UNIVERSIDADE DE ÉVORA

Is Energy Efficiency Reflected in Residential Property Prices in Portugal?

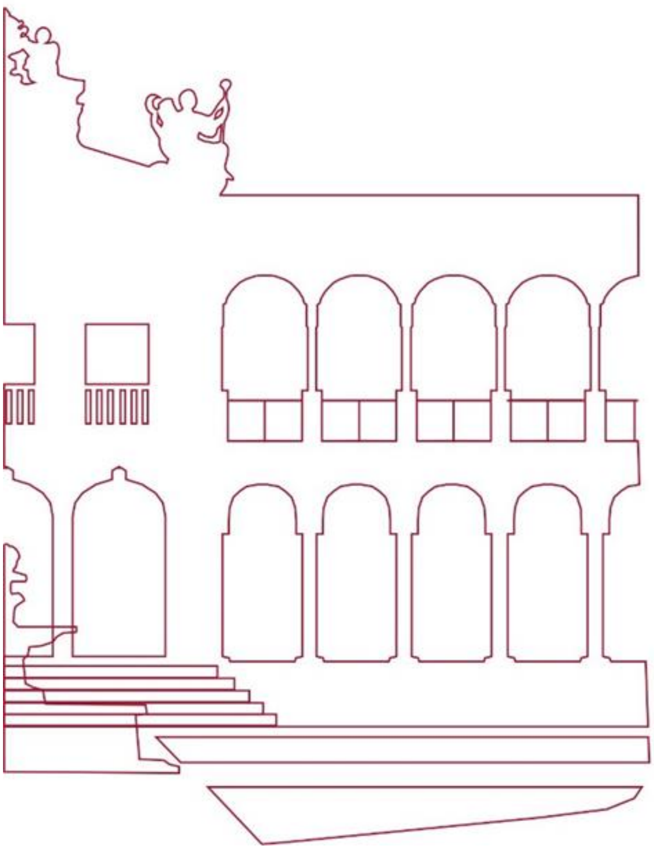
An Investigation Based on Hedonic House Price Functions and Quantile Regression Analysis

Rui Alexandre Alves Evangelista

Orientadores | *Prof. Doutora Esmeralda de Jesus Ratinho Lopes Arranhado Ramalho*
Prof. Doutor João Manuel de Sousa Andrade e Silva
Prof. Doutora Fernanda Paula Mora Peixe

Tese apresentada à Universidade de Évora para obtenção do Grau de Doutor em Economia

Évora, julho de 2019



INSTITUTO DE INVESTIGAÇÃO E FORMAÇÃO AVANÇADA

Is Energy Efficiency Reflected in Residential Property Prices in Portugal?

An investigation based on hedonic house price functions and quantile regression analysis

Rui Alexandre Alves Evangelista

Orientadores | *Prof. Doutora Esmeralda de Jesus Ratinho Lopes Arranhado Ramalho*
Prof. Doutor João Manuel de Sousa Andrade e Silva
Prof. Doutora Fernanda Paula Mora Peixe

Tese apresentada à Universidade de Évora para obtenção do Grau de Doutor em Economia

Évora, julho de 2019

Júri:

Presidente:

Doutora **Maria Aurora Murcho Galego**, Professora Associada da Universidade de Évora (por delegação do Diretor do Instituto de Investigação e Formação Avançada);

Vogais:

Doutora **Esmeralda de Jesus Ratinho Lopes Arranhado Ramalho**, Professora Associada com Agregação da Universidade de Lisboa (**Orientadora**);

Doutora **Isabel Maria Dias Proença**, Professora Associada com Agregação da Universidade de Lisboa;

Doutora **Maria Aurora Murcho Galego**, Professora Associada da Universidade de Évora;

Doutor **Paulo Manuel Marques Rodrigues**, Professor Catedrático Convidado da Universidade Nova de Lisboa;

Doutor **Luís Filipe Farias de Sousa Martins**, Professor Auxiliar com Agregação do ISCTE – Instituto Universitário de Lisboa;

Doutora **Andreia Teixeira Marques Dionísio**, Professora Auxiliar da Universidade de Évora.

To António, Ana, Margarida and Miguel.

Acknowledgments

I owe my deepest gratitude to my thesis supervisors Esmeralda Ramalho, João Andrade e Silva and Fernanda Peixe for their guidance, constructive criticism and support during my PhD. This thesis would not have been possible without the support of *Instituto Nacional de Estatística* and *Agência para a Energia*, who made possible the access to data at the premises of *Instituto Nacional de Estatística*.

I would also like to thank the comments from the participants of the 2016 Doctoral Consortium at *Universidade de Évora*. I am also indebted to Sonia Gouveia, who always had time to proofread early manuscripts, and to my colleagues at *Instituto Nacional de Estatística*, who encouraged me during these years.

Abstract

This thesis investigates the degree to which energy efficiency, as it is assessed by Energy Performance Certificates (EPCs), is reflected in residential property prices in Portugal. Its results are based on the analysis of a comprehensive dataset containing information of around 256 thousand residential property sales carried out from 2009 to 2013, a period largely characterized by depressed market conditions. This is the first large-scale study for a southern European country in this area of research. For the first time in this context, the impact of energy efficiency is analyzed along the distribution of residential property prices, using the unconditional quantile regression framework. The findings disclose a 13% sales premium for most energy efficient apartments (i.e. those bearing an A or B EPC rate) and a 5 to 6% market price premium for houses. However, quantile regression results show that the value attached to energy efficiency is not always positive across the distribution of prices. In particular, houses located at or below the 0.2th price quantile display clear energy efficiency price discounts. The use of different energy efficiency scales and cross-country comparisons support the view that energy efficiency price premiums are higher in the Portuguese residential market than in northern European markets. These results contribute to a more comprehensive understanding of the impact of energy efficiency on the real estate market and provide important messages to all political decision-makers interested in improving energy efficiency standards in Portugal.

Keywords: *Portugal, energy efficiency, residential property market, hedonic price models, cross-country comparisons, quantile regression*

JEL classification: C21, C52, C55, Q41, R21, R30

É A EFICIÊNCIA ENERGÉTICA REFLETIDA NOS PREÇOS DOS IMÓVEIS RESIDENCIAIS EM PORTUGAL?

Uma investigação baseada em funções de preços hedónicas e na análise de regressão por quantis

Resumo

Esta tese investiga em que medida a eficiência energética, tal como é avaliada pelos Certificados de Desempenho Energético (CDE), é refletida nos preços dos imóveis residenciais em Portugal. Os resultados obtidos baseiam-se na análise de um conjunto exaustivo de dados com informação sobre cerca de 256 mil vendas de imóveis realizadas entre 2009 e 2013, um período predominantemente caracterizado pela recessão. Este é o primeiro estudo de larga escala realizado para um país do sul da Europa nesta área de investigação. Pela primeira vez neste contexto, o impacto da eficiência energética é analisado ao longo da distribuição dos preços das habitações através do método da regressão por quantis incondicionais. Os resultados revelam um prémio na venda de 13% para os apartamentos mais eficientes em termos energéticos (i.e., aqueles com CDE A ou B), e de 5 a 6% para as moradias. No entanto, a análise de regressão por quantis mostra que o valor associado à eficiência energética nem sempre é positivo ao longo da distribuição dos preços. Em particular, as moradias situadas abaixo do vigésimo percentil mostram claros descontos associados à maior eficiência energética. A utilização de diferentes escalas energéticas e a comparações entre países apoia a ideia de que os prémios associados à eficiência energética são maiores no mercado português do que em mercados do norte da Europa. Estes resultados contribuem para um conhecimento mais amplo do impacto da eficiência energética no mercado imobiliário e fornecem importantes mensagens a todos os decisores políticos interessados em melhorar os padrões de eficiência energética em Portugal.

Palavras-chave: *Portugal, eficiência energética, mercado residencial, modelo de preços hedónicos, comparação entre países, regressão por quantis*

Classificação JEL: C21, C52, C55, Q41, R21, R30

Acronyms

ADENE	*Agência para a Energia*
AT	*Autoridade Tributária e Aduaneira*
CO_2	Carbon Dioxide
CQR	Conditional Quantile Regression
EPBD	Energy Performance of Buildings Directive
EPC	Energy Performance Certificate
EU	European Union
IMI	*Imposto Municipal sobre Imóveis*
IMT	*Imposto Municipal sobre a Transmissão Onerosa de Imóveis*
INE	*Instituto Nacional de Estatística*
LAD	Least Absolute Deviation
LM	Lagrange-Multiplier
OLS	Ordinary Least Squares
UQR	Unconditional Quantile Regression
VIF	Variance Inflation Factor

Contents

List of Tables

List of Figures

Chapter 1 Introduction

In the European Union (EU), buildings account for nearly 40 percent of global energy use and the residential sector is responsible for the production of around 11 percent of total global Carbon dioxide (CO_2) emissions from fuel combustion (Directive 2010/31/EU; International Energy Agency [IEA], 2017: 61). Given the relevance of the residential sector in the total building stock, the implementation of policies aimed at increasing its energy performance is regarded as one of the most effective ways to reduce CO_2 emissions and mitigate a country's dependency on energy. In Portugal, the importance of residential buildings is also high, as they are responsible for 17 percent of the country's total energy use and for 27 percent of the electricity consumed in the country (*Agência para a Energia* [ADENE], 2015: 10, 19).

A possible way of achieving higher energy efficiency in housing is through the introduction of building codes with better energy saving requirements (see, *inter alia*, Novan et al., 2017). However, while the implementation of more and better energy saving building specifications impacts directly on the supply of (newly built) dwellings, prospective buyers may still lack information on energy performance of residential units. In the absence of this information, a price premium signal is not given to most energy efficiency properties, and less than desired investment in energy performance is achieved in residential property markets. This outcome is referred in the literature as an efficiency gap (Brown, 2001: 1198), which can be addressed through the application of policy instruments. One of the most well-known instruments consists in the implementation of energy efficiency labelling schemes, which aim at influencing buying, selling and investment decisions with the disclosure of information regarding the energy performance of properties.

Energy labelling has been applied in Europe for many years, with household appliances providing one of the earliest examples (Directive 1992/75/EEC). For buildings, energy labels were first implemented in Denmark in the 1990s (Jensen et al., 2016). More recently, with the introduction of the Energy Performance of Buildings Directive (EPBD) in 2002, which was later recasted into the Directive 2010/31/EU, Member States of the EU were required to develop and implement an Energy Performance Certificate (EPC) system. In Portugal, the EPC scheme was first implemented in June 2007 as a consequence of the partial transposition of the EPBD into national law by a decree-law *(Decreto-Lei n.º* 78/2006).

Notwithstanding the potential advantages associated with the introduction of an energy efficiency label, the key question regarding its usefulness revolves around the degree to which energy efficiency improvements are capitalized into residential property transaction prices. While the existence of energy efficiency market price premiums is in line with an anticipation of energy expenditure savings from more efficient properties, it should be noted that, due to the existence of market failures and barriers, this relationship may well be insignificant or take up the form of a market price premium or a price discount. Thus, a quantitative valuation of the relationship between energy efficiency and residential property transaction prices is extremely important for an assessment of energy efficiency labels.

This thesis addresses the question of knowing whether energy efficiency, as measured by the EPC label, has any impact on transaction prices of residential properties in Portugal [1]. This country constitutes an excellent case study for a good number of reasons. First, the compliance and public awareness of the label has always been high, with 90 percent of all building completions and transactions already being done in 2010 with an issued EPC (Buildings Performance Institute Europe [BPIE], 2010: 59). This result stems not only from the mandatory status of energy efficiency certification, but also from its effective promotion by *Agência para a Energia* (ADENE), the national supervision body responsible for the implementation and administration of the European EPC system in Portugal [2]. As a consequence, selection problems associated with the generation of data by a non-mandatory system, such as the one that was implemented in the Netherlands (Brounen and Kok, 2011), are avoided when the Portuguese case is analyzed. Second, Portugal was one of the first countries to establish a national database in which the information generated by the EPC system was registered, analyzed and stored. This database, when combined with residential property transactions data, gives rise to a rich source of information for research purposes.

[1] The focus is on the sales residential market, which is much more important than the rentals residential market. According to the 2011 Census, in Portugal more than 72 percent of all family dwellings were not rented (*Instituto Nacional de Estatística* [INE], 2012).

[2] This agency is designated as ADENE even in texts written in English. As such, this acronym will be used throughout this work.

Finally, Portugal represents an interesting case study in its own right since the overwhelming literature in this area refers to northern European countries, which have climatic idiosyncrasies different from those located in the south of Europe (e.g., Fuerst et al., 2016; Högberg, 2013).

The investigation of the relationship between the EPC label and residential property prices in Portugal was carried out through the econometric estimation of the hedonic price model (Rosen, 1974), which has been used as the workhorse in this area of research. The hedonic model is based on the assumption that the value of a good, such as housing, is ultimately determined by the bundle of its quality attributes. Central to the hedonic price model is the existence of a functional relationship between prices and attributes, which could be estimated econometrically. Under this framework, the coefficient of the hedonic function associated with the variable measuring energy performance reveals important information on how the market rewards energy efficiency levels. Fesselmeyer (2018) and Ayala et al. (2016) are two recent examples of the application of this methodology for the residential property market.

This research is based on a comprehensive dataset with more than 256 thousand residential property sales carried out from 2009 to 2013. It is the result of the combination of information taken from ADENE and the *Autoridade Tributária e Aduaneira* (AT), the Portuguese Tax and Customs Authority [3], on two taxes: the *Imposto Municipal sobre a Transmissão Onerosa de Imóveis* (IMT) [4], and the *Imposto Municipal sobre Imóveis* (IMI) [5]. The tax sources cover the population of residential property sales since it is not possible to carry out a transaction without a proof of payment of the IMT. The IMT and IMI data are currently employed in the compilation of the residential and commercial property price indexes for Portugal (INE, 2017a; 2017b). A subset of the data available in the IMT and IMI records was also used in an empirical application to produce hedonic price indexes (Ramalho et al., 2017).

This thesis makes use of one of the largest datasets ever used in this research context and constitutes the first large-scale study on the impact of energy efficiency on residential

[3] The Portuguese Tax and Customs Authority is generically identified by AT, its abbreviated form which will be used in the text whenever there is a need to mention it.

[4] The real estate transfer tax is designated as *Imposto Municipal sobre a Transmissão Onerosa de Imóveis* or simply as IMT. Following the same approach that was used for AT and ADENE, the Portuguese abbreviated expression will be used throughout the text to designate this tax.

[5] The local property tax is designated as *Imposto Municipal sobre Imóveis*. Its abridged name IMI will be used in the text whenever the tax needs to be identified in the text.

property transactions in a southern European country. Although some evidence exists for meridional markets (e.g., Ayala et al., 2016), it is based on small samples and on proxy transaction prices. It is also the first study that investigates the impact of increased energy efficiency across the distribution of residential property prices. Overall, the findings confirm that energy efficiency is positively rewarded in the Portuguese residential sales market. However, the results also show that the EPC label impacts differently across the price distribution, throughout time and according to dwelling categories, with apartments yielding higher price premiums than houses. Given the importance of the former dwelling category in the Portuguese housing stock [6], the findings not only contribute to the growing literature on the effect of energy labels on market prices, but also provide new and important policy messages to all interested in enhancing energy efficiency in Portugal.

This thesis is organized as follows. Chapter 2 presents the hedonic price model as a measurement instrument for the assessment of the impact of energy efficiency on market residential property prices and addresses the association between energy efficiency and property prices. A literature review of 21 studies conducted in this area of research from 2008 to 2018 is presented in this chapter. Chapter 3 describes the sources, variables and information available in the dataset gathered for this research. Chapter 4 presents the application of the hedonic price models and provides the Ordinary Least Squares (OLS) results of the effect of energy efficiency in property prices in Portugal. Chapter 5 assesses, through 12 different experiments, the coherence and sensitivity of the energy efficiency estimates to modifications in the estimation context provided in Chapter 4. These include changes in sample size, the use of error-prone variables, and the comparison across different measurement scales and country results. Chapter 6 focuses on the question, which is investigated through quantile regression techniques, of how energy efficiency and other dwelling attributes vary across the residential property price distribution. Finally, the last Chapter summarizes the research findings and draws directions for future research in this area.

[6] The Portuguese housing stock amounts to 5,859,540 classic residential dwellings (INE, 2012). Of these, 52 percent refer to residential single family (detached, semi-detached and row) houses (author's own calculations based on Census data).

Chapter 2 Measuring the impact of energy efficiency on residential property prices

2.1. Introduction

This chapter overviews the hedonic price model as a measurement tool to investigate the impact of energy efficiency on residential property prices. It presents the results of a comparative literature review of a total of 21 studies that apply the hedonic price model in this research context. This comparative exercise, which covers the most important papers and reports produced from 2008 to 2018, offered the research directions that were followed in this thesis.

While the idea that housing markets attribute a price premium to energy efficiency is consistent with an anticipation of future energy savings, it remains an empirical issue to investigate whether and to what extent markets capitalize this expected outcome in practice. However, as it will be seen, the relationship between these two variables is far from being straightforward. In practice, due to factors such as the anticipation of higher future costs in maintaining (more expensive) energy efficiency technology, price premiums can be reduced or even take the form of price discounts. This literature review reflects this mixed outcome, with the majority of the studies pointing out to the existence of a market price premium and some other unveiling price discounts or an indifference to energy efficiency as a variable to explain residential property prices.

This chapter is divided into three sections. Section 2.2 points out the main links associating energy efficiency and property prices. Section 2.3 presents the hedonic price model, which has been used as the main instrument to measure the impact of energy efficiency on real estate market prices, and provides the results of the comprehensive literature review. Finally, a summary is given in section 2.4.

