Energy conservation and cost estimation analysis in Spanish households in regards to the Energy Performance Certificate rating.

Title: Energy conservation and cost estimation analysis in Spanish households in regards to the EPC rating.

Theme: Master's Thesis

Project period, 4th semester SEPM:

AALBORG UNIVERSITY
STUDENT REPORT

Supervisor: Jakob Zinck Thellufsen

Eduardo Lezama

Copies: 1

Pages, total: 42

Appendix:

Supplements: 1 CD – includes model and calculations

Following the directive of the EU Energy Performance in Buildings directive into the Spanish law, all properties, new and existent are obliged to have an energy performance certificate. This thesis analyzes the effects of energy efficiency rating on the sale prices of dwells in Spain. Using the Rosen's hedonic model in a semi-log approach, this thesis focus in modeling the impact of having an EPC certificate on the final property values. Two models are conducted in order to set a minimum and a maximum spectrum of variation of the premium prices paid for houses with a higher energy rating. The results of such models shows that there is a positive effect on the transaction price of properties when the energy efficiency rating is higher. This study also shows the impact of having dummy variables in the model and what other characteristics affects positively or negatively the value of the property.

(This page is left blank intentionally)

Preface

This report is composed trough 4th semester of the master programme Sustainable Energy Planning & Management at Aalborg University as the final master thesis of the program during the period from 2nd of February 2018 to 2nd of June 2017. The project is the culmination of the education program including all the knowledge gathered in the previous semester of the master and implementing them in this final master thesis.

Acknowledgements

I would like to thank the following people, who helped me doing this project by providing information and/or contributed with a valuable insights:

- Maria Luisa Jimenez, Ayuntamiento de Barcelona
- Xavier Marin, Schneider Electric
- José Cantero Muñoz, Ayuntamiento de Madrid
- Marc Gonzalez Tejero, Instituto Nacional de Estadística - Fomento

I would also like to express my gratitude towards my supervisor Jakob Zinck Thellufsen for the knowledge and helpful supervision throughout the project period.

Last but not least, I want to thank everyone from Aalborg University, specially Poul Alberg Østergaard for the dedication and support throughout all the period that I was not able to do my thesis for personal reasons. Also thank all the AAU staff that participated in my education.

Acronyms

CTE	Código Técnico de la Edificación (Technical Building Regulation)
EE	Energy Efficiency
EPC	Energy Performance Certificate
EU	European Union
GHG	Greenhouse Gas
IDAE	Instituto para la Diversificación y Ahorro Energético
IEA	International Energy Agency
LSM	Least Squared Method
RD	Royal Decree
RITE	Reglamento Instalaciones Térmicas en Edificios (Building's Thermal Installation Regulation)
SSE	Square Sum Error
SST	Square Sum Total
Toe	Tonnes of equivalent oil

CONTENTS

1 INTRODUCTION

One of the main reason of global greenhouse gas (GHG) emissions is the high production and high consumption of energy. Within the current context of climate change and security, almost all developed countries are attempting to decrease fossil-fuel use in all the different economic and industrial sectors, such as transportation, industry and buildings. In comparison to the baseline scenarios in where the surface temperature was going to increment 2ºC by the end of the 21st century, global measures regarding energy efficiency (EE) in the sectors aforementioned are expected to grow by a 336$ billion annually over the next two decades 2010-2029 (IPCC, 2014)

Buildings and household sector is one of the biggest contributor of GHG emissions. According to (IPCC, 2014) when the emissions coming from electricity and heat production are qualified to the sectors that use the final energy, industry accounts for a total of 32% of the global greenhouse gasses emissions, followed closely by agriculture and others land uses, with an amount of 24,8% ,buildings and households with 18,4%, while 14,3% and 11% are from transportation and other energy usage respectively. Along the whole European continent, residential buildings and other commercial buildings accounts approximately 40% of the total energy consumption, and 36% of CO_2 emissions. If the focus is on Spain, households and other residential buildings represent about 17% of the final total energy consumption, whereas the contribution from tertiary buildings contribute with a 9% (IDAE, 2011). To sum this data up, 65% of the building energy use is attributed to the residential sector, whereas 35% to non-residential buildings (offices and commercial buildings, not including the industrial ones) (Chuchi & Sweatman, 2012). The energy that is directly used in households represents one fifth of the total GHG emissions in Spain. If the emissions from the construction process are also considered, the residential sector contributes to a third of the total GHG emissions. (WWF, 2012)

Nowadays buildings have become a crucial field for the environment and energy policies due to their important impact on the demand and the emissions they produced. Indeed, commercial and residential buildings accounts for almost 40% of the final energy consumptions in countries with a high level of industrialization. Unlike other important sectors, the effects of buildings is of particular importance given the fact that many of them were built under old (or sometimes without) codes and therefor often without any attention to the energy efficiency or other environmental issues. Buildings are a long –term prolonged good; hence, their participation to the future energy consumption and emission are likely to be one of the largest in any energy paradigm of the future unless new and specific actions and policies are introduced within this matter. (Gago et al., 2013)

Energy efficiency (EE) gives an opportunity to change the current trend by applying measures based in It's cost-effectiveness to reduce the energy consumption (Levine et al., 2007). International institutions as the International Energy Agency (IEA) have remarked during the last year the importance of the potential savings In energy achievable from building design, and have emphasized on governments to introduce new policies to promote the energy efficiency and savings in this sector (IEA , 2007) like for instance: standards and codes, subsidies and taxes (Markandya et al., 2015). Some of the measures mentioned in that study consist in the improvement of the technical conditions of the buildings in order that they require less energy to operate and provide the same service, for instance, trough insulation, improving heating and cooling system efficiency or the use of better and improved energy appliances. In addition, users can reduce their consumption of energy by adopting a habit which satisfy the energy-saving mindset such as switching lights off, change the premises in interior temperature during winter and summer and so.

Even though energy efficacy adoptions have been widespread in other fields, it also has ben hardly observed in this sector given some strong factors that avoid agents from talking and aware the advantages of energy efficiency policies and their potential (Levine et al., 2007). Some of these important barriers are information and means imperfection and behavioral failures due to the fact that this is a complex field with a great number of agents and high cost with a limited access to capital ((Jaffe & Stavins, 2010). However, these effects may be less noted for households that show a more environmental concern, but still face other problems such as a bounded rationality toward the change and/or long payback periods. A new mindset related with houses that are pro-environmental might add value not only to the monetary savings coming from reducing the energy use but also the improvement associated with the environment; therefor they may benefit from EE more than the rest of the households. This can lead to increase the awareness when buying new appliances, by reducing the payback periods or just to pay more attention to energy performance attributes of appliances (energy performance certificates and labeling systems). Moreover, individuals of a building are more keen and aware to save energy at home by implementing eco-friendly habits such as regulating more efficient the thermostat even if they not pay the bill directly (IEA , 2007).

Being consistent with the EU's targets for climate and energy (European Comission, 2017): 20% improvement in energy efficiency and 20% greenhouse gasses emissions reduction by 2020; 27% energy efficiency goal and 40% emission reduction by 2030, the building sector have a large latent potential for cost-effective energy savings achievable by EE measures. The reason behind this, mentioned a few paragraphs before is that the majority of buildings that are constructed are based in old code regulations that didn't have many restrictions in terms of energy efficiency for buildings. This restriction falls under the category of insulation, materials, characteristics of the windows and so. Some of these EE measures that can bring benefit to the sector are informative, while others are educative or mandatory, and some use incentives in the form of money to promote EE behaviours (de la Rue du Can et al., 2011)

The main EU policy instrument to develop and improve the energy performance of buildings is "The Energy Performance of Buildings Directive (EPBD) (Directive 2002/91/EC)", which take into account cost−effectiveness and local conditions and requirements (Bio Intelligence Service, Ronan Lyons & IEEP, 2012). If this directive is properly and fully implemented, energy savings from the use of the EPBD are expected to decrease the final energy demand around 96M tonnes of equivalent oil (Mteo) in 2020.

EPBD makes sure that when a new building is being constructed, rented or sold, a certificate of energy performance (EPC) is made available to the tenant, owner or the potential buyer. This certificate illustrates an energy efficiency rating for the performance of a home, from an scale of "A" to "G", where "A" is the most efficient type of building and "G" is the most inefficient one. The idea is more or less the same as the one applied to the home appliances but, in addition, EPC must include the information on the needs of consumption in terms of energy for the building including reference values, as well as recommendations for improvements that can bring a cost-effective turnover to raise the current building rating. The reformation of the EPBD in 2010 (Directive 2010/31/EU) explained some aspects that weren't clear enough, promoted the paper that the public sector needs to play, and reinforced the role of EPCs when advertising a building for its rental or sale by demanding periodical publications rather than only at the time of signing the rental contract or the purchase contract.

Even though the implementation of the EPBD across Europe is a real matter, it may vary from region to region depending on multiple factors which include: local context and political situation, characteristics of the housing market of the country/region and so. Focusing in Spain, the EPBD 2002/91/EC was implemented trough the Royal Decree 47/2007 (Real Decreto 47/2007), in which it is stated that a certification on energy efficiency is required to be provided by new buildings only. Afterwards, in the Royal Decree 235/2013 (Real Decreto 235/2013) the modifications that the EPBD suffered in 2010 (EPBD 2010/31/EU) were included and extended the EPC scope to all type of buildings, including the ones that already exist. Consequently, from June 2013, all properties owned, offered or promoted to rent or sale in Spain are required to have this energy efficiency certificate. According to the "Real Decreto 235/2013", the methodology to generate the certificate assessment must be carried out but an authorized personal or technician. The person in charge will collect information related to the building's characteristics such as exterior exposure, windows and door openings, year of construction, orientations and so and generate calculations of CO_2 emissions by using a computer software developed by the "Instituto para la Diversificación y Ahorro de la Energia" (IDAE) (Institute for Energy Diversification and Saving) named "CE3X". This software gives an assessment of the situation of the building plus some recommendations for improving the current building rating grade (between A to G). This certification, then, is validated by the competent local administration and a copy is delivered to the owner which needs to be kept for 10 years or until re-assessment.

