). The socioeconomic gradient in health: how important is material deprivation? *The Journal of Economic Inequality*, 12(2), pp. 239-264.

Marcel Bilger and Vincenzo Carrieri. (2013). Health in the cities: When the neighborhood matters more than income. *Journal of Health Economics*, 32(1), pp. 1-11.

Mark McGillivray, Indranil Dutta & Nore Markova. (2009). Health inequality and deprivation. *Health Economics*, 18(S1), S1-S12.

Mark N. Harris and Xueyan Zhao. (2007). A zero-inflated ordered probit model, with an application to modelling tobacco consumption. *Journal of Econometrics,* 141(2), pp. 1073-1099.

Mary E. Kelley and Stewart J. Anderson. (2008). Zero inflation in ordinal data: Incorporating susceptibility to response through the use of a mixture model. *Statistics in Medicine*, 27(18), pp. 3674-3688.

Michaela Benzeval and Ken Judge. (2001). Income and health: the time dimension. *Social Science & Medicine*, 52(9), pp. 1371–1390.

Ministry of Transport, Mobility and Urban Agenda, Directorate General of Housing and Land. (2021). *Áreas Urbanas en España, 2021.* Madrid: Ministry of Transport, Mobility and Urban Agenda, Secretaría Central Técnica, Centro de Publicaciones.

OECD . (2019). *Health for Everyone?: Social Inequalities in Health and Health Systems.* Paris: OECD Health Policy Studies, OECD Publishing.

Padeiro, M. (2018). Geographical accessibility to community pharmacies by the elderly in metropolitan Lisbon. *Research in social and administrative pharmacy,* 14(7*)*, pp. 653-662.

SabinaAlkire and JamesFoster. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, Vol. 95, Issues 7-8, pp. 476-487.

Satya R. Chakravarty and Conchita D'Ambrosio. (2006). The measurement of social exclusion. *Rev. Income Wealth*, 52(3), pp. 377-398.

Stone, M. (1979). Comments on model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 276-278.

Townsend, P. (1987). Deprivation. *Journal of Social Policy,* 16(2), pp. 125-146.

## ANNEX 1: Variables

| Variable name | Variable description | % or Mean (SD) |
|---|---|---|
| Identification number | Number associated to identified each of the households | |
| Gender of the reference adult | =1 female; =0 male | 0.359 (0.480) |
| Age of the reference adult | Age in years | 56.814 (16.280) |
| Number adults | Number of adults in the household | 2.100 (0.958) |
| Number minors | Number of minors in the household | 0.357 (0.719) |
| Employed status | =1 employed<br>=2 unemployed<br>=3 retired<br>=4 studying<br>=5 incapacitated<br>=6 house work<br>=7 others | 51.88%<br>8.12%<br>34.64%<br>0.43%<br>2.16%<br>2.59%<br>0.17% |
| Education | =0 less than primary school<br>=1 primary<br>=2 secondary<br>=3 bachillerato<br>=4 professional training, medium level<br>=5 professional training, high level<br>=6 higher education<br>=7 NA | 12.02%<br>22.18%<br>19.77%<br>12.07%<br>6.81%<br>7.39%<br>19.52%<br>0.24% |
| Highest level of education in the household | =0 less than primary school<br>=1 primary<br>=2 secondary<br>=3 bachillerato<br>=4 professional training, medium level<br>=5 professional training, high level<br>=6 higher education<br>=7 NA | 7.48%<br>15.29%<br>17.77%<br>12.62%<br>9.03%<br>10.11%<br>27.11%<br>0.59% |
| Lowest level of education in the household | =0 less than primary school<br>=1 primary<br>=2 secondary<br>=3 bachillerato<br>=4 professional training, medium level<br>=5 professional training, high level<br>=6 higher education | 34.52%<br>27.21%<br>26.46%<br>8.46%<br>3.00%<br>2.90%<br>7.45% |