2.2. Links between energy efficiency and residential property prices

Achieving energy efficiency gains is generally perceived as providing several benefits, from which participants in the real estate market may profit. Perhaps the most obvious stems from the fact that, with increased efficiency, homeowners may benefit from lower utility bills (Dinan and Miranowski, 1989). Conversely, particularly in markets where energy efficiency standards are perceived as high, extra efficiency gains can be regarded not as a benefit, but as imposing high additional technological maintenance costs (Yoshida and Sugiura, 2015). In general, the idea underlying the relationship between energy efficiency and housing prices is anchored in the notion that markets are able to internalize (in market prices) the benefits associated with lower energy consumption patterns. Following this reasoning, more efficient properties would have higher future energy saving benefits and should, for this reason, have a market premium when transacted on the market.

Unfortunately, the relationship between energy efficiency and prices is far from straightforward. Due to the existence of market failures, less than desired investment in energy performance often surfaces on the market, a situation that is described in the literature as an efficiency gap (Brown, 2001). Since energy efficiency is an attribute that can only be experienced after purchase (e.g. through the comparison of utility bills), buyers tend to concentrate on more immediate and tangible dwelling characteristics and will not be willing to provide a price premium to any potential future benefits from a dwellings' greater energy efficiency performance. Homeowners, on the other hand, will not have an incentive to invest in energy efficiency, if markets fail to translate them into higher property prices. More recently, behavioral failures, such as low energy literacy rates and households' energy expenditure awareness, also started to receive attention from researchers (Brounen et al., 2013). For a summary of informational and behavioral failures see Ramos et al. (2015a).

From a theoretical point of view, it is possible to identify and estimate all potential factors linking energy efficiency and residential property prices. However, some factors, such as reputation or image effects (e.g., those stemming from owning an eco-certified property) are, due to its intangible nature, extremely hard to be measured and it is only possible to estimate their combined or aggregated effect. Table 2.1 summarizes the most important factors that are associated with energy efficiency and that have a potential impact on residential property prices.

Table 2.1: Most important factors linking energy efficiency and property prices

Individual factors	Likely impact on residential property prices
Capitalization of future lower utility bills	+
Perceived future maintenance cost increase	-
Reputation/halo effects	+
Energy literacy and other behavioral effects	+/-

As individual factors can have conflicting signs, the relationship between property prices and energy efficiency is far from simple and may well be insignificant or take up the form of a market price premium or discount. Under this framework, it is of paramount importance to measure the way markets value energy efficiency and investigate the degree to which policy instruments, such as energy efficiency labels, contribute to close efficiency gaps.

2.3. Assessing the impact through the use of the hedonic price model

2.3.1. The hedonic price model as a measurement tool

The hedonic price model (Rosen, 1974) rests on the assumption that the value of a good such as housing is ultimately determined by the bundle of its energy efficiency and other quality attributes. On the basis of this premise, which is referred to in the literature as the *hedonic hypothesis* (see, *inter alia*, Triplett, 2006: 91), price differences amongst similar varieties of a good, such as housing, can be explained by the different (quantities of) quality attributes found in each one of them. Central to the hedonic price model is the idea of the existence of a functional relationship between prices and attributes that, for the present research context, can be expressed in the following manner:

$$p_{i,t}^{*} = f_t\left(E_{i,t}^{*}, x_{i,t;1}^{*}, \dots x_{i,t;k}^{*}, u_{i,t}\right). \tag{2.1}$$

In (2.1), p and E stand for the price and energy efficiency characteristics and p^* and E^* represent some transformation of p and E, respectively. Moreover, the $x_{i,t;j}^{*}$, $j = 1, \dots, k$, corresponds to the remaining housing attributes, with the *i* and *t* subscripts identifying the dwelling and the time it was transacted, the $*$ signals the fact that the variable could have been transformed and u is a term representing additional random factors, which are not measured by the k+1 attributes included in the functional form. In the housing context, typical examples of x^* are the location of the dwelling (Kiel and Zabel, 2008), its area or floor space (Colwell, 1993) and age of the residential structure (Goodman and Thibodeau, 1995). As noted in the literature (see, *inter alia*, Malpezzi, 2003; Cropper el al., 1988), theory sheds little light on the selection of the appropriate functional form of (2.1). In the absence of such guidance, the derivation of the hedonic function tends to be essentially seen as an empirical matter, which should be guided with the help of statistical tests and, where possible, economic and engineering considerations.

It is possible to use econometric techniques to estimate the hedonic price function (2.1), where the coefficient associated with energy efficiency can be interpret as the implicit (shadow) price for that characteristic. In this context, the parameter estimate is read as the additional (or partial) effect of adding that covariate to the model, when all other explanatory variables were already accounted for. Applying the chain rule to calculate the first derivative of p^* with respect to the energy efficiency attribute, one obtains the partial effect of E on p^*in the following way:

$$\frac{\partial p^*}{\partial E} = \underbrace{\frac{\partial p^*}{\partial E^*}}_{\beta_E} . \frac{\partial E^*}{\partial E}, \tag{2.2}$$

where β_E corresponds to the coefficient obtained from the hedonic model, which is associated with the variable measuring energy efficiency. In a linear additive model, the partial effects are constant and equal to β_E only when $p^* = p$ and $\beta_E^* = \beta_E$. Except for some unusual transformations of E, such as its reciprocal where dE^*/dE is negative, β_E determines the direction of the impact of energy efficiency on p^*. Market price premiums (discounts) are thus associated with statistically significant positive (negative) partial effects for higher levels of energy efficiency. Due to its frequent utilization in the literature dealing with the impact of energy efficiency on dwelling prices, the case in which p^* assumes a logarithmic transformation of p assumes particular relevance. In this situation, the relative effect of E on dwelling prices can be calculated by $\left[exp\left(\hat{\beta}_E\right) - 1\right]$ and the percentage change is obtained by multiplying this expression by 100 (Halvorsen and Palmquist, 1980). However, for small values of $\hat{\beta}_E$, which typically occur when changes in E are of an infinitesimal nature (e.g., when this variable assumes a continuous form), it is possible to use the coefficient result as a reasonable approximation of the relative effect on dwelling prices (Megerdichian, 2018).

The econometrics associated with the estimation of hedonic price functions has been covered in several textbooks. A classical reference is Bernt (1991), which tackles, among other issues, the effects of omitting relevant covariates from the regression model, heteroskedasticity and other model specification issues [7]. More recent reviews are provided in the fifth and sixth chapters of Triplett's (2006) handbook on hedonic price indexes and in Ramalho and Ramalho (2010).

2.3.2. The roots and applications of the hedonic price model

The origins of the hedonic price method can be traced back to the pioneering work carried out by agricultural economists, of which Waugh (1928), who analyzed the impact of quality factors on the formation of fresh vegetables prices on the Boston wholesale market, constitutes one of the earliest published references (Colwell and Dilmore, 1999). The term *hedonic* was coined in Court (1939), which is considered to be the first published paper to use hedonic regressions in the compilation of price indexes (Stapleford, 2011) [8]. Although with some notable exceptions, of which Stone's (1956) report on the compilation of price and quantity price indexes for national accounts is an example, the use of the hedonic price model remained largely unnoticed until the beginning of the 1960s, when Griliches (1961) revived the interest in the method and prompted its application in a vast body of empirical work, which can be broadly divided into two main areas [9].

The first area, in which the present research can be included, addresses the estimation of the hedonic function and the measurement of the impact of a product's attribute (or group of attributes) on its price. This may include not only an investigation about the statistical

[7] It is interesting to note that, in addition to model specification problems, data issues, in particular those caused by the existence of measurement errors, a topic which has been overlooked in the subsequence literature on hedonics, are not ruled out from this early reference textbook. See, on this particular point, Bernt (1991: 129) and the note that identifies Amel and Bernt (1986), and the work cited in it, as first references in this area.

[8] Although firmly rooted in the literature, this designation is seen as to be a bit of a misnomer since hedonic coefficients are a reduced form solution, which do not generally provide information on the structure of the demand side alone but on the intersection of the demand and supply curves.

[9] Goodman (1998) and Stapleford (2011) provide tentative explanations as to why there was so little follow-up to Court's (1939) work. While the former focuses on practical issues (e.g., rudimentary state of the art in computing, data collection and coding activities), the latter attributes the lack of interest to the fact that the paper was written at a time were the economics profession was essentially dominated by a knowledge-based (or expert) rather than a market-based view of the relationship between price and quality.

significance of the impact, but also the estimation of intervals for regression coefficients and carrying out consistency checks on its sign and dimension relative to other studies. There are almost an uncountable number of applications of this type of analysis, ranging from the estimation of wage regressions (Montgomery el at., 1992) to the identification of wine price-determining factors (e.g., Combris et al., 1997). For housing, the use of the hedonic price model has a long tradition. Nelson (1982), who summarizes nine studies estimating the relationship between traffic noise and property values, is an early example. A more recent illustration of this approach is given by Stanley et al. (2016), who estimate the effect of increased energy efficiency on the value of housing. Chin and Chau (2003) and Malpezzi (2003) are two excellent reviews of the application of the hedonic price model to housing.

The adjustment of observed prices from changes in quality attributes in the construction of quality-adjusted price indexes is the second area in which the hedonic method is widely used. This application stems from the idea that virtually all empirical applications of the hedonic function can be conceived as an index number problem (Triplett, 2006). Oaxaca (1973), who applies hedonic functions to estimate male-female wage discrimination in the labour market, provides a notorious example in this area. Further applications include the use of the hedonic methodology in the construction of deflators for national accounts aggregates (Bover and Izquierdo, 2003) and in the compilation of price indexes for goods as different as cars (Santos and Coimbra, 1995) and paintings (Collins et al., 2009). A description of the methods available for the construction of price indexes using the hedonic method is available in Tripplet (2006). Specific reviews for residential housing are available in Hill (2013) and in Eurostat (2013).

In a more theoretical fashion, following the seminal work of Rosen (1974), a strand of the literature has been concerned with the use of hedonic models as a tool for the identification of supply and demand functions for product characteristics. Examples include Epple (1987), Arguea and Hsiao (1993), Ekeland et al. (2002), Bajari and Benkard (2005) and, more recently, Kuminoff and Pope (2012). Starting from the premise that hedonic models are essentially underidentified (i.e., consumer preferences or technology parameters are generally not directly obtained from hedonic implicit prices), this body of literature attempts to find the particular market conditions under which it is possible to estimate, from the information taken from hedonic regressions, supply and demand functions for product characteristics. Some of the special cases under which it is possible to derive a product's supply and demand functions are provided in Bernt (1991: 130). Since this thesis is focused on the assessment of the impact of energy efficiency on residential property transaction prices (rather than on the derivation of demand or supply curves for energy efficiency), no further explanations will be made on this topic.

2.3.3. Survey of empirical evidence

The first papers investigating the extent to which markets signal energy efficiency in housing prices were published after the energy crisis of the 1970s in the 1980s. This literature suffered from a number of limitations, the most important being the use of very small, highly localised samples, which were typically taken from subsidized or non-market environments (e.g., Laquatra, 1986), the lack of characteristics available for regression analysis and difficulties in finding a proper energy efficiency measure, which was usually built upon information taken from proxy variables, such as energy utility bills (e.g., Dinan and Miranowski, 1989). A comprehensive critical review of the earlier literature applying the hedonic method in this research context is available in Laquatra et al. (2002). Due to these shortcomings, any general conclusions drawn from this early body of research have to be considered with care. However, it can be said that it provides some evidence that improvements in energy efficiency were, at least to a certain extent, capitalized into property prices.

The last decades have witnessed a renewed interest in this research topic. This has been driven, at least in part, by the introduction of energy label schemes in housing markets, something which had already motivated the investigation of the impact of energy standards on the prices of certain goods such as household appliances, computers and electronic equipment (Greening et al., 1997; Howarth et al., 2000). However, despite the importance of the residential sector, commercial office buildings have attracted a much larger number of academic studies than its residential market counterpart. The reasons for this situation may rest, on the one hand, in the characteristics of the markets (e.g., the housing market is usually more regulated and, as such, more subject to inefficiencies, which makes it more problematic to estimate the effects of energy labels on prices) and, on the other, in difficulties in obtaining relevant data for research purposes (e.g., the supply and the demand of residential properties is far more atomised than in the office market segment). Among the papers that investigate the effect of green labels on commercial property prices, Eichholtz et al. (2010; 2013) and

Fuerst and McAllister (2011a; 2011b) stand out as references in this area. On the whole, these studies provide qualitative evidence supporting the idea that green office buildings display a premium when compared to conventional office space.

For the residential housing market, Soriano (2008) constitutes one of the first examples of the renewed research interest on the relationship between energy efficiency attributes and residential house prices. Based on 2005 and 2006 samples of around 2,400 and 2,700 sales of detached houses sold in Canberra, this work reports the existence, for both years, of a significant positive relationship between energy efficiency, as measured by the Australian Energy Efficiency Rating system, and transaction prices. Using the Dutch experience in the implementation of a non-mandatory EPC label scheme in 2008, Brounen and Kok (2011) find that greener properties obtain over less energy efficient homes, a 3.7% price premium in sales prices. An interesting feature of this paper has to do with the use of the Heckman (1979) two-step procedure to tackle sample selection bias arising from systematic differences in characteristics of certified and non-certified homes (of the 177 thousand transactions available for the study, only 32 thousand had an EPC label). Hyland et al. (2013) applied the same estimation procedure for the Irish residential market. Although based on list and not on transaction prices [10], the conclusions of these authors reinforce those of Brounen and Kok's (2011) in regard to the existence of a price premium associated with more energy efficient homes. Hyland el al. (2013) find evidence supporting the idea that the sales market segment rewards energy efficiency more strongly than the rental market segment and that the price premium attached to energy efficiency was stronger when the market was depressed. Interestingly, Fuerst et al. (2015) suggest that the price effects of energy efficiency are lower for detached and semi-detached houses than for flats or terraced dwellings and that they are not constant across English regions. For detached houses specifically, the paper reports no significant price effects, something which the authors explain by the influence of a small and atypical portion of the sample of detached dwellings located in rural areas.

Using a dataset with single-family house sales in an area of Florida, USA, Bruegge et al. (2016) find that the price premium associated with the transaction of new properties disappears in the resale market. While these results reinforce the idea that energy efficiency price premiums may vary across time, space and market segments, they also point out

[10] List or asking prices are derived at earlier phases of the buying and selling process and are, for this reason, generally different from final transaction prices.

directions for future research. In particular, given the characteristics of the housing stock in Portugal, it would be interesting to investigate whether energy efficiency is rewarded differently across the houses' and apartments' and new and existing market segments. Moreover, it would be important to investigate if energy efficiency price effects varied from 2009 to 2013, a period in which there was a strong contraction of the housing market. Högberg (2013) explores a sample of around 1,100 family house sales in the Stockholm area with the main aim of seeing whether energy efficiency is reflected in market transaction prices. An interesting feature of this paper is the inclusion of explanatory dummy variables describing categories of energy performance improvement recommendations found in EPC data. While this study corroborates the idea that the market signals improved energy efficiency with price premiums, the author finds out that buyers require price discounts from suggested energy improvements, thus revealing that sellers may have an incentive to improve energy standards prior to sale. Unfortunately, the dataset used in our research does not include information on improvement recommendations and it will not be possible to investigate this issue for the Portuguese market.

While the omission of relevant covariates appears as the elephant in the room problem in hedonic regression applications, it is interesting to note that the literature dealing with the estimation of the impact of energy efficiency on residential transaction prices does not provide many examples of omitted variable bias and little, if any, evidence on its direction and size. Stanley et al. (2016) highlight the importance to include controls for the age of the dwelling, since their omission could lead to biased energy efficiency estimates. For the Helsinki second-hand apartments market, Fuerst et al. (2016) obtain a significant price premium of 3.4 percent for the three most energy efficient EPC classes, which reduces to 1.3 percent after location, neighborhood and maintenance costs attributes are added to the regression specification. The results point to the existence of omitted variable bias when location and other quality attributes are missed out from the hedonic models. Further examples include the omission of neighborhood covariates from the regression (Fuerst et al., 2016) and the non-inclusion of hard-to-measure factors, such as buyer's predisposition to environmental ideology (Brounen and Kok, 2011) and developer's reputation (Zheng et al., 2012). All in all, it can be said that, in addition to omitted variable bias, the sensitivity of energy efficiency partial effect estimates to data and model specification issues has not been much subject to empirical research. An example of an overlooked researched topic has to do with the fact that some studies use appraisals or list prices as a proxy for transaction prices (e.g., Ayala et al., 2016; Hyland et al., 2013). Although it can be argued, at least on theoretical grounds, that the use of list prices may not bias partial effect estimates, this assumption is seldom, if ever, tested. In practice, it may be possible that the differences between true and proxy prices are correlated with some of the regressions' covariates (e.g., dwelling dimension or energy performance) thus leading to biased coefficient estimates. In this context, part of the differences found in the literature on the price impact of energy efficiency may be attributable to the use of different price measurements. As the database used in the present thesis includes appraisal values and transactions prices, it is possible to look into this issue and investigate whether or not the use of mismeasured or surrogate price variables could bias energy efficiency estimated impact.

Additional evidence on the existence of price premiums is provided in Kahn and Kok (2014), for the Californian housing markets. By the same token, Cajias and Piazolo (2013) quantify the energy premium in the German residential market measured on a continuous scale based on 2,630 building observations from 2008 until 2010. According to these authors, a 1 percent increase in energy conservation produces a 0.45 percent increase in market value and 0.08 percent increase in rent prices. Fuerst et al. (2016) found no significant price premium for energy efficiency classes below the comparison rate. According to these authors, this is justified with the buyer's predisposition to pay more for more energy efficiency, which are not interested with below the average energy efficient properties (green clientele effect). Fuerst et al. (2015) constitute the largest study made to date in this field. Based on a sample of more than 333 thousand dwellings that were sold at least twice in the period from 1995 to 2012, the results of this study suggest the existence of a positive relationship between the EPC rating and prices per square meter of transacted dwellings in England.