The principal idea behind this EPC certification system is that clients and consumers can evaluate and comprehend the characteristics of energy consumption of buildings in their decision-making process before they sign a contract. This information that EPC gives can be used as an incentive for builders and owners of buildings to invest in energy efficiency measures, as it can be theorized that the improvement of energy performance of buildings may lead also to a higher price and rents within the market.

Researches have depicted that the effects of energy efficiency rating in building's price is focused on real estate market (Reichardt et al., 2012). Little by little, though, the studies regarding home building are growing in the last years, and they depict a positive relationship between energy efficiency rating on households and home prices and/or rental prices. However, some studies show that the energy efficiency influence is stronger in the sales market (Fuerst et al., 2015). Due to the existing gap in the residential market literature compared to the commercial one and the greater regulation barriers to energy efficiency in homes (Chuchi & Sweatman, 2012), this thesis study will try to focus just on residential and household sector. In Spain, building used for living represent a total of 85% of the build surface, where the 15% remaining is used for other tertiary purposes. (WWF, 2012)

Finally, it's important to mention that Energy Performance Certificate (EPC) has been newly stablished as a must for new and current existing building in Spain, thus the Spanish household market lacks in evidences of this new certification diffusion rate. The idea of this report is to give a glimpse of this situation and analyse the impacts of increasing the energy performance of buildings in the price and make a model that describes future prices based on the energy label variable.

1.1 Studies That Shows a Link Between Energy Labeling and Price and The reason Of Using Hedonic Price Modeling

In 2010 (Sundberg, 2010) discovers that most of the people is willing to pay an extra price for buildings that have a better energy performance which can be seen once the energy performance certification is done. Some of the studies conducted in 2010 and 2011 by (Nanda, A et al., 2011) and (Pivo, 2010) are based in a different way of interpretation of the property value. Nanda, A. base the study in the appraised value instead of the real exchange whereas Pivo, G. thinks that the real price is a mixture of appraised value and transaction value.

Back in time, in USA there are some studies that show the impact of EPC rating and the house value exchange, in which they depict a positive relationship between the value and the EE rating. (Dinan, 1989) (Longstreth, 1986). The same happened years after in Europe where some of the first studies which intended to analyse the outcomes of the EPC in the market price (Brounen, 2011) in the Netherlands. The conclusions of such studies were very significant on this matter because it discovered that in Netherlands the certification rating of A, B and C had a 10%, 5% and 2% extra premium price compared to D-rated homes of the similar features. The method used by him was a hedonic pricing method with a broad number of variables that were used to control all the aspects of dwellings. In Germany, (Cajias, 2013) using a different methodology based on a continues scale of over 2.500 buildings between 2008 and 2010, estimated that an 1% increase in energy savings could produce 0,5% value increase in the price of the house and 0,1% in rent prices.

Outside the scope of Europe there are also some studies that shows the same idea. One of this first papers was done by the (Australian Bureau of Statistics, 2008) used a hedonic price for real estate transaction in 2005-2006 and concluded that the premium price of houses with better EE rating was 2% in 2006. Moreover (Nobe et al., 2011) made a hedonic pricing model using as variables the ENERGY STAR certification and compared the ones that were certified against the one that were not. The results showed that homes with ENERGY STAR certification had a premium price per square meter of approximately $9.

As it can be seen from these study, the standard approach to understand the contribution of certain feature of a property to the transaction value (in specific, the EPC) is some sort of modeling using hedonic regression. Hedonic regression was theoretically formalised as a statistic tool in 1970 by study of Sherwin Rosen. This tool has been broadly used in different fields, from environmental economics to urban, for instance, to estimate the potential comfort and discomfort aspects of transportation for instance in rural villages. Hedonic regression method is widely used to assess the internal and external attributes of a property, the price of which is implicit in the price but not observable. This method offers a broad spectrum and quality and clarity information with robustness, however is much subjected to the changes in the initial data and other statistics aspects that can distortion the model such as collinearity between variables.

1.1.1 What Links Between EPC and Building Price Cost?

As mentioned previously, EPC brings a new whole idea on how to improve the energy efficiency of buildings especially trough the renovation and refurbishment of existing one. When an EPC is issued it must include values of references, like the current legal standard, which provides a way of comparing and evaluate the performance of the building in terms of energy. Moreover, it must bring a series of recommendation for cost-effective upgrades in order to raise its performance and therefor the rating of it. The EPBD directive recast in 2010 reinforced the role of the EPCs by demanding publication of the energy performance of the building at the time of advertising for rental or sale (European Comission, 2010).

Providing information that is reliable and clear at the appropriate time is vital for making energy efficiency measures attractive in terms of investment. EPC is built in order to provide an incentive for owners and future builders for improving the efficiency as it can be thought that an improvement in terms of energy efficiency may lead to a higher price in the housing market and renting market.

In terms of value of a property, there are a few ways to think about it. Regarding the scope of this study the idea of the value of a property is reflected as a value that comes from a transaction, either sales or rental. For this type of value, it is important to have in consideration that the price of the housing market depends on a series of complex social and economic processes in which different type of agents can interfere. The idea and the interest between the relationship that may exist in EPC and price of a building isn't new: for instance, there are two studies from 1980s performed by (Laquatra, 1986) and (Gilmer, 1989) in which it is depicted what can be one of the first links stablished between the energy efficiency in households and the pricing of the dwell in the market. Both of this study refer to the same problem there is now, whether an investment made to improve the EE of a house will carry a positive reflection in the value of the building.

There are many reasons why to be interested on this. One of the reasons is that buildings with a better performance in energy efficiency may cost more to build, even though it doesn't have to be necessarily the case, it is important to understand if this investment made in energy efficiency can be recovered in the future and can lead to a better income for the investor (Brounen, 2011). This same argument may be used for renovation and refurbishment of existing houses. For this reason, the focus of this study is also to know if a buyer is willing to pay an extra price, called "premium" price (an extra of the conventional price) for buildings that have a better performance in energy and if it's the case, this information can encourage investors to invest in increasing the energy performance of the buildings.

People tend to expect that energy performance of houses will impact directly in the value of it because it brings a saving to the owner and also a contribution to the environment. In the same way, the monetary value of a building is affected because there are other type of benefits associated with energy-efficient buildings as for instance a bigger options in term of services that it can provide (Ürge-Vorsatz, 2010). For example, in this latter study it is said that a house is a composition of various items (features), such as number of rooms, bedrooms, size, bathrooms, garden and so. These features are the ones that provide to the consumer services from the house (for instance: showering, cooking, sleep, etc.). For a number of services that a home provides, the energy performance of it affects the running cost. If two different homes provide the same identical services but the running costs associated with them are different, this difference should be reflected in the price that the owner or tenant pays for living in it.

Additionally, most countries nowadays still reside in fossil fuels as their primary energy source which means that increasing the energy performance of the residential sector may lead to a better way of fighting against the climate change. It should be considered that if the society values the environment and take into consideration the impacts of pollutant energy solution, they also will value how energy efficiency is important when planning to rent or buy a house (Ürge-Vorsatz, 2010).

2 Research Question

Based on the information aforementioned, a subjacent research question is formulated, in order to determine the situation this sector is going through and how this new changes and this new label system is affecting and will affect the Spanish household market.

Based on the information presented above, how the EU directives was introduced in Spain and what are the cost benefit in prince of household for increasing the energy rating of a dwell?

The main focus of this research question is to ensure that the necessary information about the residential sector are gathered and analysed with the target of providing an understanding scope of the situation and the development that are done and being done throughout this field. By doing this, it should be possible to concrete on the analysis and how to face the challenges that raise from the future evolution of this sector.

3 METHODOLOGY

The main aim of this thesis is to study whether there is a relationship between energy performance certification of building (expressed by the EPC label) and their monetary value when it is sold or bought. The achievement of this goal is done through an analysis of combined datasets of EPC information and transaction price of the dwells with an EPC certificate. Moreover, the underlying European policy is analyzed as well as how it has been implemented in the Spanish household regulation policies.

The chapters are structured as follow: The Chapter 1 is a literature review of different studies and describes a brief introduction into the energy labeling system across Europe and why is the link between EPC label and prices interesting at all. There is a subchapter where some studies that show the link between prices and EPC are described and moreover Chapter 1.1.1 gives the main findings in what are those relationships. As said, this chapter has a review of three different approaches to the topic as well a real cases studies from USA, Netherlands, Germany and Australia. The way it is done is first: to set up the reader in the matter of GHG, energy efficiency and dwells and at the same time introducing him to the possible benefits that other studies has shown in the matter.

Next, it is outlined what is the question that arise from the previous chapter and make the focus in the Spanish housing market, where this study will be based. This research question is based on the main findings state on Chapter 1 and focus the attention on what exactly will this thesis will try to answer.