*Source: Own elaboration, data from INE*

## ANNEX 2: Proposition *FD* curves

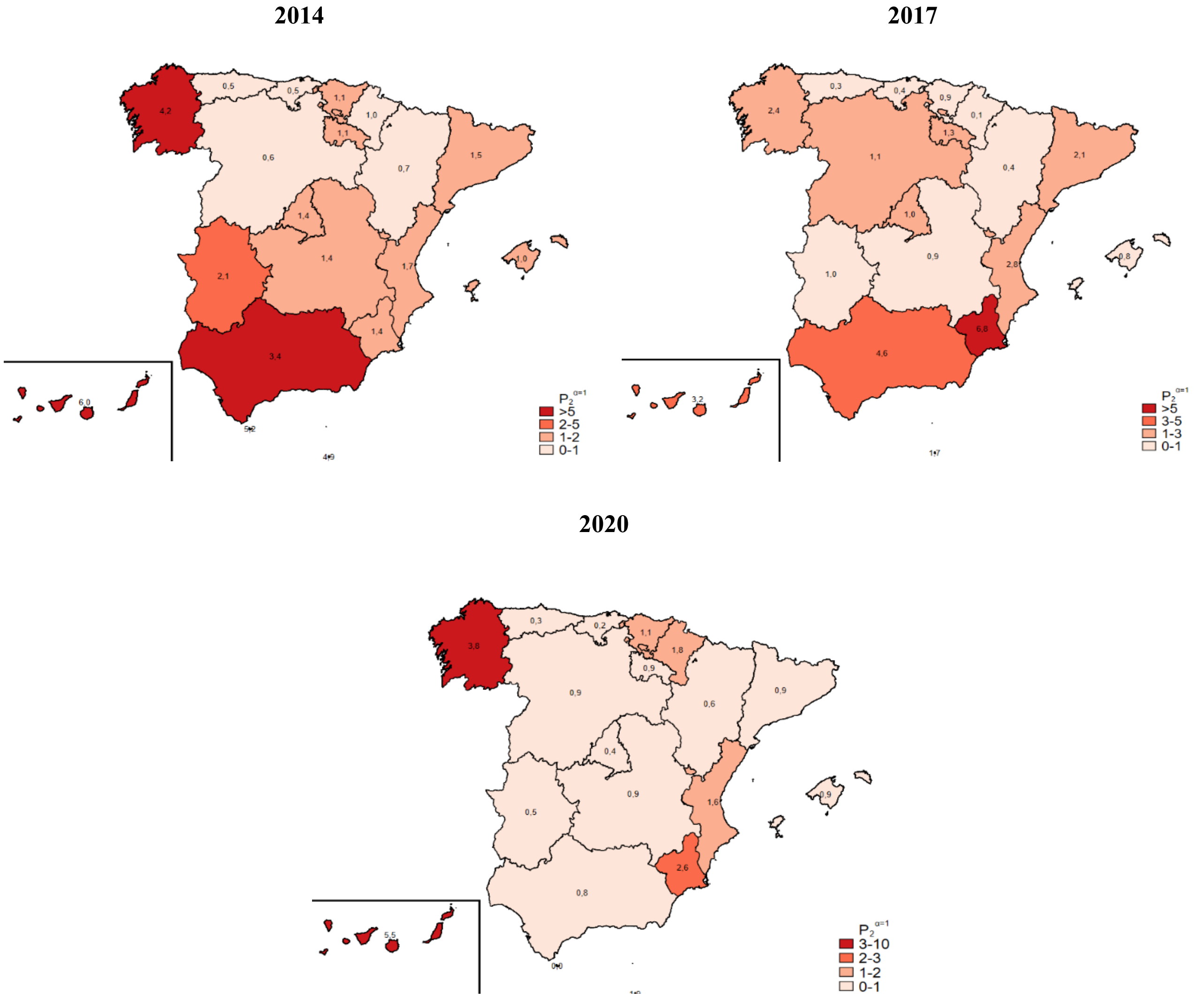
**Prop. 1.** *For any* $\boldsymbol{d}, \boldsymbol{d}' \in G$ *vectors of weighted deprivation counts:*

$$FD(\boldsymbol{d}'; p) \geq FD(\boldsymbol{d}; p) \textit{ for all } p \in [0, k]$$

*If and only if* $P_m(\boldsymbol{d}') \geq P_m(\boldsymbol{d})$ *for all* $P_m \in \boldsymbol{P}_1$ *and for all cut-off* $m \in (0, k]$.