Evidence on southern European countries is scarce and based on small samples and using proxy transaction prices. Ayala et al. (2016), finds a 10.3 percent price premium for the properties displaying one of the three top EPC ratings (A, B or C). However, the conclusions of this paper are based on a small sample of 1,507 observations and use stated housing prices as a surrogate of transaction prices. Ramos et al. (2015b) presents the first study for the Portuguese residential market. Based on a sample of 21 thousand dwelling adverts taken in March 2015 from an internet real estate portal, the authors find a 6.1 percent price premium for properties with an A, B or C EPC rating. This result is essentially driven by most efficient dwelling adds, with the most efficient rate showing a 0.404 coefficient and the second one 0.141 (Ramos et al., 2015b: 33) [11]. Interestingly, the results suggest that the Portuguese residential market rewards energy efficiency more than in northern European countries (e.g., Brounen and Kok, 2011; Fuerst et al., 2015). A growing awareness of the benefits associated with energy efficiency, which could have been triggered by increasing electricity prices, and an effective divulgation of the EPC label in Portugal are referred in the paper as possible explanations for this result. However, part of the explanation may also rest on the use of list (instead of transaction) prices, as these are typically set above real transaction values and its use may cause bias to estimated coefficients [12].

Recent research has not only been confined to the American and European contexts. Deng et al. (2012), Deng and Wu (2014) and, more recently, Fesselmeyer (2018), present evidence on the impact of the voluntary Green Mark label on residential property prices in Singapore. Covering different time length periods, the papers support the idea that green properties receive market price premiums over non-labelled dwellings. Zheng et al. (2012) build a Green Index using information about residential project attributes found on the internet and investigate whether the emerging real estate market for environmentally friendly properties in China sell for a price premium. After controlling for the time in which apartments are sold (i.e., before and after green certification), Fesselmeyer (2018) obtains a 3 percent premium the properties that were sold after the attribution of an energy certificate. This suggests that the market attaches value to energy efficiency labels, a finding that is also supported in Jensen et al. (2016), where a higher premium is found in Denmark for the period starting after the display of the EPC was made obligatory in advertisements of property sales.

Although the majority of studies investigating a link between energy efficiency and transaction prices suggest that the former is associated with price premiums, empirical evidence on this matter is not unanimous. In particular, Zheng et al. (2012) report the

[11] These coefficients represent price premiums of 49.8 percent and 15.1 percent for the A and B EPC rates over the hold-off category considered in the study (i.e., the D rate).

[12] Ramos et al. (2015b: 13) report list prices that are 35 percent higher than those obtained from mortgage loan processes (a statistics provided by INE). Unfortunately, this comparison is not correct, since it compares two different means. While bank appraisals are based on the geometric mean, the paper's average of the sample of advertised dwellings is compiled using an arithmetic mean. However, it is possible on page 14 to obtain the mean of the logarithm of sampled advertised prices, which allow us to calculate a geometric average simply by taking the anti-logarithm of the reported average. With averages of the same type, list prices are 5.2 percent above bank appraisals, which, in turn, are slightly higher than the geometric average of transaction prices for March 2015.

existence of price discounts at the resale stage in China and Cerin et al. (2014) were not able to provide evidence on the full capitalization of energy efficiency gains in property prices in Sweden. For Belfast, Northern Ireland, Davis et al. (2015) finds a negative relationship between energy efficiency and price of dwellings in a model where the natural logarithm of transaction price per square meter is the dependent variable. However, this situation could be rooted in model misspecification problems caused by the absence of important variables in the data. Although the study concentrates on a small and dense area, the models do not control for location and researchers had no access to variables providing insight into the condition of the property. In addition, Yoshida and Sugiura (2015) suggest that green condominiums in Tokyo are associated with price discounts. Stressing the idea that more energy efficient homes could be perceived as requiring higher maintenance costs, the authors attribute their finding to the possible low marginal benefit in investing in costly new technology in a market where energy efficiency levels are already high. However, a more recent paper based on a bigger sample (Fuerst and Shimizu, 2016), shows that green condominiums in the Tokyo metropolitan area command a small but statistically significant price premium for both list and transaction prices (higher in the former than in the latter case). By taking into account buyer characteristics in the hedonic model, this study finds evidence supporting the idea that condominiums purchased by wealthier buyers are associated with higher price premiums. Further evidence regarding this point could be taken from studying the way energy efficiency is valued across the price distribution of dwellings through the use of quantile regression (since most expensive homes are essentially bought by the households with the highest income and the less expensive homes by the poorest). However, this area has not been explored.

The investigation of the relationship between energy efficiency and the value of residential properties has typically been conducted using cross sectional (e.g., Högberg, 2013) or pooled cross sections of data (e.g., Cerin et al., 2014). Panel data has not generally been used in this context where, for the residential property market, it is difficult to obtain large samples containing repeated measurements for the same residential units across time. Moreover, given the importance of the coefficient estimate associated with energy efficiency characteristics, it is a bit puzzling to see the little attention that the literature has given to the use of tests that could help identify omitted variable bias and other hedonic specification problems (e.g., Ramsey, 1969) [13]. However, this extends to all other areas in which the hedonic regression is used and is explained, at least partly, by the work of Cropper et al. (1988) in which the good performance of parsimonious functional forms (e.g., semi-log) in presence of omitted variables is highlighted. Kuminoff et al. (2010) mark this paper as the most influential in the subsequent empirical hedonic price model literature and suggest that the reason why researchers are willing to provide results based on untested assumptions about the shape of the hedonic function rests on the idea that simpler models would hedge against the risk of omitted variable bias. Perhaps as a result of this, the overwhelming majority of the hedonic studies estimating energy efficiency partial effects typically apply a semi-log function where the variable measuring energy efficiency enters the model as a dummy or as a set of dummy variables. Another interrelated issue has to do with the extraction of the energy efficiency price effect from estimated hedonic coefficient figures. With some notable exceptions (e.g., Yoshida and Sugiura, 2015), when energy efficiency is measured as a dummy variable, most studies do not apply either Kennedy's (1981) or van Garderen and Shah's (2002) estimators for the percentage impact change of energy efficiency on prices. However, this situation is understandable since, in practice, these estimators do not yield striking differences among them or even when they are compared to simpler estimators (see Section 4.3.2). A summary of the key features of a total of 21 empirical studies on the impact of energy labels on residential property prices is presented in Table 2.2.

[13] Or, in the words of Malpezzi (2003: 83), "...it is somewhat surprising that the literature applying formal specification tests, [...], is modest, since specification is such an issue in hedonic analysis...".

Table 2.2: Summary of studies on the impact of energy efficiency on housing prices

Paper/study	**Geographical coverage**	**Energy label**	**Property type coverage and *type of property value***	**Sample size**	**Estimation method**	**Main findings**
(1) Fesselmeyer (2018)	Singapore	Green Mark	New apartments, *Transaction prices*	119,826 obs. (2000-2016)	OLS	Buyers attach a value to certification. Prices per square meter for new apartments receive a price premium of around 3% after green certification is attributed.
(2) Ayala et al. (2016)	Spain	EPC	All dwelling types, *Respondent's stated housing price*	1,507 obs. (2013)	OLS	Stated price of properties rated A, B or C obtain a 10.3% price premium in relation to less energy efficient properties. This price premium decreases to 5.5% when properties rated D are added to the three top rates.
(3) Bruegge et al. (2016)	USA, Florida, Gainesville area	Energy Star	Single-family residential houses, *Transaction prices*	5,031 to 5,528 obs. (January 1998 to August 2009)	OLS	Most energy efficient new single-family houses receive a 1.2% price premium over less energy efficient properties. Price premium vanishes in the resale (i.e., for existing properties) market. First-sale premium may have been eroded by the adoption of successively tighter building codes.
(4) Fuerst et al. (2016)	Finland, Helsinki metropolitan area	EPC	Second-hand apartments, *Transaction prices*	6,194 obs. (2009-2012)	OLS	Transaction prices of properties rated A, B or C display a 3.4% price premium over D-rated properties, which reduces to 1.3% after controlling more carefully for locational and neighboring attributes.
(5) Fuerst and Shimizu (2016)	Japan, Tokyo	Green Building Program	Condominiums, *List and transaction prices*	23,922 (2004-2011)	OLS	Evidence supporting the idea that the price premium of green condominiums is substantially higher for list prices than for transactions prices. First study controlling for buyer characteristics.
(6) Jensen et al. (2016)	Denmark	EPC	Detached single-family houses, *Transaction prices*	117,483 obs. (January 2007 to September 2012)	OLS	Price premiums of properties rated A, B or C increased from 2.4% to 10.6% after the display of energy performance rating has been made obligatory in advertisements of property sales.

Paper/study	Geographical coverage	Energy label	Property type coverage and *type of property value*	Sample size	Estimation method	Main findings
(7) Stanley et al. (2016)	Ireland, Dublin area	EPC	All dwelling types, *List prices*	2,792 obs. (2009 – June 2014)	OLS	Evidence that energy efficiency has a significant and positive relationship with list prices. The omission of age from the regression leads to downward biased energy efficiency estimates.
(8) Davis et al. (2015)	Northern Ireland, Belfast	EPC	All dwelling types, *List prices*	3,797	OLS	The baseline model shows that a one point increase in energy performance rating increases sales prices per £420. However, the authors also find the existence of a negative relationship between energy performance and the log-price per square meter.
(9) Fuerst et al. (2015)	England	EPC	All dwelling types, *Transaction prices* of dwellings sold at least twice	333,095 obs. (1995-2011)	OLS	Transaction prices per square meter of properties rated A or B obtain a 5% premium over D-rated properties. Results highlight the importance to control for property type in regression analysis.
(10) Ramos et al. (2015b)	Portugal	EPC	All dwelling types, *List prices*	21,170 obs. (March 2015)	Heckman (1979) two-stage estimation	Dwellings ranked A, B or C have a 6.1% list price premium over D-rated advertised properties. Dwellings advertised with an A show a 49.8% price premium over the D category. Higher price premiums than those estimated for northern European markets.
(11) Yoshida and Sugiura (2015)	Japan, Tokyo	Green Building Program	Condominiums, *Transaction prices*	11,933 obs. (2002-2009)	OLS	Transaction prices of newly-built green condominiums are approximately 10% lower than their non-green counterparts. Price discounts are interpreted as evidence on the capitalization of higher future maintenance costs.
(12) Cerin et al. (2014)	Sweden, large cities and commuting areas	EPC	All dwelling types, *Transaction prices*	67,559 obs. (2009-2010)	OLS	Suggest that, while it is not possible to conclude that energy performance is fully capitalized by the market as a whole, specific market segments may exhibit price premiums.
(13) Deng and Wu (2014)	Singapore	Green Mark	Apartments, *Transaction prices*	35,730 obs. (2000- June 2010)	OLS	Evidence that the market rewards in a different way Green Mark properties at presale and resale stages, with developers obtaining a lower premium (4%) than at resale stage (10%).
(14) Kahn and Kok (2014)	USA, California	LEED, Energy Star and GreenPoint	Single-family homes, *Transaction prices*	314,759 obs. (2007-2012)	OLS	When compared with non-certified homes, the most conservative estimate signals a 2% price premium in green homes.

Paper/study	Geographical coverage	Energy label	Property type coverage and *type of property value*	Sample size	Estimation method	Main findings
(15) Cajias and Piazolo (2013)	Germany, Mainly the southern part of the country	EPC	Residential buildings, *Market value of properties*	2,630 obs. (2008-2010)	OLS	A 1% reduction in energy consumption increases rents by 0.08% and is associated with a 0.45% increase in the market value of residential properties.
(16) Hyland et al. (2013)	Ireland	EPC	All dwelling types, *List prices*	11,060 in second stage equation (2008–March 2012)	Heckman (1979) two-stage estimation	A-rated properties obtain, over D-rated properties, a 2% premium in advertised rentals and a 9% premium in offer sales prices. Premium is higher when market conditions are worse.
(17) Högberg (2013)	Sweden, Stockholm municipality	EPC	Single-family homes, *Transaction prices*	1,073 obs. (2009)	OLS	A 1% reduction in standard energy consumption yields a premium of 0.04% in sale prices. Suggestions for improvements associated with price discounts.
(18) Deng et al. (2012)	Singapore	Green Mark	All dwelling types, *Transaction prices*	36,512 obs. (2000-June 2010)	OLS with the inclusion of a two stage price equation variant	Green properties receive a 4% price premium over non-rated properties in sale prices. (The two stage price equation suggests a higher premium: 15%.)
(19) Zheng et al. (2012)	China, Beijing	-	Dwellings in residential projects, *Average transaction prices*	3,171 obs. (2003-2008)	OLS	Green properties sell for a price premium. Residential projects constructed according to a Google-based green index. Green units sell for a price premium at the presale stage, but with a price discount at the resale stage.
(20) Brounen and Kok (2011)	Netherlands	EPC	All dwelling types, *Transaction prices*	31,993 obs. in second-stage equation (2008-August 2009)	Heckman (1979) two-stage estimation	Properties rated with A, B or C labels get, over less efficient properties, a 3.7% price premium in sale prices.
(21) Soriano (2008)	Australia, Camberra	Energy Efficiency Rating	Detached houses, *Transaction prices*	5,104 obs. (2005 and 2006)	Feasible Generalized Least Squares	A one unit increase in the star rating scale is associated with a 1.2% price premium in sale prices (2005 data). Using 2006 data, the premium is 1.9%.

It is possible to take from Table 2.2 some interesting points. Firstly, despite the existence of some papers with a national coverage (Brounen and Kok, 2011; Hyland et al., 2013), it is worthwhile to note that the majority of the studies are based on some specific area or market segment. Secondly, the design of the empirical exercises is dominated by applications of the OLS estimator on cross sectional data. Finally, the results of the 21 studies emphasize the heterogeneity of energy efficiency partial effect estimates, which vary considerably in size, statistical significance and even in sign. Among other factors, this heterogeneity is explained by the use of different measurement scales (e.g., continuous or discrete, choice of hold-out and reference energy efficiency categories) and energy label schemes, which do not necessarily measure efficiency in the same manner. These differences make the comparison among the different studies' coefficients a very difficult task (see Chapter 5 for more on this issue).

Based on a sample of 30 published and unpublished working papers covering the residential and non-residential (sales and rental) markets, Ankamah-Yeboah and Rehdanz (2014) apply meta-regression analysis (Stanley and Jarrell, 1989) to identify the factors explaining the variation in reported energy efficiency partial effects. The paper contains complementary information to the literature survey summarized in Table 2.2. Firstly, the heterogeneity of energy efficient partial effect results, which is evident from Table 2.2, is again emphasized, ranging from a minimum price discount of 10 percent to a maximum price premium of 40 percent. Secondly, meta-regression analysis yields a global average price premium of 7.6 percent. Thirdly, meta-regression results support the view that energy efficiency labels are more valued in Europe than in the USA, Australia and other geographical areas. Fourthly, the sales market is found to provide higher price premium than rental markets. Finally, the study underscores age of certification since introduction (policy takes time to be understood and adopted) and the mandatory or non-mandatory status of energy labels (the effects of voluntary certification tend to disappear through time) as relevant factors explaining cross-country variations in the coefficients associated with energy efficiency.

2.4. Summary

This chapter has highlighted the complex relationship between energy efficiency and residential property transaction prices. In practice, given the conflicting directions of the factors influencing the association between these two variables, its magnitude may well be insignificant or take up the form of either a market price premium or a price discount. In spite of its limitations, the hedonic price model constitutes the most powerful measurement tool capable of providing an assessment of the sign and likely dimension of the relation between energy efficiency and market prices. Although the majority of the studies applying this model point to the existence of energy efficiency price premiums, this conclusion is far from being universal.

The literature review has highlighted important research directions to be followed in this thesis. First, it would be important to investigate the degree to which different market segments reward energy efficiency. In particular, it would be particularly interesting to investigate the existence of significant differences between the apartments' and houses' market segments. Differences among new and existing dwellings should also be an interesting topic, as most of the literature concentrates on mature market countries, where newly built dwellings are a rarity (e.g., the Netherlands, England) [14]. Second, another direction has to do with the fact that energy efficiency may be valued differently across the dwelling price distribution. In particular, the quantile regression method (Koenker and Basset, 1978; Firpo et al., 2009) has never been applied in this research context, something that could provide additional evidence about the way different groups value energy efficiency (e.g., how high-income buyers, which buy most expensive properties, value energy efficiency in relation to low-income groups). Third, another direction focuses on changes across time and space (e.g., see if better or worse market selling conditions have an influence on the impact of energy efficiency on dwelling prices). Finally, it would be interesting to use the data available for this thesis to test the robustness of results to different estimation scenarios such as the use of proxy price variables.

[14] In the dataset used for regression analysis, a residential property is considered new if it had never been used for residential purposes. Therefore, it should be noted that, while older (in age) properties are expected not to be classified as new, there could be some cases of new properties with some years of existence (e.g., newly built homes that, due to the existence of a depressed market, remained on the market before they were first sold).

