Estimation of the price elasticity of the Spanish households' electricity demand

Ulysse Grand
Supervisors: Alaitz Artabe and Amaia de Ayala
July 29th 2022
Master in Economics: Empirical Applications and Policies, University of the Basque Country UPV/EHU

Abstract: In the past few months the electricity price has considerably increased in Spain due to major international crises which lead to instability in the market. To understand better how Spanish households change their consumption of electricity depending on the price, we estimate the value of price elasticity of the electricity demand to finally make some policy recommendations. For that, we have estimated the price demand elasticity through the demand function of Spanish households in 2019 with the data of the Spanish Survey of Family Budget 2019. The estimation of the elasticity has been done by Instrumental Variables after finding a problem of endogeneity in the price, and the final result is that we have an elasticity of -0.57 so the electricity demand seems to be price inelastic. Those estimations helped us to highlight that the price-based policies seem to have little effect on the electricity consumption of Spanish households in 2019, and so the government should complement tax policies with other kinds of measures, such as information policies and regulatory policies in order to help the households to pay less, but also to have a society with a lower impact on the environment. It would have a positive impact on the air conditioning high consumption in some regions of Spain and increase the energy efficiency in order to help households to reduce their expense of electricity but also, help to moderate the CO2 emissions.

Keywords: price elasticity, electricity demand, households demand, policy recommendations, electricity in Spain

Contents

1. Introduction

Household energy consumption represents around 26% of final energy consumption in Europe (Eurostat, 2020) and 17% in Spain (IDAE, 2020). Moreover, it is responsible for around 17% of global energy related carbon dioxide emissions in 2019 (IEA , 2019) and that is why, it is crucial to understand how the households use this energy. In 2020, the demand for electrical energy (249GWh) in Spain decreased by 5.5% compared to 2019. Furthermore, during this same year, 45.5% of the electricity has been generated with renewable energy, which is the maximum that has been reached so far.

The average final price of energy was 40.38 euros/MWh, 24.4% lower than the price in 2019 (Torra & Arnau, 2020). In the last months, the electricity price has considerably increased in Spain due to major international crisis and the instability of the price of other resources (Torra & Arnau, 2020). The electricity price in Spain is regulated by the Mercado Ibérico de Electricidade (MIBEL), which is the Iberian electricity market, the result of an agreement between the governments of Portugal and Spain: The price is determined for each hour and it is the same for both countries if there is not congestion in the net.

Even if the demand for electricity in Spain during 2020 and 2019 showed a decrease reaching 249.9 Gwh because of COVID-19, the current global political situation has led to historic spikes in energy prices in the international market including Spain (Torra & Arnau, 2020). Those consequences have resulted in some price policies from the Spanish government to reduce the expenditure of the Spanish households, such as taxing policy. We will try to figure out if they can be useful by checking the sensitivity of the electricity demand with respect to price.

In this master thesis, the price elasticity of the Spanish households' electricity demand is estimated for the year 2019. For that purpose, the variation of the quantity of electricity consumed in a year by a set household is analyzed concerning to the price and some other factors (Household variables, Geographical variables, Socio-demographic variables and House characteristic). We want to figure out to what extent a household changes its consumption of electricity depending on the price and how to characterize those households. This will allow us to reflect on the effectiveness of the price policies carried out so far as well as proposing some policy recommendations.

The electricity demand of a household (i.e. the quantity of electricity consumed in a specific year) is modeled as a function of the electricity price, but also other independent variables to take into account the heterogeneity of Spanish households, that can characterize the household, such as, Household variables, Geographical variables, Socio-demographic variables and House characteristic. The year 2019 has been selected because, as we have seen before, the data from after 2020 has been impacted by covid and we want to avoid this effect in our results. Although initially, the model is going to be estimated by Ordinary Least Square (OLS), we expect to have an endogeneity problem with the electricity price variable. An endogeneity test is conducted to analyze this possible endogeneity and the model is re-estimated by Instrumental Variables (IV).

During this paper, we will first introduce a literature review in order to have an overview of the electricity elasticity demand in the literature, then present the methodology we will use to estimate the elasticity in our case. After that, the data will be presented to finally describe the results we got, and make some conclusions and policy recommendations.

2. Literature review

The literature review has to depict the reality of the current research, the elasticity is a trendy tool to understand better the behavior of the consumers of electricity in order to stand up against climate change, as well as, to understand how consumers respond to changes in prices caused by different factors, such as, increase and decreases in taxes. In Spain, electricity represents 6.7%, 15.5% and 47% of all energy used for heating, water heating and cooking, respectively (Bueno et al, 2020). In Europe, some goals have been set, and in 2020 the objective was to reduce the GHG emissions by 20%, and 30% by 2030. Some improvements have be done especially in the building sector and way more fields (Ayala et al.,2016).

A high value of elasticity in this case, which is an elasticity value higher than 1, indicates that the increase in price causes proportionally a higher decrease in electricity demand. Thus, an increase in price implies a reduction of electricity demand and it could be seen as a solution to consume less. All those measures of elasticity lead us to understand the behaviors of the consumers in order to figure out global warming's solutions but also to prevent overconsumption with new policies. Furthermore, knowing the elasticity and the personal beliefs which affect the energy consumption can deliver some precious indication of the different policy which would be effective or not (Sola et al., 2020).

There is an increasing number of research in the field of understanding the factors driving energy consumption in the household sector. Many recent studies have focused on the demand price elasticity of many sources such as of electricity (Zhai et al., 2020), water (Hoyos & Artabe, 2016) or Diesel (Dahl, 2012).

The studies reviewed focusing on electricity will be divided into different groups: (i) those that apply the same methodology as the one we use in the present study to estimate the price elasticity: the OLS and (IV) estimation ; (ii) those which also try to point out an estimation of the elasticity but with another and distinct methodology (Generalized Least Squares, theoretical or mathematical methodologies,..) and (iii) those making policy recommendations.

Research on the price elasticity of the electricity demand or other resources have increased over the past 10 years, and some methodologies have been more used to estimate this value for example the OLS and (IV) methodologies. Regarding the first group of papers reviewed (i): the papers which apply the same methodology as the one we use in the present study to estimate the price elasticity, in Spain, an estimation of the price elasticity of the Spanish households' electricity demand has been done for each region (Romero-Jordan et al., 2014), in which the price elasticity is between -0.4 and -0.3. They conclude that the government should discourage taxes on the price of electricity and encourage renewable resources.

Some authors have also suggested other results for the same period but in a different European country. In Sweden, after analyzing the demand with an (IV) estimation, more precisely by Two-Stage Least Squares (2SLS), the result of the elasticity for each year from 1996 to 2009 is between -0.92 and -0.02 which shows over again that the value of the price elasticity of the demand can fluctuate quite a lot over years (Ericson & Rafatnia, 2011). Another study is focusing on European results of elasticity with the same methodology OLS and (IV) (Valitov et al., 2012). This paper, which uses data between 2010 and 2012, obtains distinct results by (2SLS) from what we have seen before: an elasticity between -1.4 and -0.8. As we can see, all those studies depict different realities of the database but also different results, and so a critical open question is what the situation with today's data is.