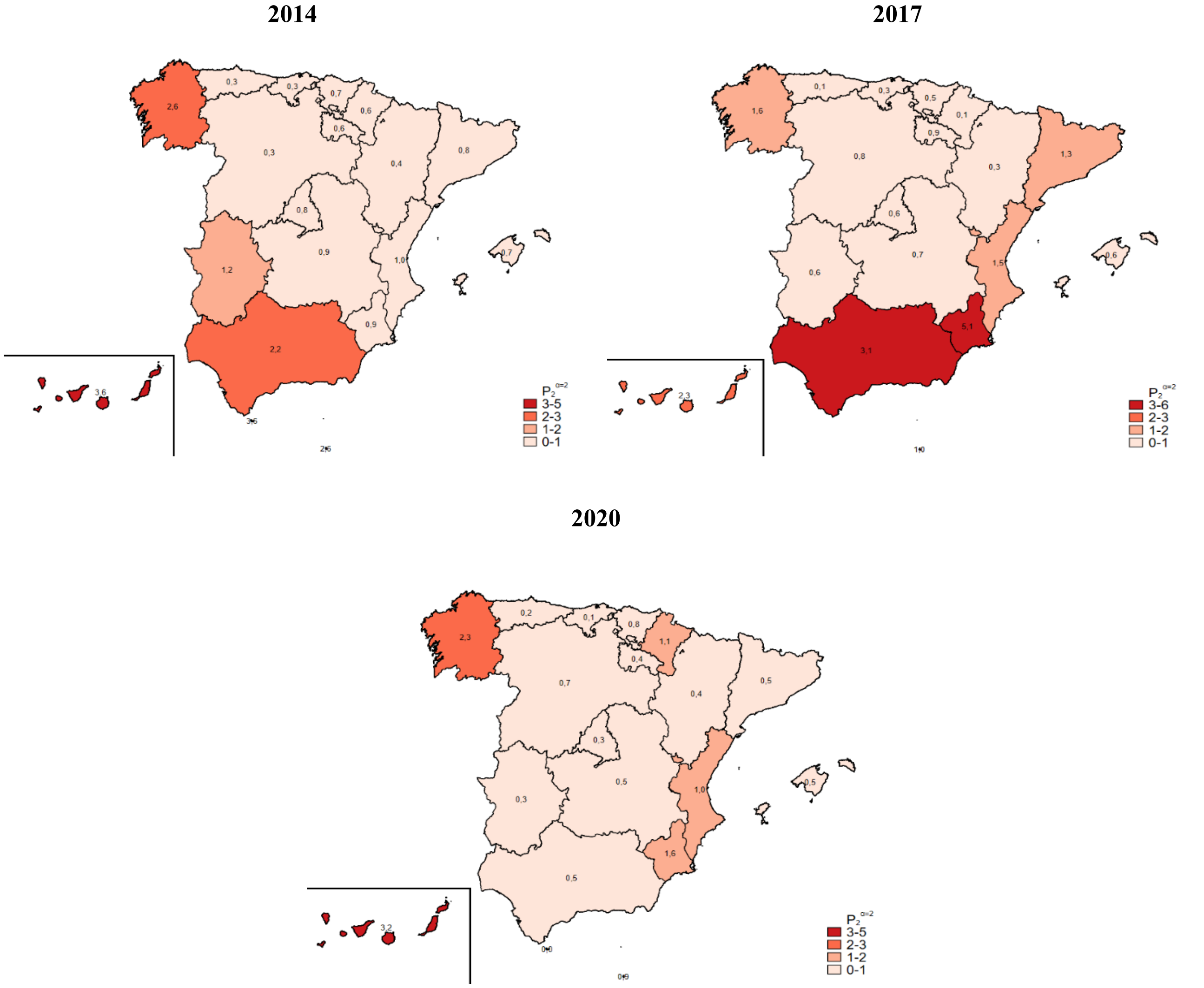
## ANNEX 3: Healthcare deprivation by region in 2014 ($P_2^1$ and $P_2^2$).