Chapter 3 Data and econometric approach

3.1. Introduction

The dataset used in this thesis, which covers the transactions of residential properties over a five year period (2009 to 2013), was derived combining three different sources. Administrative records, taken from the AT for transfer and property tax purposes, account for two of the three sources and provide the bulk of the data. Moreover, information taken from ADENE on issued EPCs constitutes the third data source. The result of the combination of these sources was a dataset with more than 256 thousand transactions prices and quality attributes of transacted residential properties, which stand out as one of the biggest ever used to investigate the relationship between energy efficiency and market transaction prices.

The objective of this chapter is twofold. First, it aims at providing an account of the information that was used for the estimation of the relationship between residential prices and dwelling energy efficiency attributes. This involves, not only the description of the process that was used to derive the employed dataset, but also the statistical analysis of the variables contained in it. The second goal of this chapter is to describe the approach that was followed to model and estimate the price-energy efficiency relationship. The quality of the available information was investigated through the help of an exploratory data analysis. Summary statistics draw interesting insights, particularly in relation to the usefulness of analyzing the impact of energy efficiency for apartments and houses separately. The approach that was chosen to specify hedonic price models underscores the idea that, in the absence of strong theoretical considerations as to the correct derivation of the hedonic function, the estimation of the relationship between energy efficiency and residential transaction prices is essentially a data-driven empirical issue, which should be guided by statistical tests and, where possible, economic and engineering considerations.

This chapter is divided into four main sections. Section 3.2 reviews the sources and matching process used to create the dataset on which the estimation of hedonic price models was based. In doing this, its dimension, coverage and richness will be emphasized. Section 3.3 provides the results of the exploratory data analysis, which was carried out prior to the specification and estimation of the relationship between energy performance and dwelling prices. This section is concerned not only with the quality of the data but also with the issue of investigating whether the available information provides plausible clues for the development of hedonic price models. Section 3.4 presents the strategy that was followed in the specification and estimation of the hedonic models. Finally, Section 3.5 provides some concluding remarks.

3.2. Data sources

3.2.1. Transaction prices and dwelling characteristics

In order to produce results, hedonic regression analysis needs reliable information on transaction prices and transacted dwelling characteristics. Prices and information characterizing property transactions were obtained from IMT records. The IMT is a tax levied on property transfers, which is calculated based on the value of the transaction (declared in the deed of sale) or on the fiscal appraisal value of the property, depending on which is higher. Under this system, which takes into account declared and appraised values, the incentive to under-declare is reduced and it is generally accepted that this tax produces information on transaction values that are the same or close to real transaction values [15]. This data source covers the population of transactions since a proof of payment of the property transference tax has to be shown by the buyer before a sale takes place. Moreover, because it represents a non-negligible cost to the buyer, the IMT is typically paid just a few days before or on the same day the property is transacted. Therefore, the date of IMT payment constitutes a trustworthy indicator of the transaction moment. An example of the IMT form is available in Appendix I. The purchase price is obtained from fields 45 and 61, which provide information on the value of the transaction. The location of the transacted property is identified through codes 20 (municipality) to 23 (individual fraction). When taken together, these fields allow for the construction of a property cadastral register identification number,

[15] A comparison between all pairs of appraisals and transaction values yielded a correlation of 0.77. Moreover, appraisals were on average 24 percent lower than transaction prices. Appraisals can be seen as minimum threshold (market) values, which play a role as a deterrence mechanism for the under-declaration of transaction prices.

which is also available in other data sources [16]. Although the IMT has been in place since the end of 2003 (Oliveira et al., 2012), the digital record and storage of IMT data covering all fiscal acts has only been guaranteed by the AT from the beginning of 2009 onwards. Accordingly, the data that were made available for analysis contained few transactions from before 2009, a fact that constrained the choice of the time period of the built dataset.

The appraisal values of properties and information on the characteristics of each dwelling is taken from IMI records. The IMI is a municipal tax which was introduced at the end of 2003 and that is levied on the value of the dwelling (Oliveira et al., 2012). Property values are appraised by means of a formula defined in the Portuguese Property Tax Code, which was introduced by the *Decreto-Lei n.º 287/2003*. This formula covers the most important price-determining factors and generates values that reflect the way the market discriminates properties. The formula is based on a minimum square meter price, which is multiplied by the property's area and by a number of coefficients that identify the use of the dwelling, the location of the dwelling, the quality and comfort of the dwelling and its age. With the natural exception of dwelling area, all other factors are defined administratively, with some being set to a specific value and other allowed to vary within minimum and maximum pre-defined values. The minimum square meter price, which reflects land prices and average construction cost estimates for the whole country, represents an example of the former values used in the formula. An example of a variable used in the appraisal formula that is allowed to vary is given by the location coefficient, which represents the price homogeneity of geographical areas. For the definition of these areas, local and regional appraisal experts have to take into account the quantity and quality of accessibilities (e.g., roads, highways), the access to social amenities (e.g., schools, public services, commerce), the offer of public transportation and the areas with high commercial values (*Direção-Geral dos Impostos* [DGI], 2011: 20). The location coefficient varies from a minimum of 0.35, attributed to rural and sparsely inhabited areas, to a maximum of 3.5, associated with areas showing high market values. The definition of the maximum and minimum values and the delimitation of homogeneous zones are revised under the proposal of the *Comissão Nacional de Avaliação de Prédios Urbanos*, a national committee for urban buildings assessment [17]. Under this framework, municipalities have the

[16] Not all information generated by this tax form is available for research purposes. For instance, field 13, which identifies the buyer, is not available for confidentiality reasons. The data sources used in this thesis do not identify dwelling buyers, sellers and owners.

[17] The delimitation of homogeneous zones has been subject to several modifications since the introduction of the IMI tax in 2003. The first occurred in 2004 with the introduction of two ministerial orders (*Portarias n.º 982/2004* and *1426/2004)*. A second change occurred in 2006 (*Portaria n.º 1022/2006*). Further minor ad-

power to choose, from given maximum and minimum values, the specific coefficients to be applied within municipality boundaries (Oliveira et al., 2012: 2). A good description of each one of the factors used in the appraisal formula is available in DGI (2011).

An example of the IMI tax form is available in Appendix I. Specifically of interest are fields 62 and 63, where taxpayers are asked to identify the elements of quality and comfort that characterize the dwelling and fields 55 to 60, which provide information on the area of the property. As with the IMT, it is possible to derive a property cadastral register identification number from IMI fields. This can be extracted from the information provided in Part III of this tax form. As a principle, the form should be submitted by taxpayers to tax authorities whenever there is the need to reassess the value of the property (e.g., due to a change in its size). The update of the IMI information can also be done by tax authorities (e.g., at the time of mass appraisal exercises). After the IMI has been submitted, the value of the property is assessed by tax authorities. Following this appraisal, taxpayers can contest the value of the assessment and ask for the revaluation of the property. The data used in this work refers only to IMI information in which the appraisal value is considered as final. Data quality, especially the accuracy of the variables associated with the factors included in the appraisal formula, gain from this verification mechanism and from the interaction between tax authorities and tax payers.

The quality of the information has also benefitted from the mass appraisal exercise that was carried out by AT in 2012 and 2013. This action was done to comply with the recommendation set out in the Memorandum of Understanding, which was signed between the Portuguese Government, the European Commission, the International Monetary Fund and the European Central Bank, to ensure that the taxable value of the housing stock was close to its market value (European Commission, 2011: 88). The IMI data almost covers the entire stock of residential properties in Portugal, with the only time period unavailable for analysis corresponding to the IMI fiscal acts recorded from the end of 2003 to the end of 2004, a period to which it was impossible to obtain records from the AT.

---

justments were introduced in 2007 and 2009 with *Portaria n.º* 1305/2007 and 1119/2009. More recently, a more comprehensive change was carried out in 2015 (*Portaria n.º* 420-A/2015). This last revision has no impact on the data used in this thesis.

3.2.2. Energy efficiency of residential units

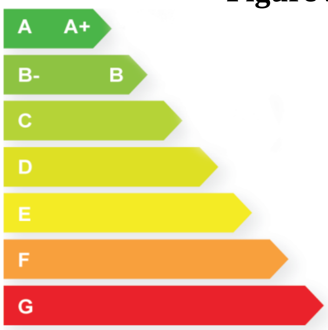
Information on the energy performance of dwellings is obtained from the data records of ADENE. Portugal was one of the first countries in Europe to implement a sound system of practical enforcement of the EPC label. With the introduction of the EPBD in 2002, which was later recasted into the Directive 2010/31/EU, Member States of the EU were required to establish a system of certification of the energy performance of buildings. In Portugal, both the EPBD directive and its recasted version were transposed into the national law by *Decretos-Lei n.º 78/2006* and *118/2013*, respectively. Contrary to other countries (e.g., the Netherlands), the EPC system was not adopted on a voluntary basis, with the mandatory status of energy certificates being applicable for all residential property transactions since the beginning of 2009 (*Portaria n.º 461/2007*) and for all advertised and rented properties since December 2013 (*Decreto-Lei n.º 118/2013*). In addition, ADENE was also able to implement effective promotional campaigns of the EPC scheme. As a result, compliance and public awareness of the EPC label have always been high in Portugal. Energy certificates can only be issued by qualified experts (usually architects or engineers), which are required to have a minimum of five years of relevant working experience. The training of these experts is supervised by ADENE, which also controls national exams qualifying experts to issue certificates. Regular control checks are also carried out by ADENE to identify any eventual professional malpractice or misconduct (BPIE, 2010: 60).

According to the EPC certification system that was adopted in Portugal for the period covered by the data, the energy performance of a property was expressed in a nine-level scale, which ranges from A^+, the most efficient level, to G, the least efficient level. Attributed scales translate a ratio between annual primary energy needs and a reference limit value, which is defined for a property of similar characteristics. Energy ratings are based on calculated rather than actually measured consumption patterns. The former measures do not vary with the size of the household or the energy consumption lifestyle of households [18]. Certified experts calculate energy needs based on three key factors: quality of the building (e.g., walls, insulation), quality of existent equipment (in terms of energy consumption) and the use of renewable energies. A property obtaining a ratio score of less or equal to 0.25, meaning that it consumes 25 percent or less energy than the stipulated reference consumption, is attributed the A^+ rate. Conversely, all properties with a ratio greater than 3, which indicates energy

[18] Not all countries have adopted a methodology based on calculated ratings. France, for instance, uses a rating system that is a mixture between calculated and measured energy needs (BPIE, 2010: 47).

needs that are more than the triple of those of reference, are attributed a G rate. Under the Portuguese EPC system, all properties with a building permit from before July 2007 are dubbed as new [19]. For these properties, the minimum energy efficiency requirement is the B^- rate. For the remaining dwellings there is no minimum energy efficiency requirement. In practice, this has improved energy efficiency standards for all newly built dwellings. Moreover, it raises the expectation that the gains from the implementation of improved energy efficiency levels should be higher and more noticeable for older dwellings, where less efficient labels are more abundant. Figure 3.1 presents the adopted label scale. For ease of interpretation, the figure also shows (on its right hand side) the percentage of the reference consumptions which are associated with each one of the energy levels.

Figure 3.1: EPC label scale



A-label: 0% - 50%

B-label: 51% - 100%

C-label: 101% - 150%

D-label: 151% - 200%

E-label: 201 % - 250%

F label: 251% - 300%

G label: More than 300%

With the transposition of the recasted EPBD, the energy certificate and the methodology underlying the measurement of energy ratings was revised, leading to the elimination of the G level. In practice, this revision introduces little noise in the data available for regression analysis since it only influences the transactions of residential properties that have been certified in the last month of the available data series (i.e., December 2013). In Appendix II it is possible to see the first pages of the EPCs that were used until November 2013 and from December 2013 onwards. As with the IMT and IMI, it is also possible to build a property cadastral register identification number using information on the location of the property from EPC data (see top part of the EPC form, fields *Localidade* to *Fracção Autónoma*).

[19] The definition used in the EPC system for new differs from the concept that is used throughout this text, where a residential property is considered new if it had never been used for residential purposes.

3.2.3. Data matching process

The first step of the construction of the dataset used to analyze the relationship between energy efficiency and residential property prices involved the matching of the information coming from the transfer and property tax records. This was done using the property cadastral register identification number, an unique identification key which is associated with each dwelling, as the matching variable between IMT and IMI records. Moreover, to restrict the scope of the analysis to the relevant set of residential property market transactions, several basic restrictions were applied to the raw IMT and IMI data. First, rustic and agricultural land, dwellings providing commercial services, parking facilities and plots of land for later construction were excluded from this first data matching step. Second, where they were possible to be identified, non-harms length transactions, such as inherited properties, were also excluded from the database. Third, after a preliminary analysis of the information, all transactions with a value equal or lower than 20,000 € or equal or higher than 3,000,000 € were filtered out from the dataset. The reasoning underlying this restriction rests on the fact that the hedonic regression analysis should not be distorted by non-market prices, which take the form of abnormally low or abnormally high transaction values.

Fourth, transactions that were carried out under a *Permuta* agreement, which involve the exchange of two (or more) properties, were also ruled out from the analysis. A typical *Permuta* transaction involves the acquisition of two residential properties, the first one purchased for a zero price and the other one for a discount. The discount reflects, on the one hand, the difference between the market value of the two properties and, on the other, differences in buyer and seller bargaining power. Since *Permuta* prices do not mirror market forces and the information provided by tax authorities does not allow the identification of the counterpart residential unit (or units) that are involved in this particular transaction, it was opted to exclude them from the analysis. Fifth, all dwellings that had less than 35 square meters of gross floor area have also been excluded from the data. The use of 35 square meters as a lower cut-off boundary was taken from the Portuguese Building Code, where minimum construction specifications are defined for new homes [20]. Although small for residential purposes, it is not impossible to find in the Portuguese housing stock dwellings with less than 35 square meters of gross floor area (an example would be a very old dwelling, for which

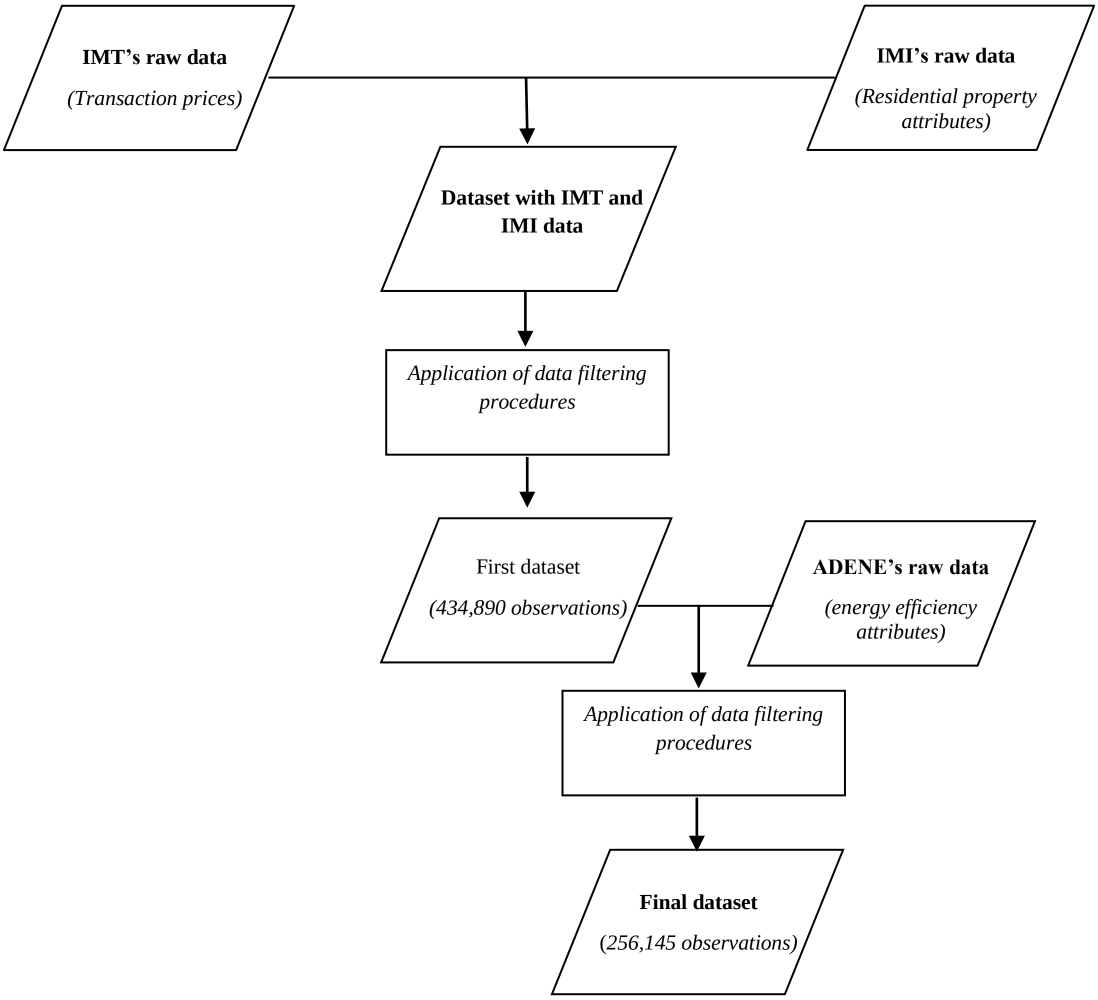
[20] The present building code is based on *Decreto-Lei n.º 38382/1951* and on its posterior revisions, of which the one introduced by the *Decreto-Lei n.º 50/2008* was the last. Minimum gross areas are described in article 67 of the code.

modern building specifications were not applied). However, these cases are expected to be the exception to the rule and its existence might be caused simply by data entry errors. Finally, all transactions of properties that had a negative number of complete years at transaction date were, in addition, excluded from the database. Although rare, this might happen in cases were dwellings are transacted before their completion date. The reasoning for this exclusion rests on the idea that these transactions are, in many occasions, carried out with an investment (and not a residential) purpose in mind. In these situations, properties are typically bought with a discount before completion, later to be resold with a profit. The application of the abovementioned restrictions resulted in a first dataset with a total of 434,890 transactions for the period starting in 2009 and ending in 2013.