Let's now try to expand the scope of results and look at a similar country in terms of social dimension but outside Europe. In the United States, a more recent paper has been done, and it gives us more insights for the period of 2003 to 2015 (Burke, 2017). Even if the paper is not taking into consideration just the household, they conclude with an OLS and (IV) estimation that the estimated elasticity is around -0.9 which is larger (in obvolute value) than expected, and it can be explained by the electricity-intense industrial activities clustering in some areas, and more precisely in the low-price states (Mississippi, Oklahoma, …). In China, the same estimations have been run with data from 2006 to 2018 which is fairly recent, and surprisingly the result of the 2SLS estimate of elasticity is similar (around -0.9) to the one of the United States (Zhai et al., 2020). It makes us think that the period of analysis is even more important than the country itself, but also which perspective we should take into consideration: long or short term, industrial or residential, and even more.

Another angle that can be chosen is to follow a more engineer approach (Muller et al., 2018). The authors of the study have driven the further development of the importance of community-scale power infrastructure and the design of the off-grid systems to finally conclude its impact on the household price. This paper addresses a new challenge of modelization of the price to finally lead by

OLS to an estimation of elasticity in 2011 of -0.14, which is definitively lower than the previous estimated done in close countries like India (-0.63) (Bose et al., 1999). Let's now step back from all those papers which are using OLS and (IV), with a paper that did use the methodology but then criticized it in consequence of its lack of robustness to changes to the set of controls or to the sample definition and then use a second methodology which is the bunching estimate (Lanot & Vesterberg, 2021).

With regards to the second group of papers (i.e. other methodologies than OLS and (IV)), we identify the Generalized Least Squares (GLS) estimation as another technique to regress a model in order to estimate the demand price elasticity of energy. This methodology is often used when the model is suffering from heteroskedasticity, and that was the case to estimate the elasticity in the United States (Labandeira et al., 2017). The result is that the average price elasticity of the electrical demand is -0.21 which is lower (in absolute value) than what we have seen in the previous papers for the United States. Using the load profile of consumers is an alternative approach to estimate the elasticity, and it has been used in a Polish database which uses data in 2016 (Andruszkiewicz et al., 2019). The elasticity has been estimated between -2.3 and –1.7, which is high enough to find an explanation: the dataset is composed of people in Poland who are actually trying to reduce their expenditure on electricity. Finally, other methods are also used to highlight different facts to explain the problem of reducing our electricity expense in order to have less impact on the environment. Since the price elasticity of the electricity demand is increasingly relevant for all kinds of analysis of electricity, but obviously also for the global warming purposes, other methods can be tried like the dynamic demand model developed by Houthakker in 1974 (Bernstein & Griffin, 2006). This method has been used with data of the residential electricity consumption in the United States from 1989 to 1999, and a fairly low elasticity between -0.31 and -0.04 is obtained, which is in line with the lower absolute values obtained from the datasets before 2000.

Finally, the third group of papers are related with different policy measures. The literature distinguishes 3 types of policy instruments that can affect household consumption: (i) command and control instruments; (ii) price instruments; and (iii) informational instruments. The informational instruments include feedback, better decision process and some guide for the users. The audits are also a way to improve the energy efficiency because it increases the awareness (Labandeira et al., 2017) as well as the energy efficient label which can be used to guide the consumers to make an energy efficient choice and it is nowadays widely used (Sola et al., 2020). Some prices policies are also possible, like the taxation, loan facilities or subsidies. Finally, we have the policy interventions that focus on the behavioral choices of users (Cattaneo, 2019), which mainly upskill the user and his/her behavior for the long term instead of focusing on overall society solutions which are often short term.

In Spain, on the one hand, some papers are arguing that the policy of taxation of the electricity in order to increase the price is not a great idea since the demand is inelastic and therefore, the price changes does not have a measurable impact on the demand (Romero-Jordan et al., 2014). But it also says that one of the solutions for the long term would be the development of more renewable energy because it would decrease the price of the electricity over years.

On the other hand, some papers argue that a corrective pricing policy is a good way to reach the climate objectives of reduction of consumption for the long term in Europe (Labandeira et al., 2017). Other policy measures can also be implemented if we want to reduce the energy consumption of households based on their behaviors. The increased awareness of our consumption with constant feedback, and the promotion of new ways of utilization while fighting the barriers of the adoption of energy efficiency and its technologies (Cattaneo, 2019) are part of the solution. Finally, looking back on the previous policy is a good way to deem the effectiveness of policy for the future to finally promote new policies or to stick to what used to work. The policies done in different countries including standard, economic, or information instruments have between 7% and 10% impact on the electricity demand, so they are effective and promote the energy efficiency (Labandeira et al., 2017)..

Based on our literature review, we can conclude that the estimate for the demand price electricity is negative and most of the time between -0.7 and -0.1. However, the magnitude varies across countries, periods of time, methodologies, profiles of consumers and assumptions made. Table 1 summarizes the estimates of the electricity demand price elasticity of the studies reviewed.

Table 1 – Literature review of estimates of price elasticity of the electricity demand

Author	Region	Elasticity [min , max]
	Europe：	
(Krishnamurthy & Kristrom, 2015)	Spain	[-1.4,-0.27]
(Pellini, 2021)	Europe	[-0.8,-0.08]
(Valitov et al., 2012)	Europe	[-1.4,-0.8]
(Çetinkaya et al., 2015)	Turkey	-0.6
(Boogen et al., 2021)	Switzerland	[-0.68,-058]
(Romero-Jordan et al., 2014)	Spain	[-0.4,-0.3]
(Andruszkiewicz et al., 2019).	Poland	[-2.3,-1.7]
(Ericson & Rafatnia, 2011)	Sweden	[-0.92,-0.02]
	North America:	
(Li et al., 2021)	United States	-0.054
(Labandeira et al., 2017)	United States	-0.21
(Bernstein & Griffin, 2006)	United States	[-0.31,-004]
(Burke, 2017)	United States	-0.9
(Ito, 2014)	United States	-0.051
(Reiss & White, 2005)	United States	-0.39
	Asia:	
(Saha & Bhattacharya, 2018)	India	-0.49
(Bose et al., 1999)	India	-0.63)
(Muller et al., 2018)	India	-0.14
(Chindarkar & Goyal, 2019)	India	-0.39
(Yu & Xin, 2020)	China	-0.8
(Zhai et al., 2020)	China	-0.9
(Wang & Mogi, 2017)	Japan	-0.51

(Ishaque, 2018)	Pakistan	-0.46

Source: Own work based on a literature review

3. Methodology

3.1 Ordinary Least squares (OLS)

The econometric model chosen is the linear function form which is known for its simpleness in terms of estimation but also for the interpretation. It is assumed a random sample of size n from the population where i= $i = 1, \dots, n$. The model specification is as follows (Wooldridge, 2009), let's name this equation, equation 1:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + \dots + \beta_K x_{Ki} + u_i$$

where, y_i is the dependent variable, $(x_{1i}, \dots, x_{ki}, \dots, x_{Ki})$ are the K explanatory variables and u_i represents the error term for the observation i. The mostly used estimator of the β is Ordinary Least Squares (OLS) (Wooldridge, 2009). Under some assumptions, such as, random sampling of observations, no multi-collinearity, homoscedasticity, the exogeneity of the explanatory variables and linearity of the parameters in the model, the OLS estimator is unbiased and consistent which means that the expected value of the estimator matches the one of the parameters and that if we increase the size of the sample, the estimates tend to the real value of the parameters. When the assumption of homoscedasticity is not satisfied, the OLS is unbiased, but not efficient. In this case, a robust variance-covariance matrix in estimated for the parameters.

When the aim is to estimate the demand model of a good, we may face the non-fulfillment of one of the assumptions: the exogeneity of the explanatory variables. Endogeneity occurs when some of the explanatory variables is correlated with the error term. When this is the case, the OLS estimator is biased (Mariel, 2021).