*Figure 2: Healthcare deprivation by region in 2014, 2017, 2020 ($P_2^1$)*




*Source: Own elaboration, data from INE*

*Figure 3: Healthcare deprivation by region in 2014, 2017, 2020 ($P_2^2$)*




*Source: Own elaboration, data from INE*

## ANNEX 4: Unidimensional Regressions Results

*Table 5: ZIOL regression odds ratios. Unidimensional case: dependent variable: dental care, mental care, medication* and *medical care*

| **Variables** | **Dental care** | | | **Mental care** | | | **Medication** | | | **Medical care** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **2014** | **2017** | **2020** | **2014** | **2017** | **2020** | **2014** | **2017** | **2020** | **2014** | **2017** | **2020** |
| Retired and incapacitated | 1.298*** | 1.211** | 1.320*** | 1.493*** | 1.478*** | 1.640** | 1.204** | 1.368*** | 1.097 | 1.108 | 1.332*** | 1.387*** |
| | (0.098) | (0.108) | (0.063) | (0.212) | (0.163) | (0.410) | (0.093) | (0.141) | (0.104) | (0.097) | (0.121) | (0.131) |
| Area | 1.455*** | 1.231** | 1.240*** | 1.101 | 1.562** | 4.067*** | 1.487*** | 1.407** | 1.111 | 1.324* | 1.485** | 1.197 |
| | (0.165) | (0.107) | (0.099) | (0.266) | (0.318) | (1.967) | (0.207) | (0.213) | (0.226) | (0.200) | (0.253) | (0.212) |
| Household size | 0.840 | 0.873 | 0.863*** | 0.525*** | 0.536*** | 0.557 | 0.853 | 0.672** | 1.322 | 0.945 | 0.677** | 0.902 |
| | (0.097) | (0.088) | (0.062) | (0.133) | (0.108) | (0.199) | (0.113) | (0.124) | (0.235) | (0.147) | (0.103) | (0.153) |
| Income | 0.395*** | 0.592*** | 0.330*** | 0.504*** | 0.678*** | 0.373*** | 0.460*** | 0.480*** | 0.409*** | 0.571*** | 0.676*** | 0.549 |
| | (0.075) | (0.110) | (0.021) | (0.092) | (0.084) | (0.119) | (0.043) | (0.073) | (0.077) | (0.066) | (0.070) | (0.067) |
| Region (“Madrid” as reference) | | | | | | | | | | | | |
| Andalucía | 1.798*** | 1.447*** | 0.602*** | 1.075 | 2.002*** | 3.236** | 2.510*** | 4.429*** | 0.561 | 2.751*** | 4.354*** | 1.424 |
| | (0.296) | (0.174) | (0.081) | (0.361) | (0.490) | (1.920) | (0.463) | (1.356) | (0.226) | (0.666) | (1.226) | (0.609) |
| Aragón | 0.368*** | 0.546** | 0.540*** | 0.277** | 0.441 | 0.000 | 0.325*** | 0.192*** | 1.102 | 0.846 | 0.720 | 2.217 |
| | (0.100) | (0.129) | (0.127) | (0.167) | (0.225) | (.) | (0.140) | (0.104) | (0.508) | (0.318) | (0.359) | (1.131) |
| Asturias | 0.224*** | 0.693* | 1.360* | 0.274* | 0.000*** | 1.234 | 0.662 | 0.432** | 0.195** | 0.777 | 0.828 | 0.859 |
| | (0.067) | (0.141) | (0.219) | (0.189) | (0.00) | (1.021) | (0.211) | (0.182) | (0.149) | (0.336) | (0.393) | (0.531) |
| Baleares Islands | 0.676 | 0.259*** | 1.103 | 0.162* | 0.126** | 8.326** | 0.863 | 0.448** | 0.766 | 1.813 | 1.885 | 1.101 |