The second and final step of the matching process refers to the inclusion of ADENE's variables in the dataset with IMT and IMI data. This was done using the same property cadastral register identification number that was used to merge IMT and IMI data. Due to the existence of incomplete (or inaccurate) information in ADENE's location variables, which are essential to construct the property register identification number, it was only possible to match 60 percent of all transactions. Moreover, a preliminary analysis of ADENE's information identified 4,832 transactions with zero annual energy needs, which, at the same time, displayed very inefficient energy performance labels. This contradiction, which stems from the fact that zero annual energy needs should be attributed only to completely efficient and sustainable buildings, is explained by the fact that the EPC system attributed, until almost the end of 2013, the zero score to derelict or dwellings in ruins. In fact, of the total number of observations falling into this category of properties, 98 percent are categorized by the EPC scheme as being in very bad conditions. As the inclusion of these transactions could have introduced additional noise in the estimation of the relationship between energy efficiency and transaction prices of residential properties, they were also excluded from the data used in regression analysis.

The end-product of this two-step matching process was the derivation of a unique dataset with information on transaction prices, energy performance and other dwelling characteristics of 256,145 residential property transactions in Portugal (59% of the transactions available at the end of the first stage). The matching of the three data sources and the analysis of the information available in them was carried out using the SAS software (SAS Institute Inc, 2015) [21]. The data matching process described in this section is illustrated in Figure 3.2.

Figure 3.2: Data flow chart



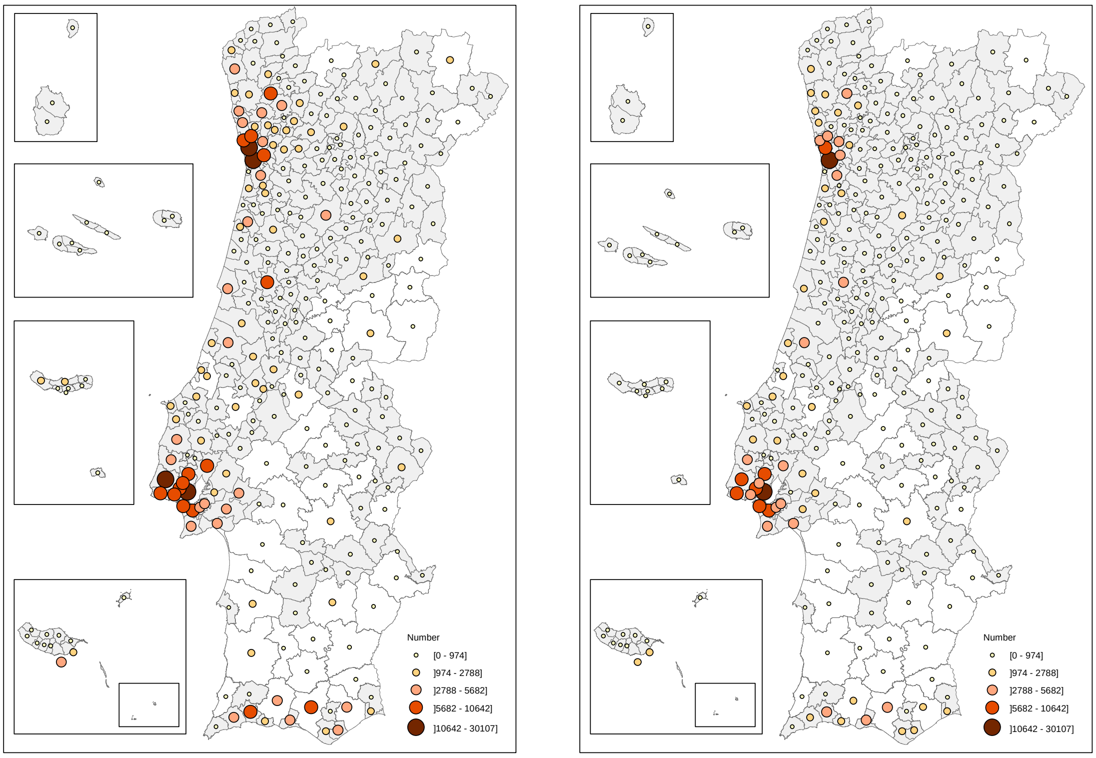
The characteristics included in the final dataset cover, in addition to the transaction price of the residential property, a wide range of continuous and categorical variables that include not only the location, age and structural attributes of sold properties, but also characteristics that are difficult to measure and that are usually absent from hedonic studies. Examples of the latter variables include, among other, measures for the availability of public goods and for the scenic and visual prominence of the location.

[21] Unless otherwise stated, all the statistical analysis presented in this thesis was carried out using this software.

3.3. Data analysis

The geographical representativeness of the data is illustrated in Figure 3.3, where the distribution of the number of sales in the dataset with the more than 430 thousand transactions (left panel) is compared with the geographical distribution of the number of transactions shown in the final dataset (right panel). The number of transactions is signaled with a circle at the municipality level. To provide a picture of the relative distribution of sales in the two datasets, the diameter of each circle (and the darkness of its color) is drawn as a positive monotonic function of the number of transactions in each municipality. The figures represent not only the mainland but also the *Açores* (three small squares at the top left part of the figures) and *Madeira* Islands (small square at the bottom left section of each figure).

Figure 3.3: Geographical distribution of transactions in initial and final datasets



The left and right panels of Figure 3.3 provide a similar picture, which suggests that the final dataset mirrors, in essence, the geographical relative importance of the real estate market in Portugal. In terms of sales importance, the same three areas emerge from the two panels. From north to south, the first one corresponds to *Porto's* and *Braga's* urban areas. The second one includes the capital *Lisboa* and its surrounding areas. Finally, the last one corresponds to the *Algarve*, whose importance in terms of sales is essentially driven by the transactions made in the southern coastal part of this region. The importance of some capitals of districts located in the middle of the country is also highlighted by the panels (e.g., *Viseu* and *Coimbra*). On the other side of the spectrum, the islands and the interior of the mainland show to be unimportant in terms of sales.

Table 3.1 provides the summary statistics for a group of selected variables available in the final dataset. In order to provide a richer description of the data, the table is presented in a form where, in addition to the totals, it is possible to obtain the descriptive statistics for existing apartments, new apartments, existing houses and new houses. The data reveals interesting differences amongst these market segments. As expected, new is generally more expensive than existing and houses are more expensive than apartments. For instance, while existing apartments data present an average value of nearly 97,000 €, new houses data provide a much higher average of approximately 180,000 €. In face of these figures, it is not surprising to see existing apartments to stand out as the most common property in the dataset (59% of all observations) and new houses as the less frequently purchased property type (5% of all transactions). Another interesting feature has to do with price dispersion, which is much higher for houses than for apartments (see the standard deviations in the table). This suggests that the former property type is more heterogeneous than the latter and that hedonic model specification may be more complex and difficult for houses than for apartments.

Table 3.1: Summary statistics of a group of selected variables

Variable description		All			Apartments						Houses					
					Existent			New			Existent			New		
		N	Mean	Stdev	N	Mean	Stdev	N	Mean	Stdev	N	Mean	Stdev	N	Mean	Stdev
Number of transactions	(#)	256,145	-	-	149,920	-	-	59,410	-	-	33,282	-	-	13,533	-	-
Transaction value	(€)	-	119,888	98,131	-	97,695	68,876	-	149,007	93,688	-	143,783	150,717	-	179,155	145,671
Gross floor area	(m^2)	-	110.6	50.7	-	96.0	34.0	-	113.1	38.0	-	148.3	79.7	-	168.6	67.4
Dependent floor area	(m^2)	-	31.1	39.1	-	18.6	21.1	-	36.1	26.3	-	60.2	66.7	-	75.4	64.4
Uncovered land area	(m^2)	-	78.2	375.0	-	2.9	15.8	-	4.9	21.6	-	441.0	797.3	-	415.6	788.0
Number of bedrooms	(#)	-	2.5	1.2	-	2.3	1.0	-	2.3	0.9	-	3.5	1.8	-	3.3	1.1
Age of property at transaction date	(years)	-	16.1	18.9	-	20.1	17.5	-	2.0	2.1	-	29.3	25.9	-	1.5	2.1
Percentage (%) of residential properties in each energy efficiency category																
Energy label A^+			0.7			0.2			2.0			0.5			0.9	
Energy label A			4.2			1.4			11.8			2.1			6.1	
Energy label B			20.2			15.9			39.3			7.3			15.7	
Energy label B^-			12.8			12.0			15.7			11.0			14.2	
Energy label C			36.3			46.8			22.8			18.5			22.2	
Energy label D			14.8			13.2			6.2			33.0			24.8	
Energy label E			8.6			9.5			1.8			16.0			9.9	
Energy label F			2.1			0.9			0.2			9.3			4.8	
Energy label G			0.5			0.1			0.0			2.4			1.4	
Properties completed after 2006			28.4			10.9			77.0			8.9			56.9	
Properties completed >1990 and <=2006			40.7			50.9			20.9			37.3			23.6	
Properties completed >1960 and <=1990			23.6			32.3			1.1			30.1			9.1	
Properties completed in or before 1960			7.3			5.9			1.0			23.7			10.4	
2009 transaction			23.5			21.7			27.9			22.3			27.2	
2010 transaction			28.5			27.6			31.5			26.4			29.5	
2011 transaction			19.4			19.6			19.0			18.8			19.6	
2012 transaction			14.2			15.0			11.8			15.7			13.0	
2013 transaction			14.4			16.0			9.8			16.8			10.7	
Improved or renewed property			5.4			1.5			8.9			4.5			35.9	

The variables reflecting the dimension of the residential unit are also in line with expected differences for the considered market segments. Gross floor area is higher for houses than for apartments and, interestingly, the data shows a clear difference between new and existing dwellings, with the latter property type being smaller than the former. For instance, while existing apartments have on average 96.0 square meters, new apartments have 113.1 square meters. An explanation for this may rest on a likely evolution in preferences, with older (and smaller) houses satisfying the needs of older household needs and newer properties essentially reflecting current needs for bigger areas. The dimension is also different for apartments and houses. This is reflected in the number of bedrooms, which is on average higher for houses (3-4) than for apartments (2-3). The statistics also reveal clear differences in uncovered land areas, with this variable being more important for houses than for apartments. While new houses had on average 415.6 square meters of uncovered area, its apartment counterpart had 4.9 square meters [22].

The summary statistics characterizing time (e.g., age and construction year) also reveal interesting information. For instance, while existing houses display an average age of 29 years, the set of transacted new houses show an average of two years [23]. As expected, the percentage of dwellings completed in or before 1990 is much higher for existing dwellings than for new dwellings. However, new houses show a relatively high percentage of properties that were completed in or before 1991 (19.5%). This is explained by major improvements and renovations, which account for 35.9 percent of the total of new house transactions. In situations where dwellings are renewed, the once before old property is put on the market as a new property and it is considered as such in the database. Finally, the mean values of the dummy variables signaling the year in which transactions occurred disclose 2010 and 2012 as the years with the highest and lowest number of sales, respectively. This is in accordance with the behavior of the residential sales market, which achieved its lowest point in 2012. The analysis of age-related variables reveals that the way time affects property prices may not necessarily be the same across the different market segments. For instance, given the smaller variability of the age variable for new dwellings, it may be argued that it may not be that

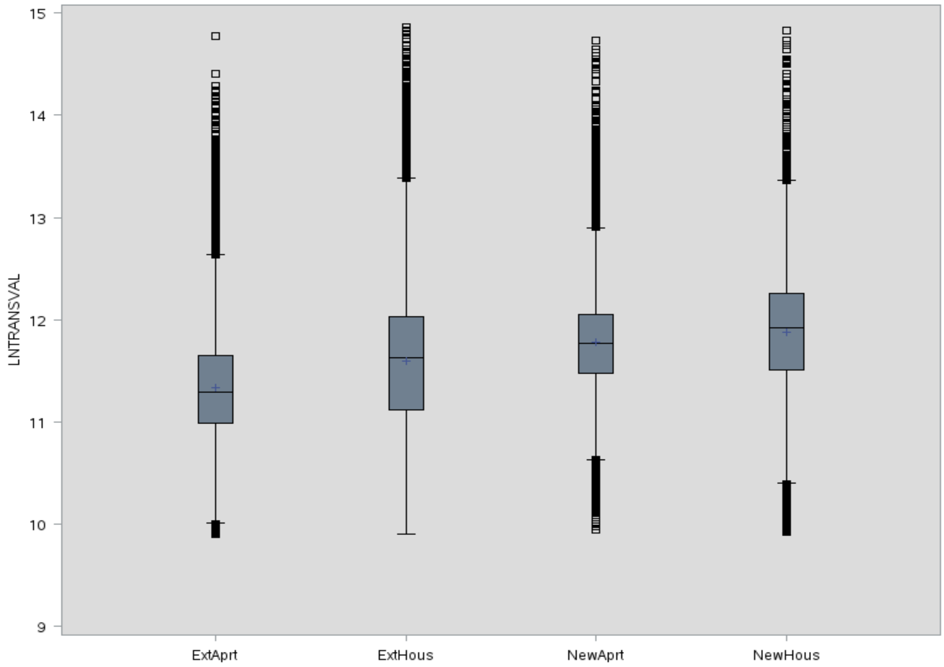
[22] An apartment may have uncovered floor area. An example would be an apartment located on a ground floor with a private courtyard.

[23] It should be noted that, while older (in age) properties are expected not to be classified as new, there could be cases of properties classified as new with some years of existence (e.g., newly built homes that, due to the existence of a depressed market, remained on the market before they were first sold). See footnote 19 where the concept of new that is used in this work is provided.

important to explain the formation of price than it is for existing dwellings. However, since renovated dwellings may display vintage effects, it may be necessary to keep in the hedonic model dealing with new dwellings variables providing information about the year in which they were completed.

The differences across the four market segments are reinforced by the visual inspection of the distribution of the logarithm of transactions prices for existing apartments (ExtAprt), existing houses (ExtHous), new apartments (NewAprt) and new houses (NewHous). The logarithm transformation of prices was used to induce some normality in the data. Figure 3.4 shows the box-and-whisker plots for the four market segments. The whiskers are drawn to the most extreme points in the group that lie within the fences. The upper (lower) fence is defined as the third quartile plus (minus) 1.5 times the interquartile range. Observations outside the fences are identified with a square and the plus sign represents the mean values.

Figure 3.4: Box-and-whisker plots of the natural logarithm of prices



The figure highlights the differences in the distributions in the four market segments and reinforces the analysis of the data that was done before. These differences are also visible in energy efficiency measures. As Table 3.1 shows, the most common rating in transactions data is C (36.3%). However, while for new apartments the most frequent rate is the B label (39.3%), for houses it is the less energy efficient D rate (25-33% of all observations).

Although higher for new properties, the percentage of A^+ or A rates is relatively low (4.9% of the total transactions). When grouped with the transactions of residential units bearing a B or B^- label, the percentage of transacted dwellings rises to 37.9 percent of total transactions. Based on these results, it was decided that, for modelling purposes, it would be better to group A and B properties in a single dummy signaling, in this way, all residential units having annual consumption energy needs that are the same of or lower than reference standard consumption values.

3.4. Modelling strategy

The use of the hedonic price model as a means to estimate the relationship between energy efficiency and residential property prices involves deciding on several important practical issues associated with the specification of the hedonic function. These include several critical issues, such as choosing a functional form, the way transaction prices are measured, which property features are included as regressors, and deciding if the same price function can be applied to all data or if, on the contrary, it needs to be tailored to different housing market segments (Ramalho and Ramalho, 2010). The approach that was followed to answer these questions was based, on the one hand, on a review of literature dealing with the specification of the hedonic model and, on the other hand, on possible indications taken from the exploratory analysis of the data available for regression analysis (see previous section).

As mentioned in Chapter 2, theory tells little about the way hedonic functions should be specified. In this context, researchers have attempted to draw conclusions from empirical studies, where the use of flexible models, such as those provided by the application of the Box and Cox (1964) procedure (e.g., Halvorsen and Pollakowski, 1981), or the semi-parametric and nonparametric estimators (e.g., Anglin and Gençay, 1996; Parmeter et al., 2007) are compared to simpler functional forms such as the linear, log-linear and log-log specifications. However, the empirical evidence stemming from these studies is mixed (see, *inter alia*, Cropper et al., 1998 and Kuminoff et al., 2010) and it is not possible to conclude about the superiority of the use of flexible functional forms in hedonic studies. Cassel and Mendelsohn (1985) specifically argue against the use of the Box-Cox methodology on two sensible grounds [24]. The first one has to do with the lack of precision in

[24] The paper presents four arguments. However, two of them are not applicable to the present research context.

parameter estimation, which is likely to occur when interaction terms and other variables are included in the model. The more variables are included in the model, the greater the correlation possibilities among them and, when those correlations involve unimportant (or the interaction of unimportant) terms, the more likely it is to have increased model flexibility at the expense of inaccurate estimates of individual coefficients [25]. The second argument in favor of the use of simpler functional forms has to do with the fact that the Box-Cox model may be seen as a too cumbersome specification, which is not particularly suitable for policy analysis purposes. Conversely, the computational simplicity and the straightforward and appealing interpretation of the coefficients of the linear, log-linear and log-log specifications are reasons that can justify its preference over more flexible, but more complex, models. Among these simpler specifications, the log-linear stands as the most used in hedonic housing studies (Malpezzi, 2003). This model uses the logarithm of price levels as the dependent variable, which helps reducing the heteroskedasticity that is intrinsic to housing prices. The hedonic price models used in this work follow the log-linear specification.