A usual way to test the endogeneity of our model and to know if a suspected variable is endogenous or not is using the Wooldridge endogeneity test (Wooldridge, 2009) where the result of the OLS and the 2SLS estimates are compared, determining whether the differences are statistically significant or not (Wooldridge, 2009). For that, we will basically test if the estimate error test of the 2SLS is significant for the error term of the OLS, in the case, the two errors terms would be correlated and we would conclude that the variable tested is endogenous (Wooldridge, 2009).

3.2 Instrumental variables (IV)

When we are suspicious of an endogeneity problem for a specific variable, OLS is a biased estimator and another estimation technique has to be implemented. There are actually multiple causes of endogeneity in a regression model: omitting a relevant variable, measurement error in the dependent variable or in the explanatory variable and including an (by definition) endogenous variable in the set of explanatory variables. The method of IV is used to estimate the causal effect of the explanatory variables when an explanatory variable is correlated with the error term. A valid instrument is needed to implement this methodology. One procedure to implement IV is 2SLS method.

Let's consider again our model specification in equation 1:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \cdots + \beta_K x_{Ki} + u_i$$

Assume that x_{1i} is the endogenous explanatory variable. Consider that z_{1i} and z_{2i} are good instruments for x_{1i}. The 2LSL method should be implemented as follows. In the first stage of the procedure, we have to specify the model of the endogenous variable x_{1i} against all the instruments and exogeneous variables as follows:

$$x_{1i} = \gamma_0 + \gamma_1 z_{1i} + \gamma_2 z_{2i} + \gamma_3 x_{2i} + \cdots + \gamma_{K+1} x_{Ki} + \varepsilon_i$$

Estimate this equation by OLS, obtain the estimation coefficients and generate predicted observation of $\widehat{x_{1i}}$

$$\widehat{x_{1i}} = \widehat{\gamma_0} + \widehat{\gamma_1} z_{1i} + \widehat{\gamma_2} z_{2i} + \widehat{\gamma_3} x_{2i} + \cdots + \widehat{\gamma_{K+1}} x_{Ki}$$

In the second stage, we substitute in equation 1 x_{1i} by their predicted values in the first stage and estimate by OLS the following model:

$$y_i = \beta_0 + \beta_1 \widehat{x_{1i}} + \cdots + \beta_k x_{ki} + \cdots + \beta_K x_{Ki} + u_i^*$$

The instruments z_{1i} and z_{2i} should validate some assumptions in order to be good instruments and to obtain unbiased estimates in the estimation procedure. There are three assumptions that a good instrument must satisfy: i) The exclusion restriction, ii) the relevance condition and iii) exogeneity assumption. The exclusion restriction indicates that the instrument cannot have a direct causal effect on the dependent variable. Moreover, the relevance condition assumes that the instrument does have a direct causal effect on the variable that we are instrumentalizing. Finally, the exogeneity assumption considers that the instruments must be random variables.

So, to sum up, our two instruments z_{1i} and z_{2i} must satisfy:

- cov($x_{1i}, z_{1i}) \neq 0$ and cov($x_{1i}, z_{2i}) \neq 0$
- cov($y_{1i}, z_{1i}) = 0$ and cov($y_{1i}, z_{2i}) = 0$

When we are estimating by IV, it is very important to check if we have chosen good instruments. For that we should implement the overidentification restrictions test in order to check the exogeneity. If we have more than one instrument, it is important to test whether some of them are uncorrelated with the structural error. However, if we have just one instrument, there is not an overidentifying restriction and this test cannot be run. The overidentification test consists of :

- Estimate the structural equation by 2SLS and obtain 2SLS residuals, $\widehat{u_i}$.
- Regress $\widehat{u_i}$ on all exogenous variables in order to obtain the R-squared, named R_i^2
- Under the null hypothesis that all instrumental variables are uncorrelated with u_i, $nR_i^2 \sim \chi_q^2$, where q is the number of instrumental variables from outside the model minus the total of endogenous explanatory variables and n is the number of observations. If we reject Ho at the $\propto$ significant level, we conclude that at least some of the instrumental variables are not exogenous at the $\propto$ significant level (Mariel, 2021).

In addition of that, in order to check the goodness-of-fit of our instruments, which means to check if our chosen instruments are good and appropriate for the endogenous variable that we are trying to explain, we should calculate some statistics.

The goodness-of-fit refers to statistics that determines how good the instruments are and a goodness-of-fit measure is a statistic that summarizes how well a set of explanatory variables explains a dependent or response variable (Wooldridge, 2009). In our case, in order to check the goodness-of-fit, we analyze the Adjusted R-Squared which is an alternative of the usual R-squared for measuring goodness-of-fit which penalizes additional explanatory variables by using a degree of freedom adjustment in estimating the error variance (Wooldridge, 2009). Moreover, we also analyze the the F-statistic obtained in the during the F-test goodness-of-fit which indicates whether the linear regression model provides a good fit to the data.

4. Data and model specification

4.1 Database

The dataset used to make the analysis of the price elasticity is from The Spanish Survey of Family Budget 2019. This survey is one of the oldest in Spain and it is carried out by the National Statistics institute[1] (INE). The main purpose of this survey is to provide information about the expense of households in many different types of ways, that is why we can say that it is a multi-objective survey. The results are published annually and it gathers the answers of nearly 24,000 families randomly selected that cooperate for two consecutive weeks in each of the two years they will be part of the survey.

Those families are questioned about all parts of their life, their expenditure of housing (water, electricity), food but also health and at the same time we have all the essential information that can

[1] https://www.ine.es/en/

help us to make the analysis: number of people in the household, wage, location as well as the education level etc. Each expenditure is named with a five digits code which is COICOP, it is very helpful to know in which category and sub-category part of the given expend is. Another important point to understand how to handle the expense of energy, each observation represents the expense for a set of families in Spain, and the variable “FACTOR” indicates the weight of a specific household for each observation. So, if we need to find how much a specific household spent on electricity, we should divide the given observation by the weight which is its corresponding value of “FACTOR”.

During this whole analysis, the dataset used is the one from the 2019 survey which has been manipulated to have only what we need. Indeed, at the very beginning of the research we had access to three datasets that describe each member of the household, the household in general and the expense of the household. One of the first things which was done, was merging the data of the household with the data of the expense of the household for electricity in order to regress the quantity of electricity (Kwh) used during the year 2019 on many other variables (number of people in the household, income, location, …). In this final dataset we have 20,817 households and for each one we have the quantity, the expense of electricity they paid in 2019 but also all the information to characterize the household.

Figures 2, 3 and 4 give us a clear view of the distribution of our variables. We should notice that the average price paid for 1 Kwh is 0.30 euros and the standard deviation is 0.11. Moreover the average quantity of Kwh consumed during the year of 2019 by the households is 2787 with a standard deviation of 2000. Finally, the average income is 2205 euros with a standard deviation of 1387. All those numbers are in line with official statistics (Eurostat, 2020).