Chapter 4 and 5 gives a glimpse into the energy labeling system implementation in Spain and what were the main tools used by the government to implement the EU directive. A brief description of the main policies involved is carried out by reading part of the documentation that support this policy and by investigating how the deployment of the EU directive was carried out in Spain. Chapter 0 a brief description of the current Spanish house environment is described as well as the main findings of the current state of the sector. This analysis was done looking at the main housing indicators in real estate and going through several articles that talked about the housing bubble in Spain and how it affected the whole economy once it blew up.

Chapter 6 is where the core of the project is located. This chapter is divided in three different sub-chapters all of them linked to the idea of explaining, understanding and making of the model. The model itself is based on the theoretical framework of hedonic pricing for dwells in the real estate market. Even though this master thesis is not analyzing the impact of the price in the real estate market, the model can still be utilized to analyse the effects of characteristics that are intangible in the overall price of the house. In the first part of the chapter, the regression analysis theory is explained as well as how to know whether a model based in a regression is good or not. The main indicators, R^2 and adjusted R^2 are explained. Moreover, in this chapter the hedonic modeling is explained and analyzed with some example that depict the idea of how the model works. The second part of this chapter is dedicated to the analysis and explanation of the data used in the modeling. Where and how the data was obtained, how the data was filtered, what items were selected and so, are things that are explained in this chapter to make clear what is the scope of the information that will be used to carry on the models. Finally, the last part is where the model is build and how the variables and predictors of the model are integrated to generate the model. In Chapter 7 the results of such models are presented and an explanation of the main results are obtained. In this chapter, the results obtained are used to answer the research question as well as pointing out some other interesting results that might have importance for the reader.

Finally, in Chapter 8, the conclusions of this thesis are presented as well as giving the last answer to the research question of this project.

3.1 Position of This Thesis

One of the things that can be expected from this thesis is the position assumed in order to develop this thesis. First if all a great amount of time was invested to gather as much information as possible in the topic of EPC and the impact of energy rating in the housing market. The initial idea was to understand up to what extend something like a energy performance certificate can affect the prices and what are the consequences in term of price increase. For that reason, the main theoretical framework used along the project is the idea of making a model based in statistics regressions, particularly the hedonic price method. This method, in which the main features of this project are based, is explained in chapter 6 where the reader is introduced first to the linear regression, how it works, and then brought to the idea of a regression model based in the hedonic theory. This theory is just a way to understand the economic impact of things that cannot be really compared in terms of pricing. This was chosen mainly because some studies presented in Chapter 1.1 reflected that there was a clear relationship between this two aspects of the same reality. For that reason, this thesis is presented and positioned in the framework of an economic crisis in Spain and in particular a crisis of the building environment caused by the explosion of the housing bubble. The idea behind all of this, then, is to emphaticized in the importance of increasing the energy efficiency of homes for the contribution with the climate change, as mentioned in the introduction, but at the same time, the thesis pretends to give reasons on why should people start looking at investing in energy efficiency. Because this study shows that there is a positive benefit from doing it in economic terms for the owner of a dwell and for the rest of the world translated in less pollution and less GHG.

4 Household Energy Efficiency Labelling System in Spain

As mentioned in Chapter 1, GHG have become one of the most important issues for European countries. The building sector is responsible for the greatest consumption of energy produced in Europe, being 37% of the total energy consumption. However, it is also the sector where reduction of GHG emissions can be reduced significantly and in an efficient way, thus offering a great opportunity in terms of reduction of CO2 emissions across the EU (Eurostat Statistics, 2012).

In the European Union, the directive in charge of develop and improve the energy performance of buildings is "The Energy Performance of Buildings Directive (EPBD) (Directive 2002/91/EC)" and the EPBD 2010/31/EU, which substitute the previous one of 2002 and extend its scope reduce the differences in energy efficiency between Member States. In Spain, the implementation of such directive is made almost in its totality by the CTE ("Codigo Técnico de la Edificación" - Building Technical Code) in the Royal Decree 314/2006 in two of its basic documents: "Documento Básico de Ahorro de Energia" and "Documento Basico de Salubridad, respectively called CTE-HE and CTE-HS. There is a significant part regarding the thermal properties which are contained in the "Reglamento de Instalaciones Térmicas de los Edificios" (RITE), and also the Royal Decree 47/2007 in which the basic procedure for the certification of new buildings is approved. In order to understand this, the following figure depicts the adaptation of the EU directive into the Spanish policy.

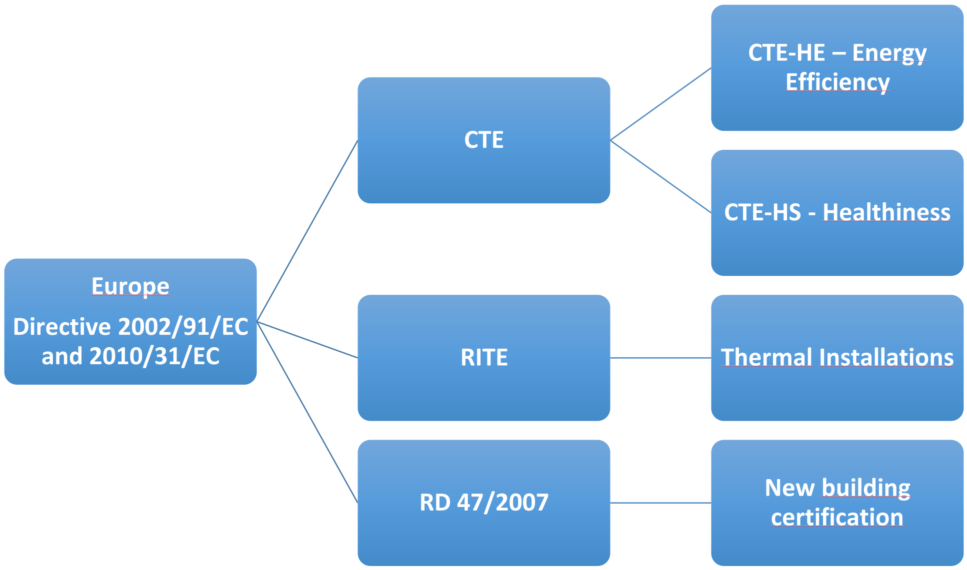

Figure 4.1. Adaptation of EU Directive 2002/91/EU into the Spanish policy

The second EPBD 2010/31/EU was adopted in Spain by enlarging the CTE with the Royal Decree 235/2013, in which the energy efficiency measures for new buildings should be also applied to the existing ones. For that reason, the government developed several programs for which existing buildings and new buildings must accredit an EPC ("Certificado de Eficiencia Energetica"). This programs are structured in two groups. The first group, which are the new buildings in the development and ending phase, will use a program called CALENER VYP, and the second group, which are the existing buildings, will have 3 different programs: CE3, CEX, CERMA, which applies the simplified method to obtain the certification.

The certification that comes as result of the energy efficiency assessment in the program is valid for 10 years and it must be renewed once this period has finished. Nevertheless, if the household happens to have a major refurbishment that may lead to an impact in the energy efficiency of the building, It will be the owner who decides freely if a new EPC must be conducted.

5 DESCRIPTION OF THE SPANISH BUILDING ENVIRONMENT

Buildings will be one of the most important sector in terms of energy and environmental improvements due to the fact that they constitute a “stock” for the future development of the cities, thus stablishing a future energy consumption and environmental emissions (Gago et al., 2012). In Spain there are 2689 million m^2 built in total, with an 85% of it representing the residential buildings and the other 15% estimated for other uses (IDAE, 2011). In terms of households, there are 25M of it, accounting for an 18% of the final Spanish energy consumption (Eurostat & IDAE, 2011). If a 9% is added to this 18% coming from the tertiary sector, the total of 26% differs significantly from the average of the EU, where the building sector accounts for a total of 37% of the final energy consumption (Eurostat Statistics, 2012). However, it is true that some studies has shown that there is a convergence between EU levels and the Spanish ones (IDAE, 2011).

The following figures depicts the energy consumption in the residential sector of Spain since 1990 and energy demand per use in a middle year (2010) respectively.

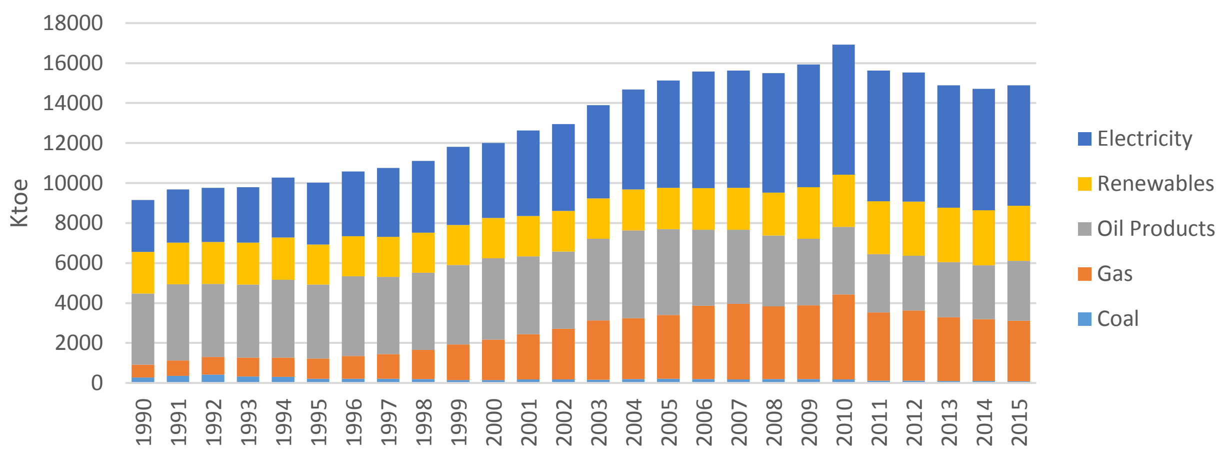

Figure 5.1. Final Energy Consumption in Spain for the Residential Sector (MINETUR, 2017)

It is clear by the graphic that the energy consumption in dwellings has increased 50% from 1990 up until 2006 due to the fact that household income increased during this period (Eurostat & IDAE, 2011). This increase was followed later by a decline, due to crisis and because the following winters where warmer than the previous ones, but is important to remark that it wasn’t because of the energy efficiency measures (WWF, 2012).