As suggested at the outset of this section, some consideration needs to be given as to the choice of the independent variables that should be included, at least on *a priori* grounds, in the specification of a hedonic price function. A summary containing a list of most important variables is available in the Handbook on Residential Property Price Indices (Eurostat, 2013: 25). According to this international manual, the most important price-determining characteristics of a dwelling are the area on which the dwelling structure is built, the area of the land, the location of the property, the age and type of the structure, the materials used in the construction of the property and attributes such as the number of bedrooms, the existence of a garage and the distance to amenities. A more comprehensive list of commonly used housing attributes in hedonic price models is available in surveys on housing, such as Chin and Chau (2003). The dataset that was built based on IMT, IMI and ADENE's records has variables covering all price-determining characteristics that were identified in these surveys. With all these data at hand, it is important to have a clear idea of the most influential dwelling characteristics, and to classify them into relevant groups of housing attributes. Table 3.2 provides the classification that was used to organize the huge

[25] Kuminoff et al. (2010: 153-4) provide evidence of a bias-variability trade-off associated with the adoption of more flexible functional forms (i.e., increasing the flexibility of the functional forms reduces the bias linked to the parameter estimate and simultaneously tends to increase its standard deviation).

amount of information available in the final dataset containing prices and characteristics of residential property transactions in Portugal.

Table 3.2: Most important dwelling attributes by group of variables

Variable group (describing)	Example of variables which are typically used to represent the variable group
(1) Dimension and other basic dwelling attributes	Floor areas, Land area, Number of rooms, Number of bathrooms, Type of dwelling (e.g., apartment or house); Presence of garages, swimming pools
(2) Durability attributes	Age of the residential property; Existence or inexistence of regular maintenance of the dwelling, major repairs or reconstructions
(3) Location attributes	Quantity and quality of accessibilities (e.g., roads, motorways); Access to social amenities (e.g., schools, public services, commerce); Offer of public transportation
(4) Energy efficiency attributes	Actual consumption of energy; Estimates of energy needs
(5) Transaction attributes	Purchaser's characteristics (e.g., bargaining skills), Type of deal involved (e.g., exchange or swap of properties)

In addition to the organization of the information, the use of variable groups was helpful as a working strategy, since it made it easier to identify if, for a given model, there was an important price determining attribute missing from the regression or (what was worse) from the dataset. The group of variables describing the location of residential properties provides an example of this strategy. As it is clear from Table 3.2, location encompasses several distinct levels, which range from the broader identification of the geographical area in which the property is, to the quality of the neighborhood of the transacted dwelling. Following this reasoning, special care was taken as to the inclusion of variables covering all these levels in the specification of the regression models used to capture the impact of energy efficiency on residential property prices.

In hedonic regression models, it is usual to apply the logarithm transformation to control variables since it induces normality in positively skewed variables such as prices or areas. However, in situations where the variables could take the null value, this transformation is not possible to be applied and a usual alternative is to use the square root transformation. The variables measuring dependent floor space and uncovered land area provide two examples for which it is not possible to apply this transformation. Contrary to gross floor areas, which are always greater than zero in residential properties, dependent and uncovered land areas of residential properties could be non-existent. In face of this situation, it was chosen to apply the square root transformation to all area variables.

The modelling approach took into consideration the question of whether it was possible to use a single hedonic model to all data or if, on the contrary, there was the need to developing different models for different market segments. Partly due to the lack of data, many hedonic studies on housing are focused only on a segment of the market (e.g., the housing market segment of the capital city of a country) and do not have to address this issue. In the present situation, the exploratory data analysis provides some strong evidence that apartments and houses are different products and should, for this reason, be specified separately. This issue was investigated further by means of a robust (to heteroskedasticity) version of the Chow (1960) structural break test (see Chapter 4). Although the specification and the estimates of energy efficiency partial effects on residential prices were subject to careful analysis and to a battery of overall specification and individual tests, it was also decided to subject the conclusions drawn from the use of the OLS estimator to a number of comparisons and sensitivity analysis. Following this line of reasoning, the soundness of OLS estimates was also assessed empirically through the use of several exercises simulating different estimation contexts and by means of a comparative study, which assesses the coherence of results across different energy efficiency scales (see Chapter 5).

3.5. Summary

This chapter describes the sources and process associated with the derivation of a dataset with information on transaction prices, energy performance and other dwelling characteristics of 256,145 residential property transactions for the 2009-2013 period. As it is possible to see from the literature review presented in the previous chapter, this is one of the biggest datasets used in this area of research. What is more important, the database reflects in an accurate way the population of transactions and key dwelling transaction features such as the transaction moment and transaction prices.

The exploratory data analysis provides evidence about the good quality of the variables available for regression analysis and pointed out directions for the specification of hedonic price models. Since the distribution of price, age, area and other important dwelling features are different for existing apartments, existing houses, new apartments and new houses, there could be some gains in modelling these four submarkets separately. The differences amongst them are also visible in attributed energy ratings. While for new apartments the most frequent rate is the B label, for houses it is the less energy efficient D rate that is the most frequent. The percentage of top energy rates is low, with less than 5 percent of the total transactions displaying an A^+ or A rate. However, when grouped with the residential units bearing a B or B^- label, it is possible to signal all properties having annual consumption energy needs that are the same of, or lower than, reference standard consumption values, and raise the percentage to 37.9. Based on these results, it was decided to group all A and B properties into a single dummy, which would be used as the key variable to assess the impact of energy efficiency on transaction prices in this thesis [26].

The last subsection of this chapter described the approach that was followed to establish the hedonic price model functional form. Following the strand of the literature on hedonic price models, it was based on the application of the log-linear specification and the OLS estimator. Overall, the chosen line of action highlights the idea that, in a context where theory sheds little light on model specification, the derivation of the hedonic function is essentially a data-driven process, which should be based on data considerations, statistical testing and, where possible, preliminary considerations about the relationship between prices and price-determining characteristics.

[26] Note though, that alternative measurement scales are considered in the robustness analysis of obtained results in subsection 5.5.

Chapter 4 Impact of energy efficiency on residential property prices in Portugal

4.1. Introduction

The overwhelming majority of the empirical studies assessing the impact of energy efficiency on residential transaction prices support the idea that energy efficiency is positively rewarded by the market (see Chapter 2). For Europe, the evidence provided on this issue has been essentially focused on the experience of northern European countries, where, due to existent climatic conditions, issues such as the thermal comfort of the properties and their energy efficiency characteristics are relatively present in the minds of all economic agents participating in the housing market. Unfortunately, for southern European countries, where the climate is generally milder and such issues are not so present, the empirical evidence on this matter is scarce, based on small samples and in information with limitations (e.g., absence of transaction prices). The work presented here helps to close this gap since it provides the first large-scale study on the impact of energy efficiency on residential property transaction prices in Portugal.

As mentioned earlier, the relationship between energy performance and property prices was investigated using the hedonic price model (Rosen, 1974). The sign, magnitude and statistical significance of the impact of energy efficiency on prices were analyzed with the help of a grid of six working hypothesis, which were developed based on the research directions drawn from the comprehensive literature survey presented in the previous chapter. The working hypotheses not only cover the key issue of seeing if energy efficiency has an influence on transaction prices but address also other interrelated issues such as knowing how energy efficiency partial effects vary through time and across dwelling categories. The first two working hypotheses address this last point and address the question of whether the same hedonic price model can be applied indifferently to all dwelling categories (or if, on the contrary, it is better to derive different models for different market segments).

After hedonic price models have been specified, estimation can take place and it is possible to address the main research question of this thesis, which asks whether or not energy efficiency has an impact on transaction prices. This issue is tackled by the third working hypothesis of this chapter. An interrelated issue has to do with knowing if the valuation of energy efficiency varies according to market conditions. Portugal constitutes a very interesting case study, as the period covered by the data broadly corresponds to the years in which the housing market has contracted due to the imposition of severe mortgage credit restrictions. This issue is covered by the fourth working hypothesis. Finally, the last two hypotheses have to do with knowing whether energy efficiency partial effects are invariant to supply side characteristics of purchased residential properties. In particular, the data allows investigating if construction requirements, which evolved from a building construction technology time period to another, and outstanding construction quality, have an impact on energy efficiency partial effect estimates.

One of the main problems associated with the application of the hedonic price model has to do with omitted variable bias. The richness of the data available for research allowed the inclusion of control variables in the hedonic models that potentially cover all price determining factors of a dwelling. In addition, the models that were used to investigate the above-mentioned grid of hypotheses were subject to (and had to pass a) battery of tests, which included the specification RESET test (Ramsey, 1969). As noted by several authors (e.g., Ramalho and Ramalho, 2010), this test is useful to detect possible model specification problems due to the omission of relevant covariates and was used as a safeguard against the use of biased estimators. All the tests were carried out using robust to the presence of heteroskedasticity statistics. In particular, the structural break (Chow, 1960), specification (Ramsey, 1969) and heteoskedasticity (Breusch and Pagan, 1979) tests were based on a procedure that uses Lagrange-Multiplier (LM) statistics developed by Wooldridge (1991).

The key findings of the research can be summarized in the following way. First, statistical tests and the analysis of the data strongly support a model specification approach that takes into account the specificities of existing apartments, new apartments, existing houses and new houses. Second, there is a clear indication of a significant price premium associated with most energy efficient property units. Third, this price premium is higher for apartments and, to a lesser extent, new properties than for houses and existing dwellings. Fourth, the effect of the EPC label on dwelling prices has grown in a systematic way from 2009 to 2013, a period predominantly characterized by recession. Fifth, the impact of energy efficiency on residential property prices is not the same for the properties built before 1960, with apartments signaling a price premium and houses a discount. Finally, although more evident for new than existing properties, dwellings with above than average quality thermal and insulation materials seem to be rewarded with an energy efficiency price premium by the market. These results add to the growing empirical evidence that support the idea that energy efficiency is positively rewarded by the housing market, and shed light on how energy efficiency price premiums vary across market segments and time. Given the relevant share of houses in the total dwelling stock in Portugal [27], the finding that apartments have a higher price premium than houses provides an important message to all those interested in policies to enhance energy efficiency levels in this country. Finally, this study contributes for the clarification of how particular supply side factors, such as building technology at time of construction and quality of building materials, impact on the way the residential property market values energy efficiency.

This chapter is structured as follows. Section 4.2 presents and explains the grid of hypotheses in which the empirical analysis and hedonic price models are anchored. Section 4.3 provides and discusses the results, which are based on the application of the OLS estimator using hedonic price models. Finally, Section 4.4 presents a summary of the research findings.

[27] For a quantification of the importance of houses in total dwelling stock, see footnote 6.

4.2. Working hypotheses

4.2.1. Market versus sub-market hedonic model specification

The estimation of hedonic price models requires careful consideration as to whether the same price function can be applied to different housing market segments or if it is better to derive sub-market specifications. While the estimation process may benefit from the use of all data into the same econometric model in terms of efficiency, it may also profit from the flexibility that is inherent to the use of different hedonic functions, as they could easily accommodate sub-market peculiarities, which would otherwise be almost impossible to handle in a single hedonic expression. The incorporation of area provides a good example of one of these peculiarities. While for houses, the area of free land may be considered an important characteristic, it is unimportant or inexistent for apartments. In many studies market versus sub-market model specification is not an issue since the data available refers to a single location (e.g., Högberg, 2013) or specific market segment (e.g., Yoshida and Sugiura, 2015). Partly due to data unavailability, this issue has not been much investigated. However, Malpezzi (2003: 82) notes in his survey on hedonic house price models that, when data allow for the testing of the existence of sub-market segmentation, the assumption concerning the constancy of the hedonic coefficients over different housing markets is usually rejected.

To investigate this issue, it was chosen to use the framework proposed by Chow (1960) for the detection of structural breaks. Two different market segmentations were analyzed, the first one exploring the separation between new and existing properties and the second one among apartments and houses. In this framework, a restricted model, which assumes parameter stability, was tested against an unrestricted model, where parameters were allowed to vary over the considered sub-markets. The restricted hedonic price model, which was used to test the existence of structural breaks, was formulated in the following way:

$$\ln(p) = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + u, \qquad (4.1)$$

where p is the transaction price, $\ln(p)$ its logarithm transformation, x_k^* refers to the k explanatory variable, β_k corresponds to the parameter of the hedonic model and u represents the error term of the econometric model. The specification that was used as a basis for (4.1) was the one that best described the logarithm of prices for existing apartments, a sub-market that accounts for nearly 59 percent of all transactions available for hedonic regression analysis (see Table 3.1). Similarly, the unrestricted model was defined as:

$$\ln(p) = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + \gamma_0 . S + \sum_{k=1}^{K} \gamma_k . (x_k^* . S) + v, \tag{4.2}$$

where S is a dummy variable identifying the hypothetical structural break. When the existence of the structural break was investigated for the separation between existing and new residential properties, the S was represented by a dummy variable, which assumes the value 1 when the property is new and the 0 value otherwise. Similarly, when the constancy of parameters was tested over apartments and houses, S was represented by a dummy variable, which assumes the value 1 when the property is considered to be a house and the value 0 when the property is an apartment. Under this framework, the null, or parameter stability hypothesis, was defined as $H_0: \gamma_0 = \gamma_1 = \ldots = \gamma_k = 0$. It follows from the above exposition that the constancy of the population parameters was tested in the form of the following working hypothesis:

Hypothesis 1: *The parameters do not change over existing and new residential properties; and*

Hypothesis 2: *The parameters do not change over apartments and houses.*

The specification that was used in both hypotheses was submitted to a process that involved the comparison of many alternative specifications and in which a battery of individual and joint significance tests were used to assess the quality of the model. The size of the samples used to investigate the two hypotheses was not the same. This has to do with the fact that, while the first hypothesis was investigated by extending the coverage of the existing apartments' model from existing apartments to all (i.e., new and existing) apartments data, the validity of the second hypothesis was examined by extending the coverage from existing apartments to all (i.e., apartments and houses) transactions of existing properties.

4.2.2. The value of energy efficiency

After deciding whether or not the same price function can be used to all data, it was possible to tackle the key research question of this work, which was formulated in the following manner:

Hypothesis 3: *Other things equal, increased energy efficiency has a positive impact on the transaction price of residential properties in Portugal.*

This hypothesis states an inequality assumption. When quality and other characteristics of transacted properties are controlled for in the hedonic model, the partial effect of a change from a less to a more energy efficient residential property will be greater than zero. As explained in Chapter 3, the group of most energy efficient dwellings includes all properties with an A^+, A, B or B^- EPC rate [28]. Conversely, properties with other EPC rates (i.e., those with estimated annual energy needs that are higher than those of reference) are identified as to be less energy efficient. In practice, the hedonic price model used to test this hypothesis was formulated as:

$$\ln(p) = \beta_0 + \beta_1 . E + \sum_{k=2}^{K} \beta_k . x_k^* + u. \tag{4.3}$$

The parameter of interest is β_1, representing the partial effect of E, on the natural logarithm of transaction prices, which can be interpreted as a relative or percentage price change as section (2.3.1) points out. For reasons associated with the definition of the zero value as the borderline situation in the inequality, it is convenient to formulate Hypothesis 3 as a one-tailed test. Thus, the null was defined as $H_0 \colon \beta_1 \leq 0$, which will be rejected if the test statistic associated with β_1 is lower than the critical value found for a chosen significance level [29]. In this framework, the null hypothesis represents the idea that energy efficiency has either no or negative impact on prices.

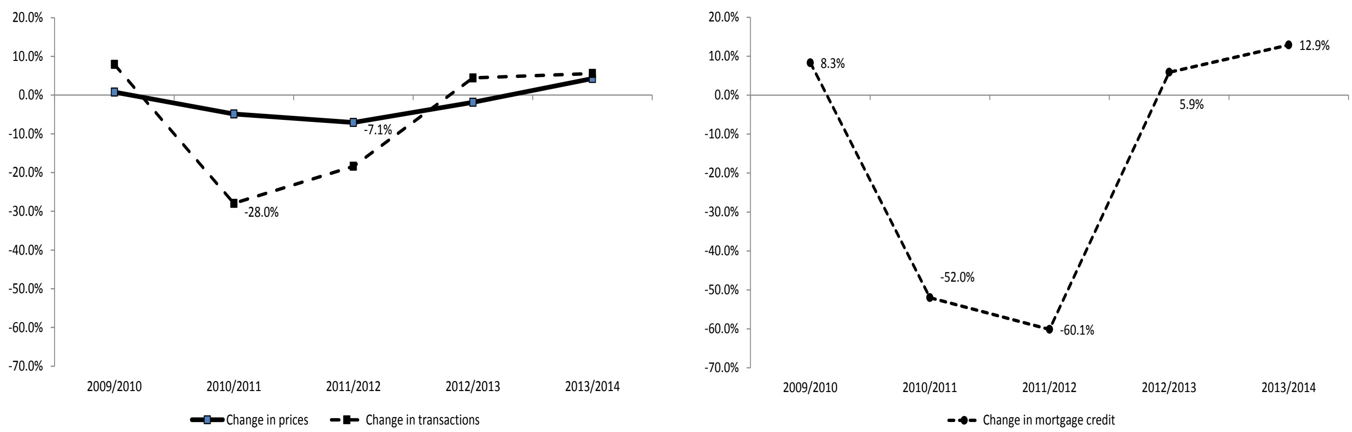
[28] For the sake of simplicity, the A+, A and B, B^- rates will hereafter be designated as A or B rates.

[29] For an excellent explanation on one-tailed tests, see Griffits et al. (1993: 139:140).