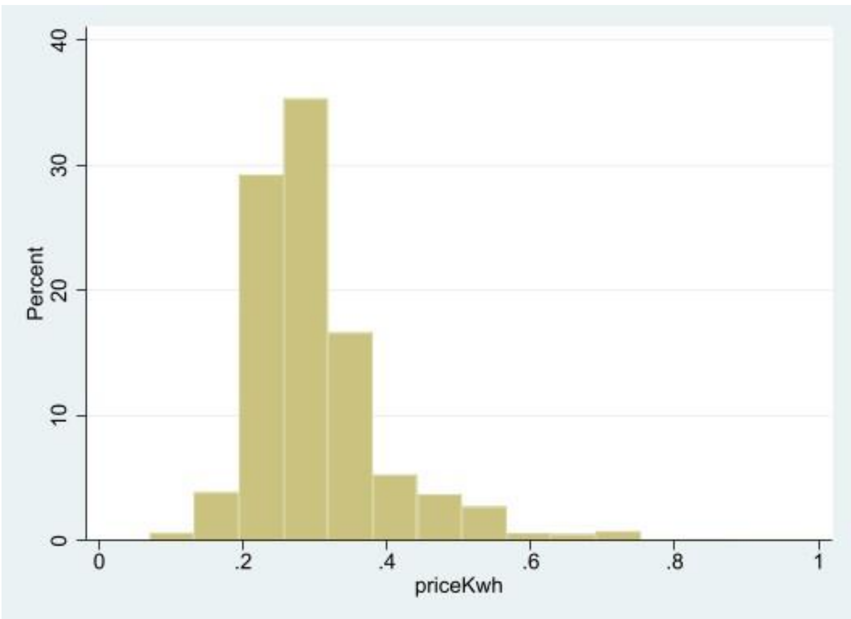

Figure 1 – Graph of the distribution of the average price paid during 2019 for 1 Kwh in percent

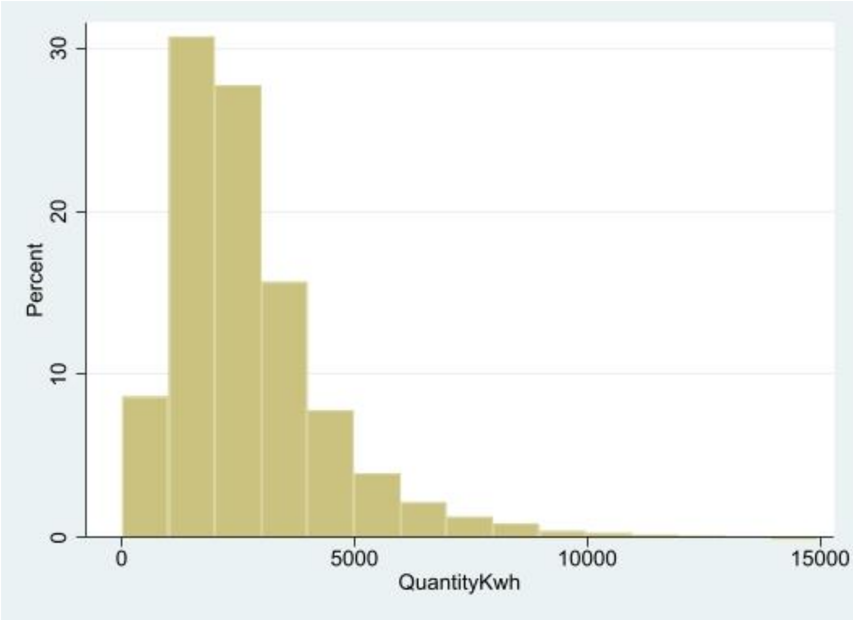

Figure 2 – Graph of the distribution of the quantity consumed in Kwh by the households in percent

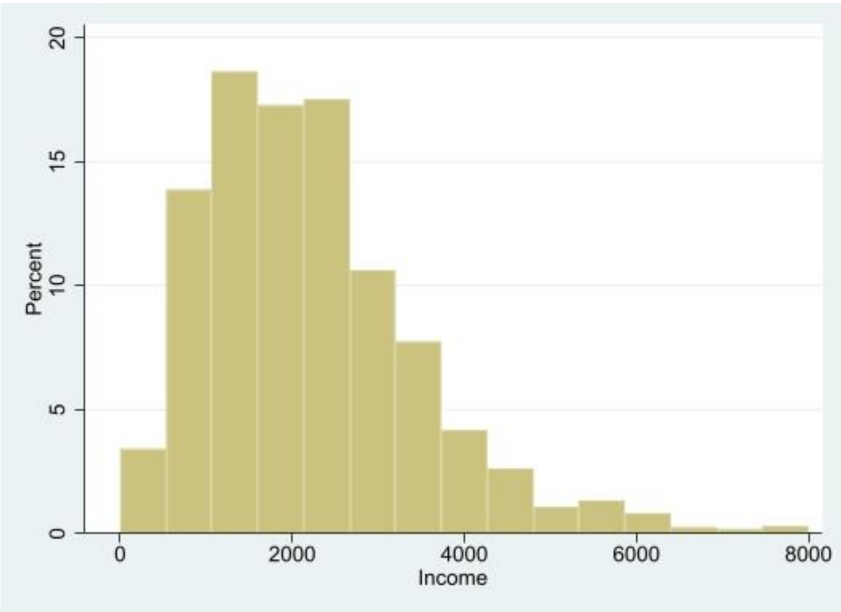

Figure 3 – Graph of the distribution of the income of the households in percent

The final independent variables selected can be seen in table 2. The explanatory variables selected can be divided into different groups, the first one is relative to the household itself: the household variables (i), then, we have also the house characteristic (ii) variables which describe the house where the members of the household live. Furthermore, we have some geographical variables (iii), which give us some information about the location of the dwelling. Finally, the socio-demographic variables (iv) contain information of the characterization of the population, they are relative to the education, income and all the things that can define our household inside the society.

Table 2: Selected independent variables and the dependent variable.

Variable	**Codification**	**Description**	**Obs.**	**Mean (Std. Dev.)**	**Range [Min Max]**
Dependent variable *LnQuantityKWh*	Continue	The logarithm of the quantity of electricity consumed within one year by a household (Kwh)	19,234	7.74 (0.623)	[2.48 10.91]
Independent variables					
lnpriceKw	Continue	The logarithm of the average price the household have paid for 1 Kwh during this year (€)	19,234	-1.24 (0.016)	[-2.65 1.55]
Household characteristics					
(i) numberofperson	Categorical	Number of members living in the household (=6 if 6 or more)	20,817	2.59 (1.200)	[1 6]
(i) lower18	Categorical	Whether one member or more of the household is 18 years old or less	20,817	0.48 (0.824)	[1 6]
(i) numberofTimeat	Quantitative	Number of times the household has eaten at home during the year (Lunch and dinners)	20,817	64.3 (33.53)	[0 337]
House characteristics					
(ii) hugeHouse	Dummy	Whether the main dwelling of the household is a chalet or a big house	20,817	0.13	[0 1]
(ii) medHouse	Dummy	Whether the main dwelling of the household is as medium house	20,817	0.81	[0 1]
(ii) ElecWater	Dummy	Whether the main dwelling is using electricity to boil the water	20,817	0.25	[0 1]
(ii) ElecHeater	Dummy	Whether the main dwelling is using electricity to heat the house	20,817	0.14	[0 1]
(ii) NumberRooms	Categorical	Number of rooms in the main dwelling	20,815	5.12 (1.183)	[1 8]
(ii) recentBuilding	Dummy	Whether the main dwelling was built less than 25 years.	20,817	0.32	[0 1]
(ii) hugecity	Dummy	Whether the city where the main dwelling of the household has more than 50 000 inhabitants	20,817	0.501	[0 1]
(ii) morehouse	Dummy	Whether the household has more than 1 house	20,817	0.16	[0 1]
Geographical variables					
(iii) Northeast	Dummy	Whether the main dwelling of the household is located in the north east of Spain	20,817	0.21	[0 1]
(iii) Northwest	Dummy	Whether the main dwelling of the household is located in the north west of Spain	20,817	0.14	[0 1]
(iii) Canary	Dummy	Whether the main dwelling of the household is located in the canary islands in Spain	20,817	0.04	[0 1]