| | (0.182) | (0.090) | (0.277) | (0.177) | (0.129) | (7.772) | (0.301) | (0.185) | (0.524) | (0.662) | (0.792) | (0.917) |
| Canarias | 2.948*** | 0.951 | 4.227*** | 0.987 | 1.046 | 3.888** | 1.357 | 1.216 | 1.258 | 23.72*** | 4.636*** | 17.271*** |
| | (0.919) | (0.157) | (0.555) | (0.421) | (0.348) | (2.573) | (0.335) | (0.331) | (0.531) | (7.466) | (1.476) | (6.563) |
| Cantabria | 0.083*** | 0.077*** | 0.087*** | 0.496 | 0.383 | 0.447 | 0.308*** | 0.238*** | 0.261* | 0.805 | 1.681 | 0.311 |
| | (0.033) | (0.039) | (0.035) | (0.284) | (0.289) | (0.536) | (0.138) | (0.129) | (0.200) | (0.370) | (0.822) | (0.251) |
| Castilla y León | 0.488*** | 0.304*** | 0.797 | 0.423* | 0.706 | 6.587** | 0.284*** | 0.574 | 1.094 | 0.607 | 2.026* | 1.275 |
| | (0.115) | (0.092) | (0.144) | (0.203) | (0.287) | (5.461) | (0.104) | (0.203) | (0.504) | (0.248) | (0.742) | (0.676) |
| Castilla La Mancha | 0.989 | 0.843 | 1.527*** | 1.096 | 0.661 | 1.984 | 0.684 | 0.877 | 0.393* | 1.169 | 1.269 | 2.048 |
| | (0.223) | (0.151) | (0.227) | (0.515) | (0.284) | (1.417) | (0.217) | (0.277) | (0.217) | (0.407) | (0.512) | (0.969) |
| Cataluña | 1.311 | 1.017 | 1.379** | 0.640 | 0.704 | 4.372** | 0.672 | 0.903 | 1.093 | 1.921** | 3.997*** | 1.436 |
| | (0.239) | (129) | (0.184) | (0.256) | (0.209) | (2.956) | (0.164) | (0.218) | (0.418) | (0.517) | (1.152) | (0.642) |
| Valencia | 0.953 | 1.484*** | 1.819*** | 0.632 | 1.729** | 9.623*** | 1.488* | 1.766** | 1.958* | 3.051*** | 2.760*** | 1.716 |
| | (0.180) | (0.185) | (0.231) | (0.276) | (0.455) | (7.171) | (0.317) | (0.416) | (0.678) | (0.792) | (0.856) | (0.759) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Extremadura | 0.714 | 0.261*** | 0.302*** | 1.352 | 1.163 | 1.004 | 0.698 | 0.379** | 0.706 | 3.637*** | 1.392 | 2.092 |
| | (0.179) | (0.086) | (0.074) | (0.632) | (0.482) | (0.01) | (0.243) | (0.171) | (0.380) | (1.117) | (0.613) | (2.092) |
| Galicia | 0.740 | 0.559** | 2.514*** | 1.160 | 0.071** | 12.949*** | 2.957*** | 1.580* | 1.447 | 2.564*** | 2.222** | 10.088*** |
| | (0.162) | (0.128) | (0.349) | (0.446) | (0.072) | (9.633) | (0.603) | (0.417) | (0.549) | (0.711) | (0.766) | (3.900) |
| Murcia | 1.081 | 2.655** | 2.355*** | 0.835 | 3.773*** | 183.590 | 0.740 | 2.157*** | 1.496 | 1.390 | 15.792*** | 4.034*** |
| | (0.242) | (1.050) | (0.343) | (0.375) | (0.978) | (630.739) | (0.237) | (0.584) | (0.602) | (0.487) | (5.417) | (1.748) |
| Navarra | 0.302*** | 0.523** | 3.046*** | 0.673 | 0.163** | 14.941*** | 0.635 | 0.133*** | 2.464** | 1.974* | 0.649 | 1.735 |
| | (0.090) | (0.135) | (0.505) | (0.355) | (0.134) | (11.507) | (0.231) | (0.083) | (1.077) | (0.697) | (0.360) | (1.020) |