4.2.3. Overall market conditions

An interrelated research question to whether or not energy efficiency has an impact on residential property prices is the issue of knowing if its magnitude, sign and statistical relevance change according to the state or condition of the housing market (e.g., depression, expansion periods). Hyland et al. (2013) analyzed the Irish rental and sales market from 2008 to 2012, and found evidence supporting the idea that the effect of energy efficiency was not stable over the years and that was stronger when the market conditions were worse. As in the Irish case, the available data broadly corresponds to a time frame in which the residential property market suffered from severe mortgage credit restrictions. This situation is portrayed in Figure 4.1, where the year on year rates of change in prices, the number of properties sold and mortgage credit are depicted for the 2009 – 2014 period. The data were taken from INE (2015) and *Banco de Portugal*'s (n.d.) website.

Figure 4.1: Mortgage credit, prices and number of transactions from 2009 to 2014



As the left panel of Figure 4.1 highlights, with the exception of 2014, the period is essentially characterized by price drops. The lowest fall occurs in 2012 (-7.1%), following the -52 percent and -28 percent year on year drops in the value of mortgage credit and in the number of properties sold in 2011 [30]. With this in mind, it is important to investigate if the partial effect associated with the dummy variable signaling dwellings bearing A and B energy efficiency rates remained constant during 2009-2013 or if, as suggested by Hylan et al. (2013), it could have been influenced by the worsening of the conditions of the

[30] Portugal requested international financial assistance from the IMF, European Central Bank and the EU in April 2011.

Portuguese housing market. The working hypothesis, which was formulated to investigate this issue, can be stated as follows:

Hypothesis 4: *The partial effect of energy efficiency on residential property prices does not change over time.*

The validity of this assumption was tested using information taken from the following hedonic price model:

$$\ln(p) = \beta_0 + \beta_1 . E + \sum_{t=2}^{5}(\beta_t . D_t) + \sum_{k=6}^{K}(\beta_k . x_k^*) + \sum_{t=2}^{5} \alpha_t . (E . D_t) + v, \tag{4.4}$$

where D_t is a dummy variable assuming the value 1 when a transaction takes place in year t (and 0 otherwise) and v corresponds to the error of the model. The parameters of interest are α_t, $t = 2, \ldots, 5$, where α_t is the coefficient of the interaction term between the dummy variable identifying most energy efficient residential properties and the dummy variables identifying the year of transaction. Since there are five years of data and 2009 is the regression's base year, there are only four time dummy variables and four interaction terms in the model (i.e., $\alpha_2, \ldots, \alpha_5$).

Contrary to the previous assumption, Hypothesis 4 is not an inequality. Thus, the null hypothesis for the joint test of significance, which states that time has no impact on the partial effect of energy efficiency, can be stated as $H_0: \alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_5 = 0$. The rejection of the null for one interaction term implies that the variable is statistically significant and that time may have an influence on the valuation of energy efficiency. If the interaction terms are not rejected, then the partial effect of energy efficiency on the logarithm of transaction price is equal to β_1 in the base year and $\beta_1 + \alpha_t$ in the remaining years.

4.2.4. Building technology and quality of construction works

When compared with technology that is used in the construction of other durable goods (e.g., cars), the one that is applied in the construction of a residential property can be considered to be relatively stable. However, since building techniques have generally improved over the years with the incorporation of tighter energy efficiency (and other) standards, it is expected that some sort of association between the time in which a property was built and its energy performance exists. It is therefore natural to raise the question of whether or not the impact of energy efficiency on residential property prices is invariant to the different building technology construction periods, which can be identified for Portugal. More specifically, the hypothesis on which one wants to shed some light can be written as follows:

Hypothesis 5: *The partial effect of energy efficiency on residential property prices is invariant to building construction technology time periods.*

To test this, four main construction technology time periods were identified. The first one refers to all completions that were carried out before 1960 and characterizes a period predominantly dominated by *alvenaria de pedra (*stone masonry buildings). The second one comprises all building completions from 1961 to 1990 and refers to a period that is typically characterized by a*lvenaria de tijolo furado* (clay hollow-brick masonry buildings). The third one covers all buildings that were completed after 1990, the year in which the first thermal building regulation was introduced [31], and before 2006, the year in which a new thermal building regulation entered into force [32]. Finally, the fourth period, which was left out from the regression to serve as the base construction time period, represents the most recent building technology and includes all dwelling completions carried out after 2006. The model that was developed to investigate this hypothesis was the following:

$$\ln(p) = \beta_0 + \beta_1 . E + \sum_{t=2}^{4} (\beta_t . Q_t) + \sum_{k=4}^{K} (\beta_k . x_k^*) + \sum_{t=2}^{4} \gamma_t . (E . Q_t) + e^*, \quad (4.5)$$

where Q is a dummy variable, which assumes the value 1 if the property belongs to the *ith* construction technology period (0, otherwise) and e^* is the error term of the model. The null hypothesis, which was used to investigate the validity of the present hypothesis for the joint test of significance, is written as $H_0 : \gamma_2 = 0,\ \gamma_3 = 0,\ \gamma_4 = 0$. In this setting, if a given γ_t is statistically different from zero, then it can be said that the partial effect of energy efficiency is not invariant to that particular construction time period.

Construction technology time periods were not the only supply side factor that was investigated as having a possible influence in the formation of the partial effect of energy

[31] *Decreto-Lei n.º 40/1990.*
[32] *Decreto-Lei n.º 80/2006.*

efficiency. Another factor has to do with the quality of the construction of the dwelling, which ranges from the excellence of the project to the quality of building materials used at latter construction phases. Since building materials and their thermal and insulation (among other) characteristics influence the energy efficiency performance of a property, there is some *a priori* belief that their quality, especially when it is higher than used in standard construction works, impacts positively on the way markets value energy efficiency. Following this reasoning, it was chosen to formulate this hypothesis as:

Hypothesis 6: *The partial effect of energy efficiency on residential property prices does not change in response to the quality of the construction works.*

In the IMI, appraisers have to take into account the quality of the project, the thermal and acoustic insulation characteristics of a dwelling and the quality of building materials used in a residential unit (DGI, 2011: 24). In practice, appraisers signal construction works by providing a value to its quality that ranges from a null value to a maximum of 0.15 (DGI, 2011: 21). Following this, three different categories were identified in relation to the quality of construction works. The first one identifies standard quality, which includes all dwellings to which the null value was attributed. This category, which includes around 90 percent of transactions, was left out from the regression for reference category. The second category signals good quality of construction works. This is defined as all dwellings to which appraisers attached a value for the quality of construction works that was greater than 0 and smaller than 0.1. Finally, the last category covers the remaining transactions, which identifies residential units with an outstanding quality in construction features. The model that was developed to investigate this issue is the same as the one that was used for Hypothesis 5. Thus, if a given γ_t is statistically different from zero, then it can be said that the partial effect of energy efficiency is not invariant to the quality of construction works and, as such, Hypothesis 6 does not hold.

4.3. Hedonic regression results

4.3.1. Evidence on sub-market model specificities

This section investigates the existence of sub-market model specificities and the reasonableness of using different models for different market segments. As mentioned in section 4.2.1., this issue was analyzed using the existing apartments' specification as a starting point for structural break testing. Table 4.1 provides the OLS results of the hedonic price model that was used as a basis for the tests.

Table 4.1: Hedonic regression results for existing apartments

Explanatory variable	*Variable description* $^{(+)}$	*Exp. sign*	*Param. estimate*$^{(.)}$	*Robust std. error*	*VIF*
Constant term	-	(+)	10.096**	.00708	0.0
DENERGYAB	*Dummy for EPCs with A or B rates*	(+)/(-)	0.118**	.00192	1.3
D2010	*Dummy for 2010*	(+)/(-)	-0.004	.00218	1.7
D2011	*Dummy for 2011*	(+)/(-)	-0.071**	.00245	1.6
D2012	*Dummy for 2012*	(+)/(-)	-0.152**	.00270	1.5
D2013	*Dummy for 2013*	(+)/(-)	-0.182**	.00270	1.5
SQRTGRFA	*Square root of gross floor area*	(+)	0.144**	.00061	1.3
SQRTDEPFLOORA	*Square root of dependent floor area*	(+)	0.024**	.00049	2.1
SQRTDWELLTRANSA	*Square root of age of dwelling, at transaction date*	(-)	-0.030**	.00118	5.4
DCSYSTEM	*Central heating and/or air conditioning*	(+)	0.079**	.00329	1.2
DABSLIFT	*No elevator in more than 3 storey high buildings*	(-)	-0.071**	.00316	1.1
DCOND	*Private condominium*	(+)	0.058**	.00591	1.4
DSWIMM	*Swimming pool*	(+)	0.153**	.00475	1. 6
DPARKING	*Parking facilities*	(+)	0.057**	.00228	2.1
DCONSTP2	*Dwellings completed > 1990 and <= 2006*	(-)	-0.113**	.00347	4.7
DCONSTP3	*Dwellings completed > 1960 and <=1990*	(-)(-)	-0.155**	.00551	8.7
DCONSTP4	*Dwellings completed in or before 1960*	(-)(-)(-)	-0.144**	.00899	5.2
DCONSTQ2	*Good quality of construction works and materials*	(+)	0.056**	.00342	1.2
DCONSTQ3	*Excep. quality of construction works and materials*	(+)(+)	0.139**	.00921	1.2
DREGION1	*North region (without Porto metropolitan area)*	(-)	-0.365**	.00303	1.3
DREGION2	*Metropolitan Porto area*	(-)	-0.296**	.00237	1.7
DREGION3	*Centro Region*	(-)	-0.252**	.00255	1.4
DREGION5	*Alentejo region*	(-)	-0.039**	.00663	1.1
DREGION6	*Algarve region*	(-)	-0.010*	.00351	1.4
DREGION7	*Açores and Madeira*	(-)	-0.005	.00783	1.1
DSEA	*Parish with access to the sea*	(+)	0.113**	.00193	1.3
DLX	*Property located in Lisboa*	(+)	0.349**	.00414	2.1
DPORTO	*Property located in Porto*	(+)	0.331**	.00448	1.2
DSCENIC2	*Visual prominence of the property*	(+)	0.100**	.00411	1.2
DSCENIC3	*Extremely good visual prominence*	(+)(+)	0.266**	.01184	1.1
DBADLOC	*Bad location, as measured by IMI's location coef.*	(-)	-0.171**	.00429	1.1
DEXCPLOC	*Good location, as measured by IMI's location coef*	(+)	0.316**	.00322	1.8

Number of obs. used in estimation:149,920
Regressions' $R^{2:}$ 0.68

Breusch-Pagan type test
H_0: Homoskedasticity
Test stat.: 3,317.61
p-value: <.0001

RESET type test
H_0: Correct specification
LM test stat.: 0.328
p-value: 0.57

Notes:$^{(+)}$ A complete description of the variables is available in Appendix IV; $^{(.)}$ p-value < 0.05; ** p-value <0.0001.*

As the table shows, the homoskedastic error term is strongly rejected (*p*-value smaller than 0.0001). This should not be seen as a surprise as heteroskedasticity is a common feature in hedonic price models. However, while its presence does not introduce bias in OLS parameter estimates (which are our main research interest), it typically invalidates standard statistical inference used in regression analysis and stresses the need to use robust statistics to make valid inference from the hedonic price models [33]. More importantly, the table provides good indications as to the statistical significance, directional sign and magnitude of parameter estimates of chosen model. With the exception of the dummy variables signaling transactions carried out in 2010 and in the *Açores* and *Madeira* islands, all the coefficients are statistically significant. While the reason for rejecting the dummy for 2010 simply highlights the fact that residential properties may not have had any price change from 2009 and 2010, the reason for the rejection of the relevance of the islands may have to do with the relatively low number of transactions that these two regions represent.

The coefficients show the expected signs. Examples include, among other, the parameter estimate for residential units with access to a swimming pool and parking facilities, which are positive as expected. Conversely, the sign of the variable signaling a bad location is negative. Another point illustrating the quality of the model is also shown in the magnitudes of the coefficients of the covariates measuring the impact of different quality levels of dwelling features. For instance, although the impact of visual prominence is estimated to be positive for the two dummy variables used to capture this feature, the price premium associated with the residential units with less visual importance is, as expected, smaller than the one associated with extremely important visual prominence (coefficients of 0.100 and 0.266, respectively). As shown in the rightmost column of Table 4.1, the variance inflation factor (VIF) does not detect excess of multicollinearity among chosen covariates. Finally, it should be noted that the null hypothesis of the correct functional form, which is tested by a Ramsey (1969) RESET type test, is not rejected (*p*-value of 0.57). All these points are evidence of the quality of the chosen specification.

Table 4.2 reproduces the results that investigate the stability of coefficients across, on the one hand, new and existing properties and, on the other, apartments and houses. This was carried out through the help of a structural break test, which compares a restricted model (i.e., one

[33] The presence of a homoskedastic error term was never accepted in the models that were used to test the set of hypotheses mentioned in section 4.2.

that assumes parameter constancy across the different sub-markets), with an unrestricted model specification (see Appendix III). The table also provides the results of the Ramsey (1969) and Breusch and Pagan (1979) tests used to check for other specification problems. All these tests were computed using heteroskedasticity-robust LM statistics. The Ramsey (1969) and Breusch and Pagan (1979) test results refer to the restricted version of the existing apartments' hedonic model.

Table 4.2: Test results for sub-market model specification

	Statistical tests			*n*
	RESET[(c)]	*Breusch-Pagan*[(d)]	*Chow*	
Hypothesis 1: New versus existing[(a)]	13.040 *(.0003)*	6,289.61 *(<.0001)*	2,416.14 *(<.0001)*	209,330
Hypothesis2: Apartments versus houses[(b)]	15.763 *(.0001)*	9,580.22 *(<.0001)*	4,184.49 *(<.0001)*	183,202
Hedonic model for existing apartments	0.328 *(.57)*	3,317.61 *(<.0001)*	-	149,920

Notes: A description of the variables is available in Appendix IV; LM test statistics. p-values in parenthesis; [(a)] H_0: Coefficient stability; over the new and existing dwellings. DNEW used for the structural break;[(b)] H_0: Coefficient stability over apartments and houses. DHOUSE used for the structural break; [(c)] H_0: Correct specification of the functional form; [(d)] H_0: Homoskedastic or constant variance of the error term.

Two points stand out from Table 4.2. First, it is interesting to note that, while the specification for the existing apartments' market segment passes the Ramsey (1969) type specification test (*p*-value of 0.57), it is rejected when new and existing apartments' data and transactions of all existing residential properties are polled together into the same model. Second, more importantly, the Chow (1960) type test clearly rejects the hypotheses of coefficient stability over new and existing dwellings and across apartments and houses. When taken together, these results support the view that the use of the same model specification for different sub-markets is unsuitable. This conclusion is in line with the findings of earlier data exploratory analysis, which revealed clear differences across new and existing dwellings and between apartments and houses (see Chapter 3). As a conclusion, it can be said that statistical tests, combined with the analysis of the data, support the use of different models for the different residential property sub-markets.

Following these findings, the models that were used as a tool to value the impact of energy efficiency on transaction prices were defined separately for four strata, each of which resulting from the crossing of the new versus existing and apartment versus house market dimensions. By looking at each one of these sub-markets separately, it was possible to increase the flexibility of the specification process and incorporate key characteristics without running the danger of having to build a too complex and cumbersome single hedonic price model. The main dissimilarities in the four hedonic specifications stem from the different treatment given to area and age variables. In relation to the area variable, it was possible to explore, for each sub-market, the three different measures available in the dataset. These were the gross floor area, the dependent floor area and the plot area. While gross floor and dependent areas are important characteristics for all dwellings, the area of the plot of uncovered land (i.e., outside space), is a feature that is essentially associated with houses and that was only included in the specification of this dwelling category. The age variable was also a feature differentiating the specifications used for new and existing dwellings. In particular, the number of complete years at transaction date was only taken into account in the models for existing dwellings. The final models for existing apartments, new apartments, existing houses and new houses are available in Appendix V.

4.3.2. Impact of energy efficiency in the Portuguese residential market

Table 4.3 presents the results of the estimated partial effects and of the one-tailed tests used to investigate the validity of the third hypothesis. These were obtained using four different model specifications of (4.3), which were applied to the existing apartments', new apartments', existing houses' and new houses' data. The explanatory variable of interest is DENERGYAB, the dummy that identifies A and B EPC rates with a 1 and other energy ratings with a 0.

As Halvorsen and Palmquist (1980) note, in a log-linear model such as (4.3), where the variable of interest is binary, the relative change in a continuous dependent variable, or r, is obtained by $exp(\beta_1) - 1$ (see section 2.3.1). Kennedy (1981) points out that this formula produces biased estimates for r, and propose an approximate unbiased estimator assuming the normality of u, the error term in (4.3). Giles (1982) and, more recently, van Garderen and Shah (2002) present numerically identical versions of an exact unbiased estimator for r. However, for small variances of $\hat{\beta}_1$, the results using Kennedy's (1981) and van Garderen and Shah's (2002) estimators are very similar to those obtained by Halvorsen and

Palmquist's (1980) formula [34]. In the present situation, the large samples used in the estimation of the coefficients contribute to the existence of small variances. As a consequence, the results for the relative and percentage change are presented using the computationally simpler Halvorsen and Palmquist's (1980) estimator.