Variable	Codification	Description	Obs.	Mean (Std. Dev.)	Range [Min Max]
(iii) central	Dummy	Whether the main dwelling of the household is located in the center of Spain	20,817	0.16	[0 1]
(iii) East	Dummy	Whether the main dwelling of the household is located in the east of Spain	20,817	0.19	[0 1]
(iii) South	Dummy	Whether the main dwelling of the household is located in the South of Spain	20,817	0.17	[0 1]
(iii) Rural	Dummy	Whether the main dwelling of the household s located in a rural area	20,817	0.17	[0 1]
Socio-demographic variables					
(iv) lnIncome	Continue	The logarithm of the total income of the household	20,703	7.52 (0.61)	[4.6 9.71]
(iv) lnIncomeSq	Continue	The squared logarithm of the total income of the household	20,703	57.04 (9.09)	[21.2 94.3]
(iv) HighEducation	Dummy	Whether the head of the household has university studies	20,817	0.24	[0 1]
(iv) working	Dummy	Whether the head of the household works	20,817	0.63	[0 1]

Source: own work

Our key variable is the price, more specifically the logarithm of the price *lnpriceKw,* which is here, the average price paid by the household during the year indeed the information given by our dataset are the expenditure in electricity and the quantity of Kwh spent during the year.

With respect to the household variables, we consider the variable *Numberofpeople*, which is the number of people living in the household. Moreover, this group considers as well *lower18*, which is the number of people who are younger than 18 and finally, the last variable of this group is *numberoftimeat* which describes the number of times the family has eaten at home during the whole year.

Regarding the house characteristic variables, *HugeHouse* is referring to a house or big flat in good condition, and *MedHouse* to a regular condition or average flat in good condition. Furthermore, *Elecwater* and *ElecHeater* are referring to houses which are using electricity to boil water and using electricity to heat the place. On top of that, the *NumberRooms* corresponds to the total number of rooms including kitchen, private rooms, living room etc. A building is considered recent if it has been built within the past 25 years and in this case, *recentBuilding* takes the value 1. Finally, *Morehouse* is referring to families with more than 1 house.

Concerning the dummy variables about the location, we took the community of Madrid as reference and the regions are allocated as follows: *Northwest*: Galicia, Asturias and Cantabria, *Northeast*: País Vasco, Comunidad Foral de Navarra, La Rioja and Aragón, *Central*: Castilla y León, Castilla-La Mancha and Extremadura, *East*: Cataluña, Comunitat Valenciana and Illes Balears, *South*: Andalucía, Región de Murcia, Ceuta and Melilla. In connection with those, *rural* considers the households living in rural areas.

For the last group, the first important socio-demographic variable is the total net income as the income level of the household which means the total income of each member of it, and then we calculate the logarithm *lnIncome* and the square of this same variable *lnIncomeSq*. Moreover, *Higheducation* is referring to a level of study of the breadwinner, it has to be more than 240 ECTS which is the equivalent of a master, Phd or every similar long study to take the value 1. Regarding once again the head of the household who is the person who is making the most money in the household, the dummy variable *working* takes the value 1 if he/she is working at least one hour (employee, apprentice, worker receiving training from a remunerated public program, employer, entrepreneur with no employees, independent worker or family assistance), and 0 if he/she is unemployed, retired, student, dedicated to household chores, having permanent work-related disability, or any other type of economic inactivity.

4.2 Model specifications

The specification of the appropriate functional form is a key issue when estimating a demand econometric model. We could expect that price and income do not linearly affect the electricity consumption level. The logarithmic functional form is increasingly used in recent years in order to specify the demand of different goods (Wooldridge, 2009). In this work, we consider the semi-logarithmic functional form of the demand function, as it allows us to have more flexibility on the specification and interpretation of the model.

$$
\begin{aligned}
Ln(QuantityElect_i) = {} & \beta_0 + \beta_1 ln(price_i) + \beta_2 ln(income_i) + \beta_3 \ln(income_i)^2 + \\
& \beta_4 HugeCity_i + \beta_5 Numberofperson_i + \beta_6 lower18_i + \beta_7 hugeHouse_i + \beta_8 MedHouse_i + \\
& \beta_9 ElecWater_i + \beta_{10} ElecHeater_i + \beta_{11} NumberRooms_i + \beta_{12} recentBuilding_i + \\
& \beta_{13} Northest_i + \beta_{14} Northwest_i + \beta_{15} Canary_i + \beta_{16} Central_i + \beta_{17} East_i + \\
& \beta_{18} South_i + \beta_{19} HighEducation_i + \beta_{20} working_i + \beta_{21} numberofTimeat_i + \\
& \beta_{22} Rural_i + \beta_{23} morehouse_i + u_i
\end{aligned}
$$

where β_0 is a constant, β_1,….., β_{23} are the coefficients to estimate, u is the error term and $ln(price_i)$,…, $morehouse_i$ are the explanatory variables defined in section 4.1 and table 2.

Price may be an endogenous variable in the demand model according to the literature (Bueno et al., 2020). Indeed, not only the electricity demand is determined by the prices but the prices are also determined by the quantity of the electricity that the households demand (Ericson & Rafatnia, 2011). Thus, we can have an endogeneity problem. Therefore, OLS reports biased and inconsistent estimates (see section 3). In light of this, we can use IV estimation procedure which is consistent under endogeneity. IV are frequently used when there is an endogenous variable. In this case we should find instrument variables for the price which means variables related to the price but not to the demand.

Since we could suspect that the price could be an endogenous variable, we have to look for one or more instrumental variables which are holding the conditions explained in the corresponding part (see section 3.2). The two instrumental variables are: (i) the log of the price of the gas and (ii) a dummy variable that is describing whether the city where the household is located is densely populated area (this variable is under revision by Eurostat: cities where at least 50% of the population lives in high-density clusters). Those two instruments are expected to be correlated with the price of the electricity but not with the error term and so not with the log of the quantity of electricity consumed by the households.

Table 3: Instrumental variables

Instrumental Variable	**Codification**	**Description**	**Obs.**	**Mean (Std. Dev.)**	**Range [Min Max]**
lnpriceGas	Continue	The logarithm of the average price the household have paid for 1 m^3 during this year (€)	8,064	0.0025 (0.38)	[-3.18 2.69]
dense	Dummy	whether the city where the household is located is densely populated area	20,817	0.47	[0 1]

Source: own work

Let's justify why those instruments could fit as good instruments, that is, they have a direct effect only on the price of electricity but not on the quantity. In section 5, some tests (see section 3.2) will be run to justify what we are trying to show now and have the adequate conclusions.