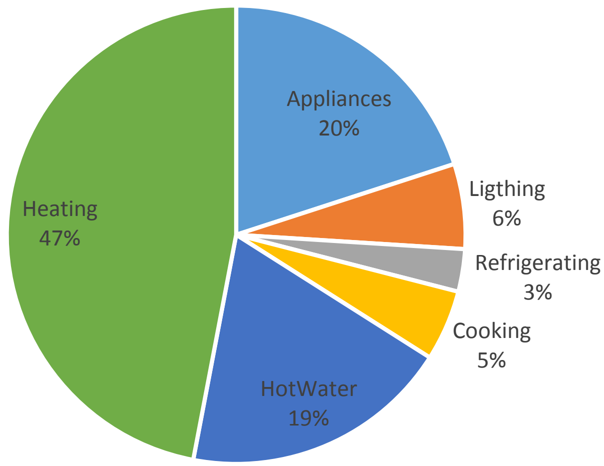

Figure 5.2. Energy consumption per use in a household in Spain (2010) (MINETUR, 2017)

It's important to mention that almost 70% of the households are block buildings and that the rental sector is rather small accounting for a 8% of the households. It means that the norm in Spain are home with ownership (Eurostat & IDAE, 2011). From the figure above, it can be seen that the main expenditure of energy is made in the use of heating for the household. Moreover, almost 45% of the households built in Spain were built before 1980 and therefor have a very low energy efficiency rating in comparison to newer buildings (Eurostat & IDAE, 2011). The current rehabilitation of buildings in Spain is 0,5% per year (WWF, 2012) and the weight of this refurbishment in the construction sector in Spain is 30%.

Spain suffered from 1998 to 2006 a housing bubble, which had a very important impact in the sector. This bubble was promoted by the favorable conditions in the political and economic spectrum which involved the adoption of the Euro as the common currency. Because of this bubble, Spain became the country of the European Union with the most number of households per thousand inhabitants, the place with the most unoccupied dwells and the place with the highest number of second residential homes. (Pérez, 2010)

Nowadays, due to the effects of this bubble and the crisis in Europe, the construction in Spain collapsed leaving the housing sector in a recession. The labor, during the bubble, increased very rapidly because there were plenty of jobs available in the construction market, however with the collapse, this labor force was released, resulting in an exponential rise of the unemployment rate, which currently is over 20% (OECD, 2011).

6 COST ESTIMATION

In this chapter, the technics used to create the model and analyse the main point of the research question is described as well as the description of the dataset. Moreover, the function and model construction is followed step by step and finally the results are presented. The reason why this method has been chosen is that many studies suggest that the theoretical framework of hedonic price model fits perfect with the idea of classifying homes as multiple features that have their inner cost and therefor it is possible to estimate the weight of them in the final price (Gago et al., 2012). As stated in the research question, the purpose of this study is to determine the economic impact of increasing the energy rating and in order to do that a good solution is to identify this qualitative variable that cannot be measured economically and estimate the impact of such variable in the overall price.

6.1 THE METHOD: REGRESSION ANALYSIS AND HEDONIC MODELING

This chapter introduces the regression analysis and a technic called hedonic price modelling for assessment of a dataset. The theories and fundamental of these tools are discussed, as well as how are they build and their potential. The regression analysis is used to determine the relationship between characteristics of buildings and the price (rent or buy) and the hedonic modeling helps to create a function that can be used to determine future prices.

6.1.1 The Regression Analysis

A regression analysis technic is a method used to define the connection between data points of a dataset. The result of it can be used to predict the results of future iterations. It is a tool that explains how different variables (called independent variables) influence another single variable (called dependent variable). These independent variables are also called explanatory variables because they are used “to explain” the behavior of another variable. The most typical regression is the so called least squared method (LSM) which measures the relation between a unique dependent variable (Dy) and a unique independent variable (Dx).

A good example to depict this theory is to establish the correlation of the speed of a car and the number of accidents in a highway. In this case, the independent variable is the speed of the car, and the dependent variable is the nº of accidents. If the speed of the car is plotted in the X-axis and the number of accidents in the Y-axis, a possible plot could be the following:

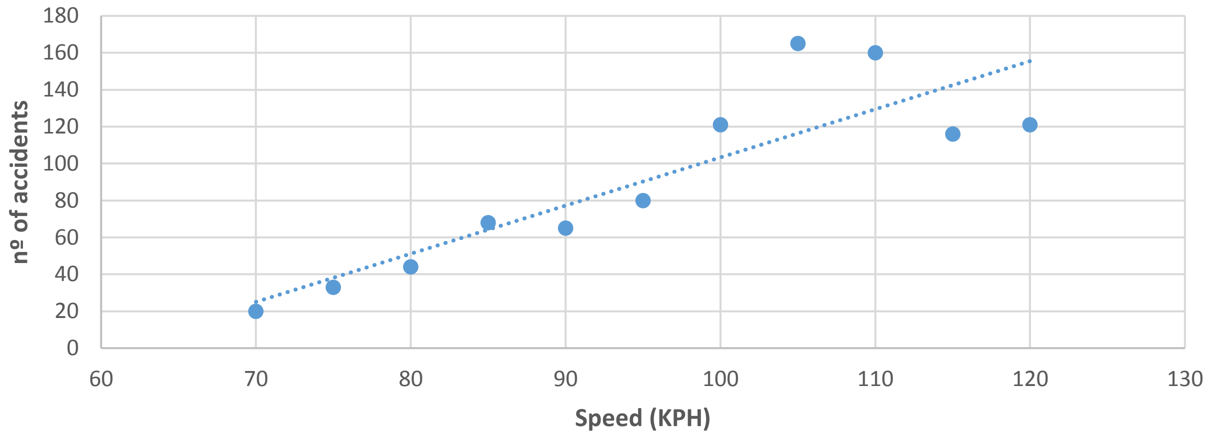

Figure 6.1. Nº of accidents vs speed of a car in a highway – example

One of the conclusion someone can draw from this plot is that as the car speed increase on the highway, the number of accidents also increase, as it can be seen in the trend line. However, it isn't totally true, because this particular case represented in the figure above only depicts one of the innumerable factors that can play a role in an accident (weather, time of the day, maintenance of the car…). For that reason, a multiple regression analysis technique needs to be conducted. This technique is used understand the contribution of different independent variables on the result (dependent variable). The result of such technic produces what is called correlation coefficients for each of the independent variables that was used.

This same methodology can be used in the context of this study, to determine the impacts of different features of houses in the market price, and specially the impact of the labelling rating in the final rent or purchase price. The housing sector can be compared to a grocery bag; each bag can be larger or smaller and they all contain different types of products which impact directly in the overall price of the grocery. By performing a multivariable regression analysis using historical data for bags of grocery, one can determine the correlation of each product with the overall price and the contribution of each item to that final price. By doing this, one could also determine which items impact the most in the price.

According to (Malpezzi et al., 1980) the housing sector is no different. A house can be larger or smaller and have a number of characteristics that impact directly on the overall price (size, nº of dorms, type of kitchen, bathrooms…). These characteristics can be understood as independent variables and therefor a regression analysis can be done to determine how much each of this feature affect the final price.

An advantage of using this method is that it is possible to quantify something that is qualitative. An example can be the EPC label rating: even though the label by itself is not a physical part of the house, it can positively influence the price of the house due to the energy savings it provides to the owner or tenant. This type of feature is called "internal characteristic" and other example are, for instance, the distance to the market, distance to the airport and so. On the other hands, there are "external characteristics" which include features as usable area of the house, building finishes, parking, dorms, etc. In terms of analysis, there are innumerable numbers of characteristics that may increase the value, and is up to every person to determine which the most interesting ones are. That's why a regression analysis using multivariable is perfect for this study, because it helps measuring such characteristics (Bio Intelligence Service, Ronan Lyons & IEEP, 2012).

This paper will analyze this possibility focusing on one special feature: the EPC label, that determines whether a house is energy efficient or not. Moreover, other types of characteristics will be analysed and a multivariable regression will be used to explain what the impacts of this features in the price are.

6.1.1.1 Evaluating whether the regression fits or not

A good model is the one that is able to predict values that are close to the data observed and the values pointed out for that observations. Normally, there is a simple way to measure the fitness by using the "mean model", which uses the mean of the values as predicted value. This model can be used when there is no information for predictor variables, however, in this case, the models is based on predictor variables and therefor there is a need to know whether the model fits better than the mean model. To do that, the most extended way to analyze a model is by analyzing the R^2 and the adjusted R^2.

R^2 and the adjusted R^2 are based on two different sums of square: SST and SSE (Sum of Square Totals and Sum of Square Errors respectively). By doing a difference between SST and SSE once can determine the improvement that the regression model has in comparison to the mean model. Dividing by the SST a R^2 is determined which then indicates the proportional improvement from the model compared to the mean model which at the same time indicates the goodness or badness of the model.