| Basque Country | 0.465*** | 1.541** | 0.846 | 0.576 | 0.162*** | 7.180** | 0.731 | 0.692 | 1.668 | 1.191 | 0.600 | 1.823 |
| | (0.105) | (0.288) | (0.142) | (0.279) | (0.090) | (5.525) | (0.210) | (0.202) | (0.622) | (0.401) | (0.256) | (0.832) |
| La Rioja | 0.605* | 0.561** | 1.389* | 0.147* | 0.419 | 5.583* | 0.919 | 0.709 | 1.467 | 1.429 | 3.174*** | 0.235 |
| | (0.168) | (0.139) | (0.253) | (0.155) | (0.235) | (5.795) | (0.341) | (0.280) | (0.722) | (0.584) | (1.235) | (0.252) |
| Ceuta | 2.230 | 0.000*** | 0.171*** | 0.925 | 0.000 | 0.000 | 4.935*** | 0.000 | 0.049*** | 3.566*** | 0.431 | 0.000 |
| | (1.105) | (0.000) | (0.089) | (0.792) | (.) | (.) | (1.598) | (.) | (0.052) | (1.746) | (0.453) | (.) |
| Melilla | 1.133 | 1.821** | 0.911 | 0.238 | 0.027*** | 0.000 | 1.092 | 0.266** | 0.814 | 13.82*** | 10.681*** | 15.029*** |
| | (0.432) | (0.533) | (0.291) | (0.231) | (0.028) | (.) | (0.467) | (0.156) | (0.627) | (5.477) | (5.147) | (7.401) |
| Constant | 0.266*** | 0.518 | 1.154 | 0.009*** | 0.252** | 0.017*** | 0.443*** | 0.152** | 0.358*** | 0.050*** | 0.034*** | 0.074*** |
| | (0.046) | (0.503) | (0.121) | (0.006) | (0.047) | (0.006) | (0.064) | (0.116) | (0.084) | (0.012) | (0.009) | (0.22) |
| $\kappa_0$ | -0.547 | 0.656 | 1.098 | 0.608 | 2.883 | 2.040 | 2.948 | 1.676 | 3.644 | 2.840 | 2.848 | 3.228 |
| | (0.046) | (1.109) | (0.161) | (0.835) | (0.380) | (0.761) | (0.297) | (1.004) | (0.409) | (0.362) | (0.521) | (0.487) |
| Households (sample) | 22,817 | 23,073 | 22,053 | 22,817 | 23,073 | 22,053 | 22,817 | 23,073 | 22,053 | 22,817 | 23,073 | 22,053 |
| Statistical representation | 183e+05 | 178e+05 | 188e+05 | 183e+05 | 178e+05 | 188e+05 | 183e+05 | 178e+05 | 188e+05 | 183 e+05 | 178e+05 | 188e+05 |
| Hausman test | 225.19*** | 112,426*** | 330.15*** | 15.57 | 58,924*** | 331.74*** | 397.90*** | 57,209*** | 125.83*** | 108.55** | 163.41*** | 87,778*** |
| AIC | 1.38e+07 | 1.26e+07 | 1.03e+07 | 1.73e+06 | 2.33e+06 | 1.67e+06 | 4.03e+06 | 4.36e+06 | 1.95e+06 | 4.16e+06 | 3.64e+06 | 3.04e+06 |
| AIC (no inflation) | 1.41e+07 | 1.28e+07 | 1.05e+07 | 1.83e+06 | 2.48e+06 | 1.76e+06 | 4.20e+06 | 4.53e+06 | 2.03e+06 | 4.31e+06 | 3.80e+06 | 3.12e+06 |
| BIC | 1.38e+07 | 1.26e+07 | 1.03e+07 | 1.73e+06 | 2.32e+06 | 1.67e+06 | 4.03e+06 | 4.36e+06 | 1.95e+06 | 4.16e+06 | 3.64e+06 | 3.04e+06 |
| BIC (no inflation) | 1.41e+07 | 1.28e+07 | 1.05e+07 | 1.83e+07 | 2.48e+06 | 1.76e+06 | 4.20e+06 | 4.53e+06 | 2.03e+06 | 4.31e+06 | 3.80e+06 | 3.12e+06 |

Note: standard errors are reported in parentheses.
Data is from INE; NHS and EHSS.
*, **, *** represent 10%, 5%, and 1% levels of statistical significance respectively

*Source: Own elaboration, data from INE*