First, regarding the first instrument, which is the logarithm of the average gas price paid by each household in 2019, it is pretty straightforward that the price of gas has no effect on the quantity of electricity spend. The only way it can have an effect would be that the price of gas is way too expensive for the household and in this specific case people would start changing their installation in order to consume only electricity. But obviously, the change of the price of gas are not enough to affect the change of heater, boiler, kitchen installation for most of the households in Spain. So, it seems that the price of gas has no direct effect on the quantity of electricity

. Furthermore, in a marginalist pricing system such as the one used in Europe's electricity markets, and specifically in Spain, it is the most costly and inefficient generation that ultimately determines the price paid to all sources of generation (Uribe et al., 2022). This means that, to the extent that thermal power plants pass on the extra cost resulting from higher natural gas prices to the price of the power generated by them, the electricity prices paid by households will be directly affected by the price of natural gas. Furthermore, we can find other papers trying to explain the price of the electricity with the price of the gas (Shiri, 2012) and even some papers use it as an instrument to explain the price of electricity as we do (Romero-Jordan et al., 2014)

Concerning the second instrumental variable, *dense*, which indicates whether the city where the household lives is densely populated areas or not, does not directly affect the quantity of electricity spent. Iindeed, the quantity depends on what the household needs inside the dwelling and its specific needs. Furthermore, the price of electricity that we are talking about is not the one directly defined by the market of electricity of Spain, Operador del Mercado Ibérico de Energía (OMIE) which is for each hour the same for all cities of Spain, but the real price paid by the household. Indeed, some autonomous communities give help to the population or some don't, some autonomous communities have more regional taxes on the energy etc.

Let's have a look at the price paid by the autonomous community, which of course contains more or less densely populated areas. Let's try to find out a link between those two variables with some well-known statistics. For that we are using figure 4 which reports the price of electricity in 2018 for each autonomous community in Spain, and figure 5 which is the map of the different density areas in Spain as we have defined it.

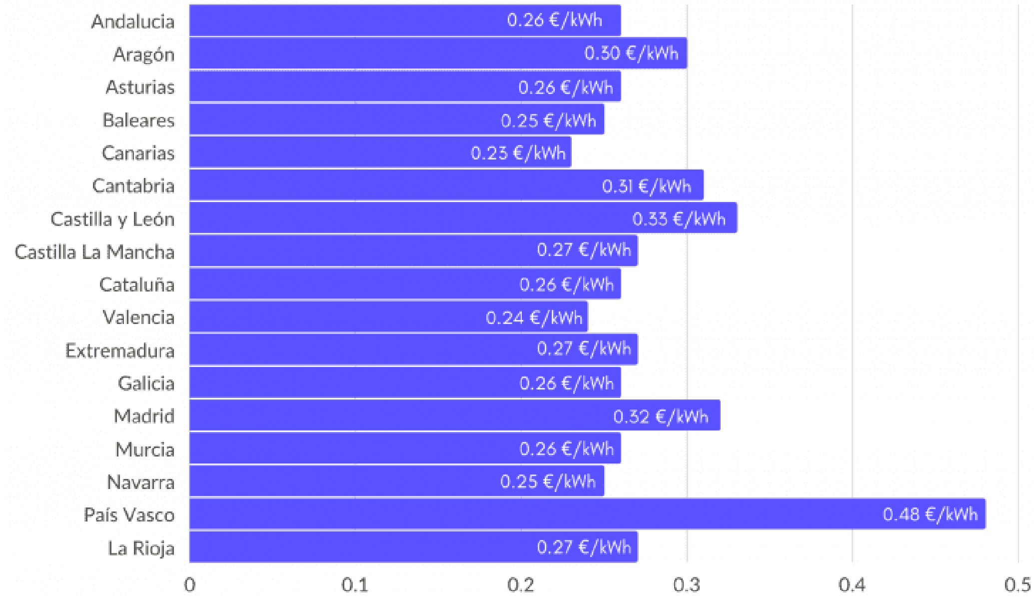

Figure 4 – price of the electricity in 2018 for each autonomous community in Spain (source : Papernest, link: https://www.companias-de-luz.com/por-que-el-recibo-de-la-luz-es-mas-alto-en-tu-ciudad/)

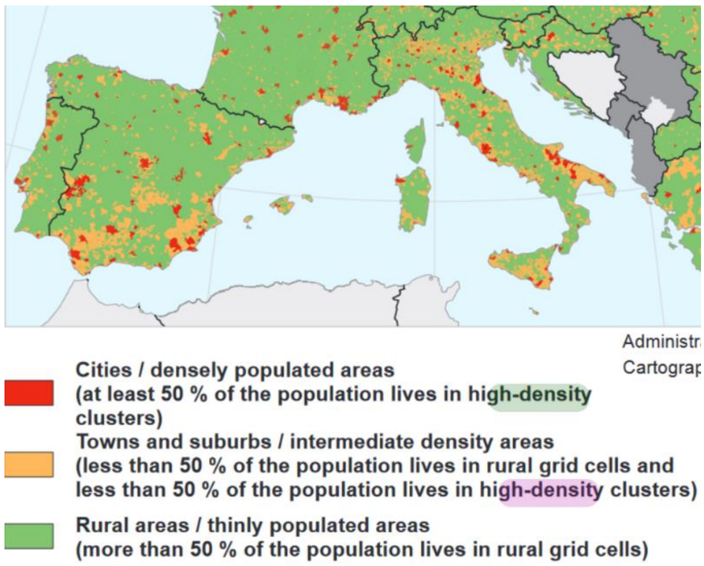

Figure 5– Map of the different density area in Spain (source : Eurostat, JRC and European Commission Directorate-General for Regional Policy, link : https://ec.europa.eu/eurostat/cache/digpub/regions/#total-population)

It seems that when we juxtapose the information of both figure 4 and figure 5, there is a positive correlation between the price of electricity by kWh and number of densely populated areas in each autonomous community. We can find some papers which also take this approach of density to explain the price, for example in Italy where the density has a clear impact on the price of electricity in each region (Folloni & Caldera, 2001). Consequently, for each region of each country, having electricity is a different challenge, and it seems to be also the case for Spain.

5. Results and discussion

The results of estimations of the demand model that identify the effects that the explanatory variables have on the demand of electricity are reported in Table 4. The second and third columns of the table report the OLS and 2SLS estimates, respectively.

First, we center on the analysis of the instruments used in the IV estimation. We check the goodness of the two instruments that we have used. The overidentification test, also known as J-test, allows us to conclude whether the instruments used in the analysis satisfy the necessary exogeneity condition. The null hypothesis is that the instruments are exogenous and the 0 .539461 (p = 0.4627) obtained in the test allows us to conclude that this necessary condition is satisfied and the two instruments are good instruments. Secondly, we should check if our instruments explained well the price of electricity. For that purpose, we can analyze the value of the simple R-squared obtained in the first stage of the 2SLS procedure. Higher values purportedly indicate stronger instruments, and instrumental-variables estimators exhibit less bias when the instruments are strongly correlated with the endogenous variable. In our case, we have a value of R-squared of 0.1495 for the first stage of 2SLS. We also can notice a F-statistic of 0, which is another way to emphasize the fact that our explanatory variables have an impact on the endogenous variable *lnpriceKwh*.