R^2 is structured in a scale that range from zero to one, which make it easy to read and asses in terms of how good or how bad it is. A scale of 0 means that the prediction model doesn't improve the predictions in comparison with the mean model whereas the one indicates that the predictions fits better than the mean model. By improving the model information and filtering it, it is possible to increase the R^2; hence it's really important for the model to have values which are within the scope of the idea of study and also filtering the data to avoid information that may be misleading for the model

A handicap from R^2 is that it only can increase as predictors are added to the model. However, this increase in R^2 can be false and actually not improving the model when the predictors added doesn't fit well. In order to avoid this problem, another statistic indicator is the adjusted R^2, which incorporates the model's degree of freedom; in other words, how can a value be restricted or not to specific value. In this thesis case, by adding coefficients to the model, the degree of freedom of the data is less and therefor the restriction in the possible values are more accurate for the model. Adjusted R^2 will decrease if predictors added to the model doesn't fit correctly and doesn't make up for the loss of degree of freedom. Likewise, it will increase as predictors are added and fitting correctly. This adjusted R^2 should be always used in models with more than one predictor and can be understood as the proportion of total variance that the model can explain (Ramos, 2017)

6.1.1.2 Hedonic Method: Pricing Modeling

The hedonic method is a particular application the multivariable regression method. The method is used to estimate the values of a good breaking it into its constituent and obtain estimations of the contributory value of each of this constituents into the final price. Hedonic modeling technique are commonly used in the real estate market to estimate the price of individual features of a property the price of which are not directly observed or quantified. It is acknowledged by some studies as the most accepted and widespread empirical method (Markandya, 2015) .

The application of the hedonic method is done by using the coefficients created by the regression analysis of each independent variable against the dependent variable. So, it can be said that market price is a "function of each internal and externa characteristic". The following equation represent this relationship:

$$Price = f(external \,\&\, internal\ features, building\ characteristics, other\ factors)$$

By collecting information and data on buildings and their transaction price, it is possible to do a regression analysis that will determine the relation between the characteristics and the price. Those correlation factors will be then used to stablish the formula (model) and to calculate the proportional weight of those variables and how the impact the final price.

6.2 DESCRIPTION OF THE DATA USED

All the data used in the modeling was obtained from different sources. First, major bulk data was obtained from Integrated Public Use Microdata Series, International (IPUMS) for the years 2001 and 2011 which required first to contact the organization and register and account with academic purpose. From this website all the data related with dwellings characteristics such as rooms, usable area, air conditioning and other characteristics were obtained. More than 600.000 register were downloaded from the website database and then filtrated in order to meet the requirements of the study. The filtering method followed the next classification:

- IPUMS offer data classified in country, state, province and municipality. The first filter consisted in selecting only a few regions from Spain that are different in terms of geography in order to have values form all the county. Four big major cities were selected: Asturias, Barcelona, Madrid and Sevilla. The main reason to not include all the register was the impossibility to filter and disaggregate the data in useful data, hence the reason why the model is based in four different provinces located in different parts of Spain.
- The data regarding these cities was then classified in municipalities using IMUMS code system. Moreover all the houses without a certification were excluded from the dataset, leaving only the houses with an energy label.
- All the data coming classified as "Other municipalities" was erased from the dataset in order to be able to link the municipalities with the amount of population of it (this variable is used as independent variable)
- The downloaded data was divided into people of the same dwell. Since the model only needs register for each house, all the register pointing to a same house were erased except the first one.
- All the buildings without information in number of rooms, area, or any other variable which wasn't know was erased from the dataset
- Buildings with a living area less than 20m^2 and more than 500m^2 was deleted from the dataset because they were irregular points that weren't consistent with the overall data

- Finally, the data was selected and classified in a total of 1871 different register from the four provinces of Spain.

From the same dataset downloaded from IPUMS it was possible to determine the municipality and province of each register, which was divided into 4 new variables each one stating the province of the house (Asturias, Barcelona, Madrid or Sevilla). This variable are the ones used a geographical independent values and have a value of 0, if the house is not in that province, or 1 if it is situated in the province.

To complement this data, information from the "Instituto Nacional de Estadística (INE)" (National Institute of Statistics) were also downloaded and scrubbed for each province. From INE, it was possible to get an estimation of distance to nearest hospital for each municipality. However, since the data was an estimation between 2 possible values, a randomize function was used in order to get different points of data for each house and therefore not having flat data which can distortion the model.

The data regarding prices was obtained from multiple sources: INE, Ministry of Development (Ministerio de Fomento), Madrid City Hall website, Sevilla City Hall website, Asturias City Hall website, Barcelona City Hall website and Real State Cadastre. In the following table, the variables and a summary of the statistics are

Variable Name	Type	Description
LN(TRANSFPRICE)	Quantitative	Natural log of the transfer price
EPC_ABC	Binary	Whether the building is rated A,B or C
BUILDYEAR	Quantitative	Building year of construction
FLOOR	Quantitative	Floor number of the dwell
LIVEAREA	Quantitative	Living area of the dwell
ROOMS	Quantitative	Number of rooms in the dwell
AIRCOND	Binary	Whether the dwell has air conditioner or not
HOTWATERHEATER	Binary	Whether the dwell has hot water installation or not
CENTRALHEAT	Binary	Whether the dwell has central heating or not
INTERNETACCES	Binary	Whether the dwell has access to internet or not
ELEVATOR	Binary	Whether the dwell has elevator or not
ASTURIAS	Binary	Binary variable that indicates if the dwelling is in Asturias
BARCELONA	Binary	Binary variable that indicates if the dwelling is in Barcelona
MADRID	Binary	Binary variable that indicates if the dwelling is in Madrid
SEVILLA	Binary	Binary variable that indicates if the dwelling is in Sevilla
MUNIEPOP	Quantitative	Population in the municipality of the dwell
HOSPDIST	Quantitative	Distance to the nearest hospital

Table 6-1. Summary of variables – Name, Type and Description

6.2.1 Main findings of the data

From 1871 register used in the model, 464 of them are from Asturias, 499 from Barcelona, 545 from Madrid and 362 from Sevilla. The information extracted involves all the possible types of housing, from flats or apartment buildings which constitute 78% of the register to detached houses and others, which involves 22% of the data. Those houses have an average number of rooms of 5 rooms counting the kitchen. The average structure in terms of people living in a house is three members being a couple with a child the most typical structure. The average number of kids per household is 1 and the average age of the elder children is 21 years old. The main head of the house typically is a man (70%) with an age between 45 to 55 years old.

As for the house characteristics, the main sources of heating and hot water are electricity and natural gas, accounting for a total of 86% of the total heating systems. Regarding this systems, 51% of the heating systems are individual heating whereas 13% are collective systems. In this feature, 31% of the houses have other types of heating system that are not central heating. In terms of air conditioner, 20% of the houses used in the model have an air conditioning system whereas the other 80% doesn't have any. Moreover 97% of the houses have an internet and/or telephone access which means that it possible for them to connect to the internet.

In terms of pricing and cost of the buildings, the range is quite broad, ranging from 30.000€ to 975.000€ with a mean value of 139.207€. These values, even though they are not used directly in the model, will be used as the dependent variable in form of natural log of the value. In this way, the results obtained in the hedonic model re easier to understand (explanations in the next chapter).

6.3 MODEL

A hedonic regression model will be constructed in the form of semi-log model using Microsoft Excel. Excels offer a really easy way to make multivariable regression and the ease of use makes this program the best choice, avoiding the time needed to understand and process the information in other types of software. Moreover, Excel's results are obtained fast and in a way, that can be integrated directly in the results of this thesis. Other options as STATA were also analyzed but the requirement of paying and the time it needed in order to be able to work with it discarded this option. The idea of building a semi-log model comes from some studies that shows that this type of model is easier and simple to understand and also offer a good relationship between the independent and dependent variables within the housing field. The theoretical foundations of such model were developed by (Rosens, 1974) and complemented by other studies. The main idea of this modeling is to capitalize a good based on the maximization of utilities and features set by the consumers. The reason for the semi-log is that when an item which price needs to be determine by a regression cannot be easily restocked, like for instance dwells, then some non-linearities arise in the model and therefor using a linear-linear model wouldn't quite fit the study. To deal with this problem, it is common to employ a semi-logarithmic functional form for the dependent variable.