Table 4: Results of the (OLS) and (IV) method for the estimation of the econometric model specified in Eq. (1)

LnQuantityKWh	OLS	(IV) 2SLS
lnpriceKw	-1.05***	-.57**
	(.01)	(.23)
hugecity	-.07***	-.08***
	(.01)	(.02)
numberofperson	.06***	.06***
	(.01)	(.01)
lower18	-.03**	-.02
	(.01)	(.02)
lnIncome	.41***	.29
	(.13)	(.22)
lnIncomeSq	-.02*	-.01
	(.01)	(.01)
hugeHouse	.23***	.11*
	(.02)	(.06)
medHouse	.07***	-.00
	(.02)	(.05)
ElecWater	.20***	.05
	(.011)	(.06)
ElecHeater	.19***	.13***
	(.01)	(.02)
NumberRooms	.04***	.06
	(.01)	(.06)
recentBuilding	.06***	.05***
	(.01)	(.01)
Northeast	-.07***	-.17***
	(.01)	(.05)
Northwest	-.09***	-.15***
	(.01)	(.02)

Canary	-.22**	.45
	(.02)	(.67)
central	-.04**	-.13***
	(.01)	(.05)
East	.01	.03
	(.01)	(.02)
South	.05***	.11**
	(.01)	(.44)
HighEducation	-.03***	-.01
	(.01)	(.02)
working	-.03***	-.06***
	(.01)	(.02)
numberofTimeat	.002***	.002***
	(0)	(0)
Rural	.01	-.05
	(.01)	(.03)
morehouse	.01	.01
	(01)	(.02)
Constant	3.78***	4.98***
	(.49)	(.94)
Observations	19.136	7814
R-squared	0.420	0.346

Robust standard errors in parentheses
*** p<.01, ** p<.05, * p<.1

After analyzing the validity of the instruments, we can focus on the analysis and the comparison between OLS and (IV). The Wooldridge's endogeneity test obtained after the 2SLS allows us to conclude that, if the test statistic is significant, then the variable being tested must be treated as endogenous. Taking into account the test value F(1,7789) = 4.45498 (p = 0.0348), we should reject the null hypothesis that the logarithm of the price of electricity is exogenous at 5%, so we must continue to treat *lnpriceKwh* as endogenous. Thus, OLS estimator gives us inconsistent estimates of the electricity demand model.

Therefore, we focus our analysis on the results obtained with the (IV) estimation method. In general, we can see that most of the variables are significant. Furthermore, the R-squared is high enough (0.35) to tell that those explanatory variables insightfully describe the quantity of electricity.

Another very important thing, and the main focus of this research, is the estimated price elasticity of the electricity demand, which is equal to -0.57. So, the demand seems to be price inelastic and the price have little effect on electricity consumption of Spanish households in 2019. Indeed, this value is aligned and got credit from the vast majority of the elasticity range given in the literature: between -0.88 and -0.07 (Labandeira et al., 2017), between -0.75 and -0.003 (Ericson & Rafatnia, 2011) and between -0.05 and -0.7 (Andruszkiewicz et al., 2019).

As we have seen before, the reference region when analyzing the effect of the location of the household is the region of Madrid. *East* and *Canaries* are the only dummy variables that are not significant for the regions, which means that our dataset does not have enough evidence of a difference behavior of the families living in those regions with respect to the region of Madrid.

For the rest, it is obtained that the families that live in the north and the center consume less than those who live in Madrid, *Northeast* (-0.07), *Northwest* (-0.09), *central* (-0.04) but a household in the *south* (0.05) consumes more electricity. We will not go too deep into the details for the case of the region of Canary islands because it is a very unique location and the number of observations is pretty low. These results coincide with the results obtained by (Morollón & Díaz, 2020) about the average annual electricity consumption of Spanish families in a municipal scale, as it is summarized in Figure 6

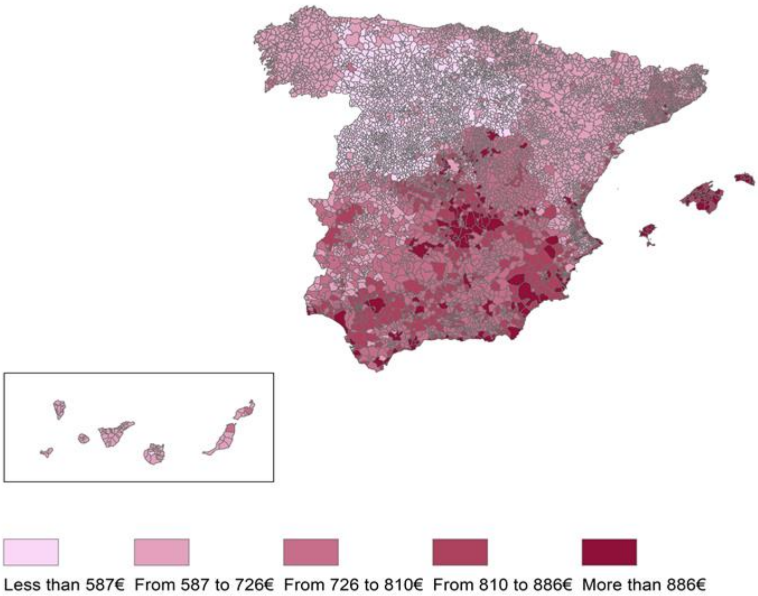

Figure 6 - Map with the average annual electricity consumption of Spanish families (2020), municipal scale

Figure 6 shows that the south region consumes more electricity on average than Madrid, and Madrid also consumes more than the north. It can be explained that the fact that 36.8% of the households in Spain use the gas to heat their home, but in the south of Spain the main issue is having air conditioning (AC) in order to cool down the temperature at home. The AC can exclusively use electricity in order to work and it consumes a lot of electricity, that is why we are observing this difference between the north and the south of Spain.

Let's now focus on some dummy variables that are highlighting some results that have to be explained more in detail. Living in a huge city, *hugecity*, has a significant and negative effect (-0.08). This can be explained by the fact that people are most likely to have hobbies outside home and spend their time at other places than their own home. For the residential electricity consumption, the higher the population where the families live is, the lower the electricity consumption (Bisello et al., 2019). So, by extension the huge cities which are also the most populated, are the ones where people are less consuming. A similar conclusion can be given for the negative and significant effect of the dummy variable *working* (-0.06). There is a tendency to spend less electricity at home when a member of the household works because the people are more likely to be out during the whole day.

In addition, let's try to figure out a disturbing result, living in a recent building, *recentBuilding*, has a positive effect on the consumption of electricity (0.05). It is a confusing result since we would expect that the most recent buildings are more energy efficient than the older ones, because beforehand, this question was not an issue for the engineers and the technological progress was very limited. But we have to keep in mind that the definition of "recent building" identifies the houses which are less than 25 years old and the energy efficiency is more recent than that. Furthermore, in Spain, according to our database, the new buildings are increasingly using electricity to heat the houses instead of gas, indeed the recent buildings use 4% more electricity to heat the houses.

Finally, the positive and significant effect of the following variables are straightforward to understand the electricity demand: number of people living in the dwelling *Numberofperson* (0.06), whether people are living in a huge house *HugeHouse* (0.11), whether people are using electricity to boil their water *ElecHeater* (0.13), and the number of times the family has eaten at home *numberofTimeat* (0.002).

6 . Conclusions and policy recommendations

This master thesis has focused on the influence of the price of electricity on the electricity consumption. The literature shows that the price of energy can slightly affect the energy consumption, but its effect varies depending on the country, the periods of time, the methodology used, but also the profile of consumers.

This study has estimated the Spanish households' electricity demand for 2019. The variation of the quantity of electricity consumed in 2019 by a Spanish household is investigated in regard to the price and some other independent variables such as the location of the dwelling, if the dwelling is in a huge city, the size of the house, the number of people living in the household, whether the breadwinner works or not, whether the dwelling uses electricity to heat the place, whether the dwelling is recent or not, and finally the frequency of the family eating at home.

Two estimators have been used to estimate the electricity demand model, OLS and then (IV) because of the endogeneity problem due to the price. After we made sure that we had a problem of endogeneity with the Wooldridge's endogeneity test and so that the OLS estimates are biased, we have used the instrumental variables regression to have valid results. In order to make sure that our two instruments which are *lnpricegas* and *dense* are good enough, we have run the overidentification test and some statistics like the Adjusted R-Squared and the F-statistic, and we finally conclude that our instruments are good enough to take into consideration the results obtained with the 2SLS.

The results could help us to figure out to what extent a household changes its consumption of electricity depending on the price and how to characterize those households. We could thus reflect on the effectiveness of the recent price policies the Spanish government is carrying out due to the increase in the energy prices.

Our results show that the electricity demand is price inelastic (-0.57). This estimate seems to be in line with the literature as it falls into the range of -0.7 and -0.2, where most of the estimation we have reviewed fall also into.

Moreover, we have seen that households with lot of members living in the same dwelling, located in the south, using electricity to boil water, in a recent building and which has the habit to eat at home often tend to present a higher electricity consumption. On the contrary, the households living in a huge city, in the north or the center of Spain and with a worker as the breadwinner tend to have a lower electricity consumption.

Taking into account our findings, we can say that policies based on price could have little effect on the demand of Spanish households (the reaction of the demand in quite inelastic) and that efforts should be complemented with other type of policies, for example, informational policies and regulatory policies.

More precisely, if we focus on policies that are aimed to reduce the expenditure of households in order to decrease their expenses, we can think for example to inform the population in order to implement habits of making multiple meals at once when it is possible in order to decrease the number of meals cooked at home. Another example could be informational policies such as awareness campaigns about how much an AC can consume, as we have seen that the air conditioning is a huge issue now but will be even more in the following years, so the users should know how to use it well.

Furthermore, if we focus on reducing energy consumption in order to reduce the CO2 emissions, the policies have to be more global, for example, the regulation and control of the use of AC for private and public use and the obligation of some energy efficiency norms for new buildings such as the Energy Efficiency Certificate for buildings.

Nevertheless, this study has focused just on the demand of one year (2019) and future work could analyze more than one year to see the evolution of the demand with respect to the price and other factors. Furthermore, it would be interesting to do this kind of work with the price of the electricity given by OMIE in order to confirm the estimation we have done with other types of datasets.

References

A Bernstein, J. Griffin, (2006), Regional differences in the price-elasticity of demand of energy.

Ali Shiri, (2012), Electricity Price Forecasting Using Support Vector Machines by Considering Oil and Natural Gas Price Impacts.

Amaia de Ayala, Ibon Galarraga, Joseph V. Spadaro, (2016), The price of energy efficiency in the Spanish housing market.

Andriano Bisello, Daniele Vettorato, David Ludlow, Claudia Baranzelli, (2019), Smart and sustainable planning for cities and regions.

Nina Boogen, Souvik Datta, and Massimo Filippini, (2021), Estimating residential electricity demand: New empirical evidence. Energy Policy.

Bose, Ranjan Kumar, Shukla, Megha, (1999), The Elasticities of electricity demand in India.

Cameron Torra, Arnau, The Spanish electricity system, (2020), ¿ Por qué ha aumentado tanto el precio de la energia eléctrica ?

Murat Çetinkaya, Alparslan A. Başaran, N. Bağdadioğlu,(2015), Electricity reform, tariff and household elasticity in Turkey. Utilities Policy.

N. Chindarkar, N. Goyal, (2019), One price doesn't fit all: An examination of heterogeneity in price elasticity of residential electricity in India. Energy Economics.

Cristina Cattaneo,(2019), Internal and external barriers to energy efficiency: which role for policy interventions.

Carol Dahl, (2012), Measuring global gasoline and diesel price and income elasticities.

David Hoyos, Alaitz Artabe, (2016), Regional differences in the Price Elasticity of Residential water Demand in Spain

Desideri Romero-Jordan, Pablo del Rio, Cristina Penasco, (2014), Household electricity demand in Spanish regions. Public policy implications.

Eurostat, (2020), Energy consumption and use by households, https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20200626-1.

Fernando Rubiera Morollón, Alberto Díaz, (2020), Obtención y análisis del mapa municipal de consumo de electricidad de las familias en España, abc, https://abcblogs.abc.es/riqueza-regiones/otros-temas/obtencion-y-analisis-del-mapa-municipal-de-consumo-de-electricidad-de-las-familias-en-espana.html?ref=https%3A%2F%2Fwww.google.com%2F.

Gauthier Lanot, Mattias Vesterberg, (2021), The price elasticity of electricity demand when marginal incentives are very large.

Giuseppe Folloni, Oriana Caldera, (2001), Size, density and costs of network services – the case of the distribution of electricity in Italy.

IDAE, (2021), Estudios, Informes y estadisticas, https://www.idae.es/informacion-y-publicaciones/estudios-informes-y-estadisticas.

IEA, (2019), Global energy & CO2 status report 2019, https://www.iea.org/reports/global-energy-co2-status-report-2019

Ishaque, (2018), Revisiting income and price elasticities of electricity demand in Pakistan. Econ. Res.

K. Ito, (2014), Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. The American Economic Review.

Javier Bueno, Desideri Romero-Jordan, Pablo del Rio, (2020), Analysing the Drivers of Electricity Demand in Spain after the Economic Crisis.

Jerzy Andruszkiewicz , Józef Lorenc, Agnieszka Weychan, (2019), Demand price elasticity of residential electricity consumers with zonal tariff settlement based on their load profiles.

Jonas Ericson, Farshid Rafatnia, (2011), An estimation of price elasticities for supply and demand in Sweden.

Jorge M. Uribe, Stephania Mosquera-Lopez , Oscar J. Arenas, (2022), Assessing the relationship between electricity and natural gas prices in European markets in times of distress.

Chandra Kiran B. Krishnamurthy, B. Kriström, (2015), A cross-country analysis of residential electricity demand in 11 OECD-countries. Resource and Energy Economics.

Raymond Li, Chi-Keung Woo, Kevin Cox,(2021), How price-responsive is residential retail electricity demand in the US?

Marc F. Muller, Sally E. Thompson, Ashok J. Gadgil, (2018), Estimating the price (in)elasticity of off-grid electricity demand.

Maria del Mar Sola, Amaia de Ayala, Ibon Galarraga, Marta Escapa, (2020),Promoting energy efficiency at household level: a literature review.

Niyaz Valitov, Sebastian Nielen, Wemer Bonte, Torben Engelmeyer, (2012), Price elasticity of demand in the European wholesale electricity market.

Paul J. Burke, (2017), The price elasticity of Electricity demand in the United States: A three dimensional analysis.

E. Pellini, (2021), Estimating income and price elasticities of residential electricity demand with Autometrics. Energy Economics.

Petr Mariel, (2021), "Regression Analysis in Practice", "Specification and measurement errors, instrumental variable (IV) estimation" and "Qualitative Dependent variables ", Master in Economics: Empirical Application and Policies in the University of the Basque country.

Peter C. Reiss and Matthew W. White, (2005), Household Electricity Demand, Revisited. The Review of Economic Studies.

Devalina Saha, Rabindra N Bhattacharya, (2018) An analysis of elasticity of electricity demand in West Bengal, India: Some policy lessons learnt. Energy Policy.

N. Wang, G. Mogi, (2017) Industrial and residential electricity demand dynamics in Japan: How did price and income elasticities evolve from 1989 to 2014? Energy Policy.

Jeffrey M. Wooldridge, (2009), Introductory Econometrics: A Modern Approach, 4th ed., South-Western College Publishing, New York.

Xavier Labandeira, José M. Labeaga, Pedro Linares, Xiral Lopez-Otero, (2017), The impact of energy efficiency policies: Meta-analysis.

H. Yu, X. Xin, (2020), Demand elasticity, ramsey index and cross-subsidy scale estimation for electricity price in China. Sustainable Production and Consumption.

Zhengyuan Zhai, Lei Zhang, Xiaochao Hou, (2020), Price elasticity of electricity demand in China.