The semi-log hedonic price has some advantages over the linear-linear model. The main advantage from log-linear (semi-log) is that it gives the possibility to a given feature (number of bathdrooms, for instance) to vary in a proportional way with the value of another characteristics (number of rooms). On the other case, in a linear-linear model, a second bathroom would add the same value to the house that has one bedroom as to one with four bedrooms

First, it is considered that the idea of housing is divided into different economic goods of multiple physical and location characteristics, which are going to be called "dwelling features". This dwelling features consist in the following variables: dwell building year (BUILDYEAR), living area (LIVEAREA),number of rooms (ROOMS), floor location of the house (FLOOR), whether the house have or doesn't have air conditioning (AIRCOND), hot water installation (HOTWATER), central heating (CENTRALHEAT), access to the internet (INTERNET) and elevator (ELEVATOR). This last 5 variables will have a value of 1 if the building have access to that characteristic or 0 if it doesn't. The most important variable for this study is whether the building have an EPC rating of A, B or C (EPC_ABC) and will also have a binary value of 1 or 0.

$$T_{price} = X_{dwellS} \, , \, Y_{geograph} \, , \, Z_{socio-economic}) \qquad (1)$$

Where:

$$X_{dwellS} = X(x_{EPC_{ABC}}, x_{BUILDYEAR}, x_{LIVEAREA}, x_{ROOMS}, x_{FLOOR}, x_{HOTWATER}, x_{CENTRALHEAT}, x_{INTERNET}, x_{ELEVATOR})$$

Moreover, for this study it is considered that houses have a geographical factor called "geographical features" which consist of four binary variables that match the four different locations of the building. This locations are Asturias, Barcelona (Nord West and North East), Madrid (Center), Sevilla (South). This location were selected in order to have buildings from different geographical location and to include any dependency that might be possibly exist between prices and locations.

$$Y_{geograph} = Y(Y_{ASTURIAS}, Y_{BARCELONA}, Y_{MADRID}, Y_{SEVILLA}) \qquad (2)$$

Finally, it is also assumed in the modeling that there are variables indicating some aspects of the socio-economic environment. In order to do that, two variables has been chosen which will include aspects that are outside the housing field but can play a role in determining a dwell price. This variable consist of: population of the municipality where the house is located (MUNIEPop) and distance to the nearest hospital (HOSPDIST).

$$Z_{socio-economic} = Z(Z_{MUNIEPop}, Z_{HOSPDIST}) \quad (3)$$

Applying a semi-log hedonic model the natural log of the transaction price for a home is a function of all the independent variables (external and internal):

$$\ln(TRANSFPRICE_i) = \varphi_i + \alpha_1 x_{EPC_{ABC\,i}} + \alpha_2 x_{BUILDYEARi} + \alpha_3 x_{LIVEAREAi} + \alpha_4 x_{ROOMSi,} + \alpha_5 x_{FLOORi} + \alpha_6 x_{HOTWATERi} + \alpha_7 x_{CENTRALHEATi,} + \alpha_8 x_{INTERNETi} + \alpha_9 x_{ELEVATORi} + \alpha_{10} Y_{ASTURIASi} + \alpha_{12} Y_{BARCELONAi} + \alpha_{13} Y_{MADRIDi} + \alpha_{14} Y_{SEVILLAi} + \alpha_{15} \mathrm{Z}_{\mathrm{MUNIEPopi}} + \alpha_{16} \mathrm{Z}_{\mathrm{HOSPDISTi}} \quad (4)$$

Where i refers to the household number, φ_i is a constant term of the equation, $\alpha_1 \ldots \alpha_{16}$ are the coefficients that will be estimated with the multivariable regression. In the following table a summary of the variables in their main statistics are shown:

Variable Name	Type	Description	Obs	Average	Max	Min
LN(TRANSFPRICE)	Quantitative	Natural log of the transfer price	1870	11,743	13,791	10,321
EPC_ABC(D)	Binary	Whether the building is rated A,B or C	1870	0,089	1	0
BUILDYEAR	Quantitative	Building year of construction	1870	1967	2001	1899
FLOOR	Quantitative	Floor number of the dwell	1870	4,833	10	1
LIVEAREA	Quantitative	Living area of the dwell	1870	90,17	500	30
ROOMS	Quantitative	Number of rooms in the dwell	1870	4,994	20	1
AIRCOND	Binary	Whether the dwell has air conditioner or not	1870	0,199	1	0
HOTWATERHEATER	Binary	Whether the dwell has hot water installation or not	1870	0,374	1	0
CENTRALHEAT	Binary	Whether the dwell has central heating or not	1870	0,634	1	0
INTERNETACCES	Binary	Whether the dwell has access to internet or not	1870	0,971	1	0
ELEVATOR	Binary	Whether the dwell has elevator or not	1870	0,282	1	0
ASTURIAS	Binary	Binary variable that indicates if the dwelling is in Asturias	1870	0,251	1	0

BARCELONA	Binary	Binary variable that indicates if the dwelling is in Barcelona	1870	0,260	1	0
MADRID	Binary	Binary variable that indicates if the dwelling is in Madrid	1870	0,294	1	0
SEVILLA	Binary	Binary variable that indicates if the dwelling is in Sevilla	1870	0,195	1	0
MUNIEPOP	Quantitative	Population in the municipality of the dwell	1870	408.122	3.165.541	16.850
HOSPDIST	Quantitative	Distance to the nearest hospital	1870	9.405	17.993	10

Table 6-2. Summary of variables and their statistics

For this study, there will be two models that will vary only in the specification of our variable of interest: the first model will take into consideration the variable EPC_ABC as a whether or not the building have an A, B or C rating whereas in the second model, the rating D will be included. As mentioned, this variable of interest is a binary variable or also called "dummy" variable. The semi-log model for the dependent variable makes it easier to interpret the coefficients as average percentage of the premium price. This interpretation is based in a simplified adjustment formula proposed by (Van Garderen, 2002) which describes the method where the proportional impact p_j of a dummy variable (our case for the variable EPC_ABC) on the dependent variable in a semi-log regression model is as follow:

$$\% p_j = e^{(c_j)} - 1 \qquad (5)$$

Where c_j is the coefficient estimated for the dummy variable in the regression model.

7 Results & Discussion

Housing stock has been receiving an increasing policy focus on trying to increase and improve the energy performance due to the constant pressure the human kind is impacting on the climate change. The main action was materialized with the mandatory energy performance certificate in the EU which is aimed to change the tenant and owners behaviours by giving clear information on the energy performance of the buildings. It is known that some of the attributes of the dwellings cannot be quantified directly like for instance energy efficiency but it is also known that the introduction of a market-based instrument as the EPC will capitalize this kind of attributes and eventually stimulate the supply of new efficient houses and the renovation of the existing ones to improve the energy performance (Amaia de Ayala et al., 2015).

Using a sample of 1.870 dwells in Spain with an EPC certificate, it is possible to see that the major part of dwellings in the dataset are clustered in the middle position of the certification range (D,E,F) which represent 70% of the houses . The category with the major number of houses is the label F with 32% whereas the least category is A with only 14 dwells (1%). As it is expected as well, there is a relationship between EPC rating and the year of construction of the dwell: in this case, 83.83% of the buildings build after 1980 have a rating of C or better whereas only 16.17% is accounted for buildings that were built before 1980. The following figures show these distributions.

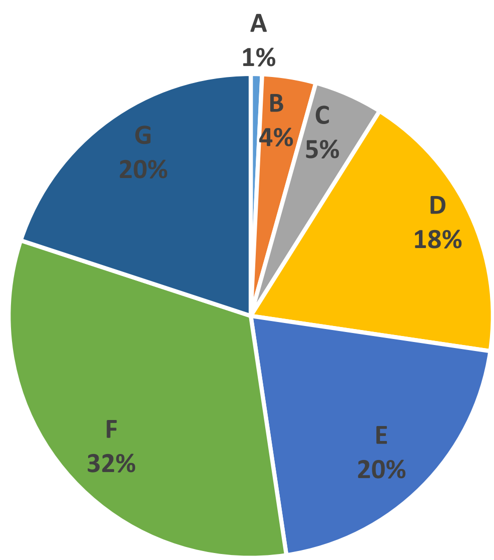

Figure 7.1. Distribution of dwells per EPC label

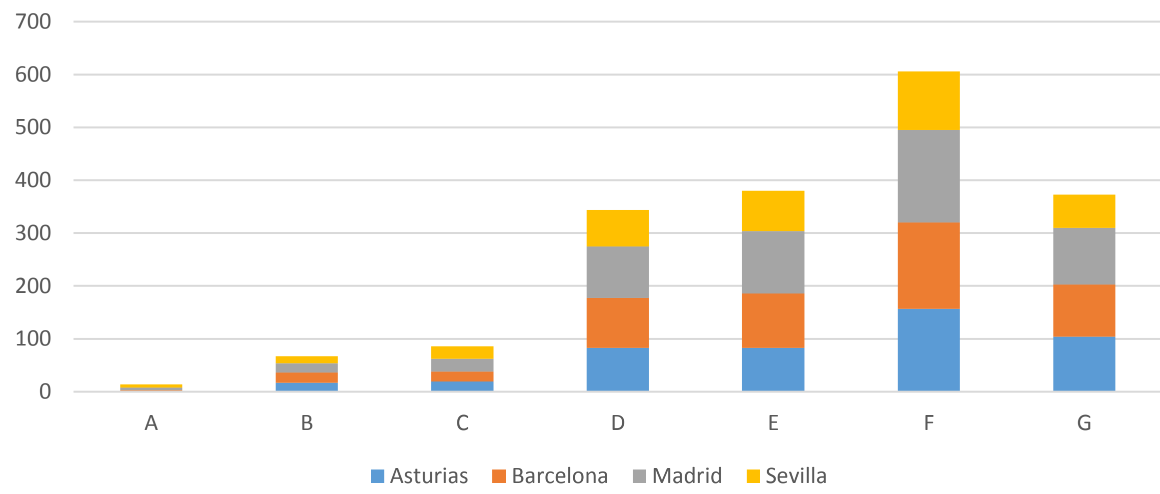

Figure 7.2. Distribution by region and EPC label

In terms of pricing, the majority of the houses are located in prices that range between 100.000€ and 200.000€, representing 56% of the total number of dwells. The second category in number of dwells is the category of "less than 100.000€. This information is represented in the following figure:

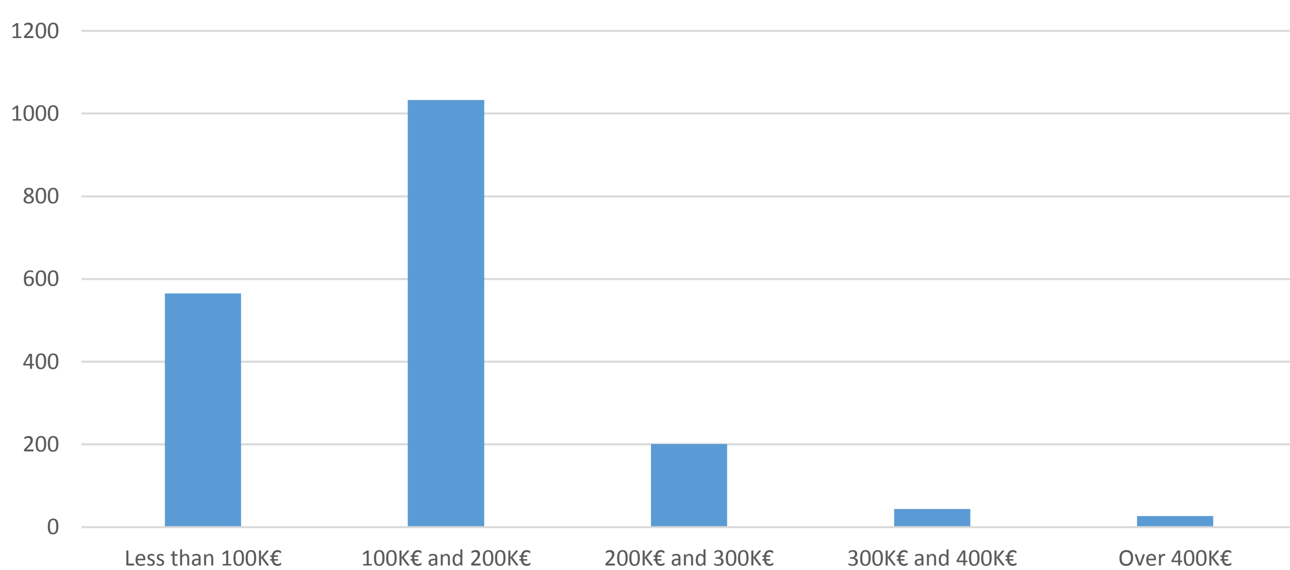

Figure 7.3. Range of prices and number of dwells

In order to evaluate and assess the impact of EPC labels on the price of houses, a hedonic price model is used. This model has been presented in the equation (4). This model is based in an ordinary multivariable least square approach using the Microsoft Excel 2015 software Analysis ToolPack add-in. In order to be able to use the regression function from Excel, it is necessary to activate this feature trough the internal option due to the fact that this feature is an "add-in" and is disabled by default. The model in equation (4) has a variable called EPC_ABC(D) which is binary and can have a value of one if the house has a EPC label of A, B or C for the first case and also D for the second case and a value of zero otherwise. Table 7-1 presents the estimation of the coefficients of the different used variables in the first model, where the control variable studied is the EPC_ABC variable. Using the formula (5), it is possible to determine the proportional impact of our variable of control for dwellings with an A, B or C label:

$$\%p_{EPC\_ABC} = e^{(0,10221463)} - 1 = \mathbf{10,76}\%$$

This means that dwells with these labels have a 10,76% higher value (premium price) compared to homes of similar features but with a lower energy EPC label. From the dataset, it was determined that the average price of the dwells was 139.207€ which means that this premium price paid for houses with a better energy label is 14.981€

Variable	*Coefficients*	*Variable*	*Coefficients*
EPC_ABC	0,102214625	**INTERNETACCES**	0,029009602
BUILDYEAR	0,004692180	**ELEVATOR**	0,068769609
FLOOR	-0,016071232	**Asturias**	0,047292999
LIVEAREA	0,008594834	**Barcelona**	-0,010988262
ROOMS	0,012420681	**Madrid**	-0,067521542
AirCond	0,015289719	**Sevilla**	0,048355736
HotWaterHeater	0,022697703	**MUNIEPop**	4,16135E-08
CENTRALHEAT	0,101122151	**HOSPDIST**	-1,28739E-06
Regression Statistics			
	Multiple R	0,89794685	
	R Square	0,80630854	
	Adjusted R Square	0,80420209	
	Standard Error	0,19091175	
	Observations	1870	

Table 7-1. Model 1 EPC_ABC - Coefficient estimation

In order to understand the impacts of a lower energy efficiency rate, a second model was developed. This second model based in the same hedonic price theory have an alternative EPC_ABC variable, which includes the label D in the set. For the second model, all the other variables remained unchanged and only the variable EPC_ABC was change adding a value of one to the dwells with a rating "D". The resulting coefficients from this second model are showed in the Table 7-2. In this case, the extra proportional impact of adding label "D" as in our variable of control is as follow:

$$\%p_{EPC\_ABCD} = e^{(0,04165846)} - 1 = \mathbf{4,25}\%$$

As mentioned in the first model, this means that in this model the price premium paid for dwells of labels A, B, C or D is 4,25% higher than those dwells of lower rating. In terms of monetary values, this 4,25% represents an additional 5.922€ to the average house relative to the dwells with same characteristics but with less energy performance.

Variable	*Coefficients*	*Variable*	*Coefficients*
EPC_ABCD	0,041658461	**INTERNETACCES**	0,033224861
BUILDYEAR	0,004678389	**ELEVATOR**	0,006827561
FLOOR	-0,015987276	**Asturias**	0,034512254
LIVEAREA	0,008635337	**Barcelona**	-0,010885797
ROOMS	0,011528125	**Madrid**	-0,067405183
AirCond	0,015207929	**Sevilla**	0,051394203
HotWaterHeater	0,021425432	**MUNIEPop**	4,17097E-08
CENTRALHEAT	0,103315659	**HOSPDIST**	-1,28281E-06
		Regression Statistics	
	R^2	0,803624117	
	Adjusted R^2	0,801495941	
	Standard Error	0,192230154	
	Observations	1870	

Table 7-2. Model 2 EPC_ABCD- Coefficient estimation

Joining together the results coming from both models, it can be seen that the houses in Spain with higher energy performance have a premium price over the average between 4,25% and 10,76% relative to the less efficient dwells. Moreover, the results of the regression models show that houses located in Asturias (Nord) and Sevilla (South) are valued on average higher than those houses located in Barcelona and Madrid. Regarding the characteristics of the buildings, the feature that values more positively the price of the house is whether it has any kind of central heating or not. The model also shows that characteristics like number of rooms, air conditioning, hot water and internet access have a positive impact in the value of the house whereas the location of the house in term of floor affects negatively the higher the house is located. In terms of the socio-demographic variables, the models have shown that they have a very little impact on the price but is important to notice that the population affects positively the higher it is whereas the distance to the nearest hospital affects negatively the higher it is, as it is expected.

Finally, in terms of the models fitting the study, it is clearly observable that the data had a very good fitting in both models where the adjusted R^2 is 80,6% in the first model and 80,1% in the second model. This means that the dataset used in the modeling can be predicted by the model with a success of approximately 81%, which is a very good result based on the many predictors that the model has used.

7.1 P-VALUE ANALYSIS AND THE SIGNIFICANCE OF THE VARIABLES USED

Before analyzing what p-value is and what is the meaning of a high or low p-value it is important, first, to clarify what is the null hypothesis to which the p-value makes reference. A null hypothesis basically says there is no statistical significance between the variables of an hypothesis. In the case of this thesis, it is trying to demonstrate that there is a significant relationship between the independent variable (natural log of price) and the rating label of the dwell. In order to claim that our hypothesis (that there is a link between this variables) is true it is necessary to disprove the null hypothesis is true, which will mean that the effect of measuring the EP label has no effect in the final price (coefficient = 0), which would mean that there is high probability to obtain the same results from a different observations. To do so it's is recommended by many studies that the maximum level of significance (that is, maximum error called "α") should be lower or equal to 0,05. If the p-value is less than 0,05 the null hypothesis can be rejected meaning that there is a strong evidence against the null hypothesis (that is, if the null hypothesis was true it would be unlikely to get the same observations) whereas if is greater, it shows that there is a weak evidence against the null hypothesis (it is likely probable to get the same results when the null hypothesis is true) (Eberly College of Science, 2017). In other words, the p-value of each predictor tests the null hypotesis in which the coeffficient of the predictor is equal to 0 (which means it has no effect in the final Price). A low p-value, indicates that this null hypothesis can be rejected meaning that any change in the predictor will carry a change for the price, whereas a high p-value means that there is no effect of the variable in the final price.

Assuming that our model will have a maximum level of significance of 0,05, it is important to show whether the variables adjust to this maximum or not. Table 7-3 shows the p-value obtained for each variable of both models. The variable of control that this thesis is studying is the EPC_ABC(D) which has a p-value that is less than our maximum level of significance of 0,05. So, it is possible to conclude that the null hypothesis can be rejected and therefor there is a strong evidence that there is a relationship between this two variables. However, it is important to notice that there are 4 variables which has a very high p-value and therefor it is not possible to disclaim the null hypothesis for them. That means that it is possible to obtain valued as the one observed because the variable doesn't have any affectation in the final price. This would mean that for future studies this variables can be rejected from the equation because they won't produce any significant change in the final price of the building.

	Model 1 (ABC)	**Model 2 (ABCD)**	**Less than 0,05?**
EPC_ABC	6,05459E-11	3,20714E-05	TRUE
BUILDYEAR	1,0455E-112	7,8814E-111	TRUE
FLOOR	1,20271E-13	2,32906E-13	TRUE
LIVEAREA	0,00113135	1,3145E-14	TRUE
ROOMS	0,001201774	0,002804768	TRUE
AirCond	0,207110868	0,212700711	FALSE
HotWaterHeater	0,022577523	0,032580187	TRUE
CENTRALHEAT	1,30513E-18	4,13197E-19	TRUE
INTERNETACCES	0,283034751	0,221981011	FALSE
ELEVATOR	0,48618439	0,492311483	FALSE
Asturias	0,00885276	0,00998138	TRUE
Barcelona	0,00032143	0,00066641	TRUE
Madrid	9,77954E-07	1,21051E-06	TRUE
Sevilla	0,008535704	0,005480856	TRUE
MUNIEPop	6,28188E-10	7,5375E-10	TRUE
HOSPDIST	0,14498324	0,149211442	FALSE

Table 7-3. p-values obtained for the models

8 CONCLUSIONS

Throughout all this thesis, it has been mentioned the importance of low carbon policies across Europe countries and specially the implications that the building sector has in it. This sector is playing an important role due to the fact that, as some studies suggest, there is a large potential within it in terms of energy conservation and energy performance. The main tool to promote the energy efficiency in houses in Europe was the EPBD 2002 and its recast in 2010, which carried with them a gradual implementation across the European countries. Focusing in Spain, this EPBD main policy was implemented in 2013 with different policies divided in the area in which it affects. CTE and RITE are the main policies that regulate the implementation of the EPC in existing and new buildings at the time of construction, rent or sell. One of the assumption that have been made during this thesis is the fact that if improvements in the energy performance are translate in higher price in the housing market, it will bring a new whole idea in terms of demand and supply of new buildings with better energy efficiency performance and promote the refurbishment of the existing ones in order to get this capitalization. In this frame, this thesis makes a glimpse into the impact of energy labels and the price of houses in Spain.

The information used for these models was obtained from different sources and databases as well as from direct contact to the relevant administrations (see chapter 6.2 Description of the Data Used). According to the first hedonic model estimations done in this thesis, in equality of features of the dwells, houses that have an EPC of A, B or C have a higher value of 10,76% in comparison with homes rated D, E, F or G whereas in the second model it is shown that if the D label is included the price of the dwells have 4,25% extra price (premium price) in comparison with the other low energy performance houses. It is important to remark that even though the results are similar to the ones observed in the literature review, the model has some flaws that still need to be corrected. This flaws are observed in Chapter 7.1 where some of the variables can be seen that doesn't have any effect in the hypothesis that this thesis is accepting. For future study this variables need to be re-adjusted or eliminated. The inclusion of new variables should be also a thing to bear in mind. Even though some of the variables doesn't have an effect in the final price, the model itself has a very good fit, accounting for a adjusted R^2 of approximately 80%.

Nevertheless, it's important to remark the limitations of this thesis. First of all, the range of scope this thesis can accomplish cannot be compared to a real statistic study due to the fact that the limited knowledge on the topic doesn't entitled this study to be completely real. There are some things that need to be improved for a further study in order to have a clear statistic view of the project. Moreover, the deepness and the reliability of the information can still be improved in different ways. One of the things this thesis struggled for a long time was obtaining the information of the dwells, their certification and the estimated transaction price of each house. This information was all obtained from databases that might be out of date and could have information that haven't been certified or checked by an official authority, however it was the only information that was available at the time of making the thesis. Second, the filtering of the information needs to be more accurate and able to disaggregate the information that has a potential use and the information that have potential harm capabilities in terms of misinforming the characteristics of the houses.

Moreover, some of the variables were not available for the observations and some estimations based on the information from the local municipality authorities and other institutions were made. Even though the information might not be the most up-to-date and the most accurate, the modeling method is still consistent and coherent with the theories of hedonic price. For the future, the only thing needed is to verify the information and integrate new variables and new datasets that brings added value to the current information and therefor improve the overall results of the modeling.

Finally, there are some aspects that need to be reminded in order to make a successful implementation of the EPC. First and most important owners or tenants of the dwells should be able to understand what an EPC is and what it is used for: they need to know what kind of information is in it, make useful choices based on the information and know what are the potential benefits from doing energy efficiency measures. This thesis exactly pretends that, to give a grasp on this regard as it can provide an estimation of the benefits from improving the energy performance in the Spanish household market.

9 REFERENCES

Amaia de Ayala et al. (2015). *The price of energy efficiency in the Spanish housing market.* Alava: Basque Centre for Climate Change.

Australian Bureau of Statistics. (2008). *Energy efficiency rating and house price in the ACT.* Canberra: Department of the Environment, Water, Heritage and the Arts,.

Bio Intelligence Service, Ronan Lyons & IEEP. (2012). *Energy performance certificates in buildings and their impact on transaction prices and rents in selected EU countries, Final report prepared for European Commission.* DG Energy.

Brounen, D. (2011). *On the economics of energy labels in the housing market.* Journal of Environmental Economics and Management.

Cajias, M. (2013). *Green performs better: Energy efficiency and financial return on buildings.* Journal of Corp Real Estate.

Chuchi, A., & Sweatman, P. (2012). *A national perspective on Spain's buildings sector: A roadmap for a new housing sector.* http://www.climatestrategy.es/index.php?id=23. Retrieved from A national perspective on Spain's buildings sector: A roadmap for a new housing sector.

de la Rue du Can et al. (2011). *Country review of energy−efficiency financial incentives in the residential sector.* Berkeley: Ernest Orlando Lawrence Berkeley National Laboratory.

Dinan, T. (1989). *Estimating the implicit price of energy efficiency improvements in the residential housing market: A hedonic approach.* Journal of Urban Economics.

Eichholtz et al. (2010). Doing well by doing good? Green office buildings. In *The American Economic Review* (pp. 2492-2509).

European Comission. (2010). *Evaluation of the Energy Performance of Buildings Directive 2010/31/EU.* EC.

European Comission. (2017). *European Comission*. Retrieved from Climate Action: https://ec.europa.eu/clima/policies/strategies/2020_en

Eurostat & IDAE. (2011). *Análisis de lconsumo energético del sector residencial en España.*

Fuerst et al. (2015). *Does energy efficiency matter to home−buyers? An investigation of EPC ratings and transaction prices in England. Energy Economics.* England.

Gago et al. (2012). *ClimateChange ,Buildings and Energy Prices.* EconomicsforEnergy.

Gago et al. (2013). *Handbook on energy and climate change.* Edward Elgar, Cheltenham.

Gilmer, R. (1989). *Energy labels and economic research: An example from the residential real .* Energy Economics.

IDAE. (2007). *PNAEE 2008-2012.* Ministerio de industria y energia.

IDAE. (2011). *Plan de ahorro y eficiencia energética 2012 – 2020. Informe Técnico.* Madrid: Instituto para la Diversificación y Ahorro de Energía.

IEA . (2007). *Mind the gap. Quantifying principal-agent problems in energy efficiency.* Paris: OECD/IEA.

IPCC. (2014). *Summary for Policymakers. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* United Kingdom and New York: Cambridge University.

Jaffe, A., & Stavins, R. (2010). *The energy-efficiency gap. What does it mean? Energy Policy.*

Laquatra, J. (1986). *Housing market capitalization of thermal integrity.* Energy Economics.

Levine et al. (2007). *Residential and commercial buildings. In Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.*

Longstreth, M. (1986). *Impact of consumers' personal characteristics on hedonic prices of energy−conserving durables.* Energy 11.

Malpezzi et al. (1980). *Characteristics prices of housing in 59 metropolitan areas.* Department of Housing and Urban Development.

Markandya et al. (2015). *Policy instruments to foster energy efficiency (in Green energy and efficiency. An economic perspective).* Berlin.

Markandya, A. (2015). *olicy instruments to foster energy efficiency.* Springer.

MINETUR. (2017, 05 03). *Consumo de energia final*. Retrieved from http://sieeweb.idae.es/consumofinal/bal.asp?txt=Residencial&tipbal=s&rep=1

Nanda, A et al. (2011). *Measuring the financial performance of green buildings in the UK commercial property market: addressing the data issues.*

Nobe et al. (2011). *Valuing green home designs: A study of ENERGY STAR® homes.*

OECD. (2011). *OECD Perspectives: Spain Policies for a Sustainable Recovery. .*

Pérez, J. G. (2010). *The real estate and economic crisis.* Susteinability 2.

Pivo, G. (2010). *Income, value, and returns in socially responsible office properties.*

Ramos, K. (2017, Mayo 28). *The Analysis Factor*. Retrieved from http://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/

Reichardt et al. (2012). *Sustainable building certification and the rent premium: A Panel Data Approach. Journal of Real Estate Research.*

Rosens, S. (1974). *Hedonic prices and implicit markets: Product differentiation in pure competition.* The journal of political economy.

Sundberg, A. (2010). *Is sustainability reflected in commercial property prices: an analysis of the evidence base.* Kingston University.

Ürge-Vorsatz, D. (2010). *Counting Good: quantifying the co-benefit of improved efficiency in buildings.*

Van Garderen, K. (2002). *Exact interpretation of dummy variables insemilogarithmicequation in semilogarithmic equations.* Econometric Journal.

WWF. (2012). *Potential energy savings and CO2 emission reduction from Spain's existing residential buildings in 2020.* WWF Spain.

WWF. (2012). *Retos y oportunidades de financiación parala rehabilitación energética en España.* Madrid: WWF/Adena.