Table 4.3: Impact of energy efficiency on property prices

	Sub-market			
	Existing apartments	*New apartments*	*Existing houses*	*New houses*
DENERGYAB:				
Parameter estimate of β_1	0.118	0.123	0.045	0.055
Estimated perc. change	12.5%	13.1%	4.6%	5.7%
One-tailed test ($\beta_1 \leq 0$)*:*				
Test statistic	61.71	50.36	7.31	8.87
p-value	<0.0001	<0.0001	<0.0001	<0.0001
Number of obs. used in estimation	149,920	59,410	33,282	13,533
Regressions' adjusted R^2	0.676	0.733	0.670	0.753
RESET type test				
LM test statistic	0.328	2.73	3.00	3.77
p-value	0.567	0.098	0.083	0.052

As the bottom lines of Table 4.3 show, all four models pass the Ramsey (1969) type specification test at the usual 5 percent level of significance. More importantly, the null of the one-tailed test is rejected for all segments and the alternative hypothesis, which holds that the impact of energy efficiency on residential property prices is positive, is not rejected. However, the magnitude of price premiums is different for apartments and houses, with apartments having a market premium of around 13 percent, and houses displaying a considerable lower premium of 5 to 6 percent.

Price premium differences across dwelling types are not a novelty and have been reported in Fuerst et al. (2015), where the impact of energy efficiency on prices is shown to be lower for detached and semi-detached houses than for flats or terraced dwellings. For detached houses specifically, no significant price effects were estimated, something which the authors attribute to the influence of a small and atypical portion of the sample of detached dwellings located in sparsely populated areas. In order to investigate if dwellings located in rural areas impose a downward effect on estimated energy efficiency partial effects, it was decided to rerun the

[34] The existence of unsubstantial differences in many empirical situations is acknowledged in Giles (1982), van Garderen and Shah (2002) and, more recently, Megerdichian (2018).

regressions using a new dummy variable, which was built to identify the transactions in urban areas, and an interaction between this covariate and the variable signaling residential units bearing A and B energy rates. The variable is based on a list of statistical cities, which was compiled by INE (2011). As it is based on a 2011 classification, the new dummy provides a static and, as such, imperfect identification of urban and rural locations, which naturally evolved from 2009 to 2013. For example, all the urban areas that have been built after 2011 are not identified as such by the new covariate. Conversely, an area defined as urban in 2011 might have been sparsely inhabited in 2009. However, given the information at hand, the new variable provides a reasonable proxy of what might be defined as rustic or urban. Accordingly, the percentage of houses falling into non-urban areas (as defined by the new dummy variable) is particularly relevant for existing houses (54%) and new houses (57%). As expected, this percentage is considerably lower for apartments (27%). The next table presents the main regression results of this exercise.

Table 4.4: Valuation of energy efficiency in urban areas

Explanatory variable	*Variable description*	*Sub-market*			
		Existing apartments	*New apartments*	*Existing houses*	*New houses*
DENERGYAB	*Dummy for EPCs with A or B rates*	*0.095*** *(.00339)*	*0.110*** *(.00423)*	*0.036*** *(.00847)*	*0.044*** *(.00829)*
DENERCITY	*Interact. between DENERGYAB and urban areas*	*0.033*** *(.00385)*	*0.016** *(.00504)*	*0.018* *(.01133)*	*0.025** *(.01141)*
	Estimated perc. change on prices for urban areas	*13.7%*	*13.4%*	*5.5%*	*7.1%*

Notes: p-value < 0.05; ** p-value <0.0001. Robust standard errors in brackets.*

The coefficients of the new interaction term suggest the existence of an energy efficiency price premium associated with urban areas. However, these are not enough to approximate apartments' and houses' price premiums. For instance, the price premium for new apartments in urban areas is estimated to be 13.4 percent, while for new houses is 7.1 percent. This yields a difference of 6.3 percentage points, which is comparable to the 7.4 percentage points difference that can be obtained with the information reported in Table 4.3 for new dwellings. Overall, it can be said that the results support the idea that the difference between apartments and houses cannot be attributed to rural or urban valuation idiosyncrasies.

The disparity found in the way apartments and houses reward energy efficiency could be justified on other grounds. A possible explanation is anchored in the uses given to apartments and houses and on how these uses reflect buyers' preferences in relation to (the same) property characteristics. For instance, buyers of houses might be more interested in characteristics such as area (e.g., plot of land for gardening activities or the latter construction of a swimming pool), than in energy-saving characteristics of the residential unit. Another possible explanation could be found in dwellings bought for seasonal or vacation purposes. Since vacation residential units are used only during part of the year, their energy saving characteristics might not be considered as important as in the case of the purchase of a dwelling that is going to be used as a permanent residence. If houses are more often bought for vacation purposes than apartments, then this could help explaining why energy efficiency attributes are less valued in the former dwelling category than the latter. Although the IMT and IMI dataset does not identify vacation from permanent residences, it is possible to estimate from the total housing stock the percentage of apartments and houses that are used on a seasonal basis. Using 2011 Census data, it is possible to see that, whereas 23 percent of all houses are used seasonally, only 16 percent of the total apartment stock is inhabited on this basis. These figures reinforce the explanation made above and help explaining, at least partially, why homes have a smaller premium attached to higher energy efficiency levels.

Another possible explanation for the apartment versus house difference stems from physical or engineering considerations and their association with the perception of higher or lower future energy bills. As houses are usually bigger than apartments, it is technically more difficult (and costly) to ensure high energy saving attributes in houses than in apartments. Moreover, the building envelope of a house (i.e., what separates the indoor and outdoor environments) does not include shared walls. Apartments, on the contrary, are pieces of a bigger envelope and are often concomitant to other buildings. For this reason, apartments are often less exposed to the external environment than houses and therefore may be associated with lower utility bills than houses in maintaining high energy efficiency standards. As a result of these factors, it is reasonable to assume that the market discounts these costs and places a smaller price premium to energy efficiency in the case of houses.

A final interesting point that emerges from Table 4.3 is the price premium attached to new and existing residential property units. Although this is more evident for houses than apartments, the results suggest that the market adds an additional premium to new dwellings. More concretely, the difference between new and existing price premiums is 1.1 and 0.6 percentage points for houses and apartments, respectively. To further investigate this issue, it was decided to run two additional regressions, one for apartments and another for houses, with two additional covariates identifying new properties and the cross product between this variable and the dummy identifying A and B energy efficient dwellings. Although pooling together new and existing transactions into the same regression is contrary to the approach that was chosen to specify hedonic price models, it is expected that, if the difference in price premiums is strong, the inclusion of the new interaction term would always capture a positive and statistically significant price premium. In addition, it was decided to explore the definition of new dwellings, which includes not only newly built units but also reconstructed or renovated properties [35], to see whether the market attaches a price discount to the latter type of new properties. To this end, the regressions for new apartments and houses were rerun with an additional interaction term resulting from the product of the dummy variables identifying renovated dwellings and the dummy variable grouping A and B energy efficiency units. Due to construction technology constraints, renovated properties may be perceived as demanding higher future costs in maintaining higher energy efficiency standards than newly built properties. As such, the expected outcome is that the interaction term exhibits a price discount and that energy efficiency is less rewarded in renovated properties. Table 4.5 details the results of these two additional exercises.

Table 4.5: Assessing the influence of being new

Explanatory variable	*Variable description*	*Exercise 1*		*Exercise 2*	
		Apart.	*Houses*	*New apart.*	*New houses*
DENERGYAB	*Dummy for EPCs with A or B rates*	*0.122*** *(.00186)*	*0.003* *(.00954)*	*0.116*** *(.00249)*	*0.069*** *(.00592)*
DENERGNEW	*Inter. between DENERGYAB and new dwellings*	*-0.002* *(.00296)*	*0.037*** *(.00614)*	-	-
DENERGEN	*Inter. between DENERGYAB and renovated units*	-	-	*0.061*** *(.00957)*	*-0.051** *(.01716)*
	Estimated percentage change on prices	*12.7%*	*4.1%*	*19.4%*	*1.8%*

Notes: p-value < 0.05; ** p-value <0.0001. Robust standard errors in brackets.*

The results of the first exercise show that the expectation formulated above on the importance of the new versus existing split is only confirmed for houses. For apartments, however, the -0.002 coefficient for most energy efficient new dwellings is statistically insignificant, suggesting that, in the case of apartments, being new does not have an impact on the formation of energy efficiency price premiums. As for the results of the second exercise, which uses new dwellings transactions data and renovated new dwellings as a proxy of

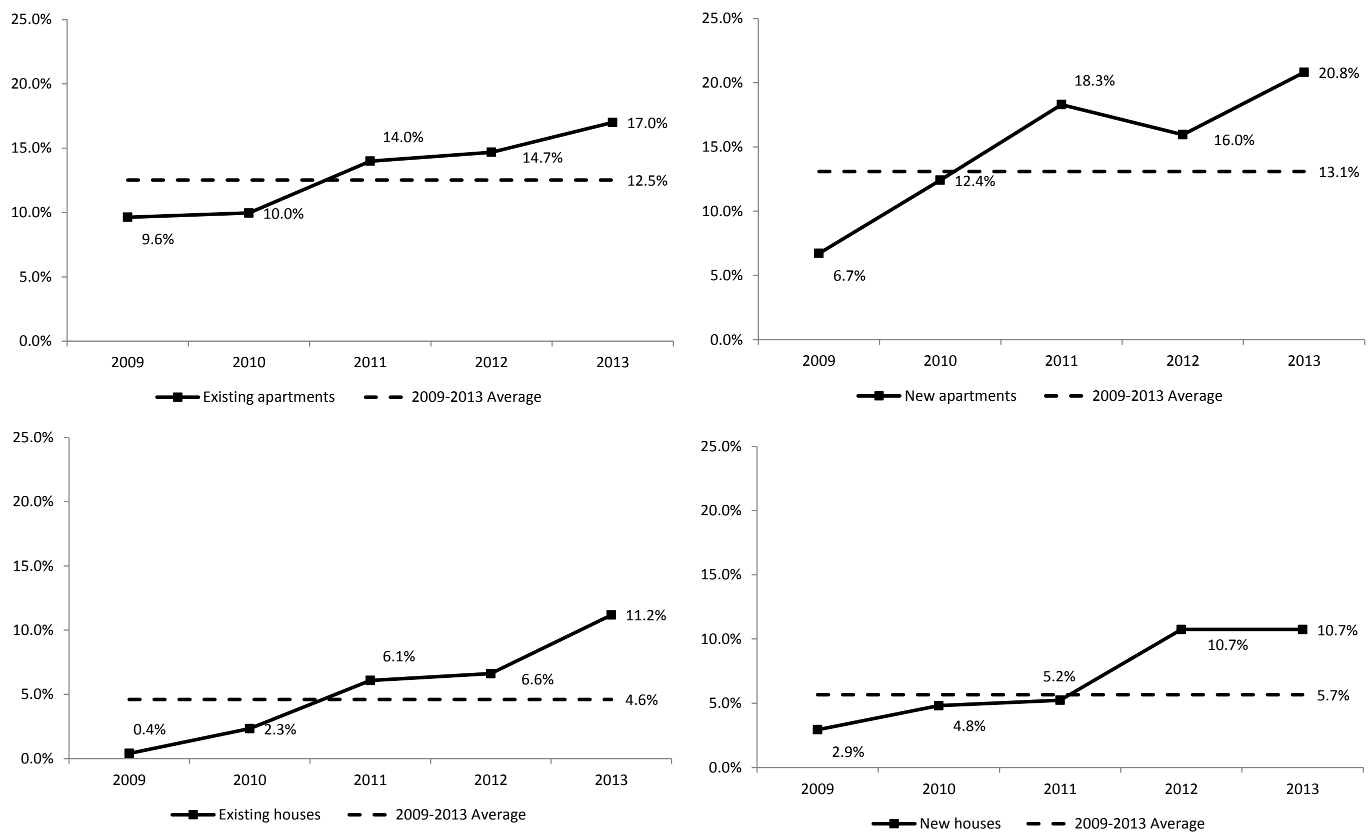
[35] Hereafter simply identified as renovated.

existing dwellings, it is interesting to note that the initial expectation regarding the existence of a price discount is also only confirmed for houses. While for new houses there is a statistically significant energy efficient price discount for renovated dwellings (-0.051 coefficient), the opposite happens for new apartments (0.061 coefficient). This result is partly explained by the weight of renovated dwellings in total sales, which is much lower for apartments than for houses (9% and 36%, respectively). Moreover, the percentage of new houses built before the introduction of the first thermal building regulation in 1990 is almost 20 percent. For apartments, this percentage is only of 2 percent. Given the relevance of older construction technologies in the total number of new dwelling transactions, it is not surprising to see the market associating a stronger energy efficiency price discount for houses, which are perceived as requiring higher maintenance costs than apartments for keeping high energy efficiency standards. All in all, these results suggest that the price premium that was estimated for new dwellings is not the same for apartments and houses, with the former dwelling category showing some evidence that energy efficiency partial effects are not influenced by the new versus existing split. For houses, the anticipation of higher utility bills and other costs, together with the different uses given to houses and apartments, seem to give rise to price discounts and to justify, at least partly, the existence of a lower predisposition to value energy efficiency.

4.3.3. Valuation of energy efficiency through time

This subsection focuses on the evolution of energy efficiency partial effects through time. The results that were used to investigate this issue were derived using equation (4.4), which used interaction terms between the dummy variable identifying A and B rated properties. With the exception of the cross product between energy efficiency and the 2010 dummy, all remaining interactions were found to be statistically relevant for all sub-markets (results shown in Appendix V). Figure 4.2 illustrates the yearly evolution of the market price premium for each one of the four housing sub-markets considered in this thesis. For comparison purposes, a dashed line was introduced in each panel of the figure, representing the average energy efficiency price premium for the 2009-2013 period, which was presented in Table 4.3.

Figure 4.2: Effect of energy efficiency on prices from 2009 to 2013



The partial effect of energy efficiency on residential property prices shows a clear upward tendency in all sub-markets. For instance, for existing properties, the price premium placed on A and B rated dwellings jumped from 9.6 percent in 2009 to 17 percent in 2013 for apartments and, in the case of houses, from 0.4 percent in 2009 to 11.2 percent in 2013. These results indicate that the hypothesis on the constancy of the valuation of energy efficiency throughout time, does not hold for the analyzed time period. This pattern is consistent with Hyland et al. (2013) where, for the Irish housing market, the effects of the energy EPC rating were found to be higher when sale conditions were worse.

The 2009-2013 period was strongly marked by a severe contraction of the mortgage credit and the worsening of the sale conditions. In situations where markets are depressed, buyers may look at more energy efficient properties as an extra quality indication and as an extra guarantee for the value of their money. In addition, as Hyland et al. (2013: 949) highlight, tighter credit constraints may limit the availability of finance to conduct major repairs and renovations on less energy efficient homes leaving investors placing a higher value on more energy efficient properties since they are perceived as not having to require investment on the renovation of the dwelling. Another possible explanation for the results rest on the idea that, while the benefits of energy certificates may have not been so evident in the eyes of the market at the beginning of 2009 (i.e., when the presence of an energy certificate began to be mandatory in all dwelling transactions), the awareness and worth of the EPC label may have been gradually consolidated from 2009 to 2013.

4.3.4. Influence of building technology and quality of construction works

Table 4.6 presents the results of the two hedonic regression models that investigated the degree to which construction technology periods and the quality of construction works influence the formation of the energy efficiency price premium (equation 4.5). For the sake of simplicity, only the estimates for the variable signaling most efficient properties and its interaction between the dummies controlling for the different construction periods and quality works are shown. The full regression outputs can be seen in Appendix V.

Table 4.6: Point estimates of relevant interaction terms

Explanatory variable	Variable description	Existing apart.		New apartments		Existing houses		New houses	
		Hyp. 5	Hyp. 6	Hyp. 5	Hyp. 6	Hyp. 5	Hyp. 6	Hyp. 5	Hyp. 6
DENERGYAB	Dummy for EPCs with A or B rates	0.130** (.00493)	0.116** (.00198)	0.122** (.00274)	0.119** (.00263)	0.106** (.01349)	0.045** (.00639)	0.078** (.00668)	0.049** (.0068)
DENERGCONSTP1	Interact. between DENERGYAB and dwellings completed > 2006 (hold-off cat.)	-	-	-	-	-	-	-	-
DENERGCONSTP2	Interact. between DENERGYAB and dwellings completed > 1990 and <= 2006	-0.010 (.00534)	-	-0.002 (.00552)	-	-0.033* (.01515)	-	-0.015 (.01361)	-
DENERGCONSTP3	Interact. between DENERGYAB and dwellings completed > 1960 and <=1990	-0.039** (.00680)	-	0.077 (.04317)	-	-0.054* (.02024)	-	-0.112* (.03955)	-
DENERGCONSTP4	Interact. between DENERGYAB and dwellings completed in or before 1960	0.064** (.01389)	-	0.152** (.03939)	-	-0.158** (.0283)	-	-0.158** (.03319)	-
DENERGCONSTQ1	Int. between DENERGYAB and standard quality of constr. works (hold-off cat.)	-	-	-	-	-	-	-	-
DENERGCONSTQ2	Interact. between DENERGYAB and good quality of construction works	-	0.033** (.00630)	-	0.025* (.00665)	-	-0.027 (.02017)	-	0.042* (.0162)
DENERGCONSTQ3	Interact. between DENERGYAB and very good quality of construction works	-	-0.021 (.02691)	-	0.011 (.01893)	-	0.163* (.06339)	-	0.084* (.0369)
Regression's adjusted R^2		0.676	0.676	0.734	0.734	0.671	0.670	0.754	0.753
RESET type test									
LM test statistic		0.008	0.043	4.153	2.332	5.878	2.910	6.731	4.711
p-value		0.931	0.836	0.042	0.127	0.015	0.088	0.010	0.030

Notes: p-value < 0.05; ** p-value <0.0001. Robust standard errors in brackets.*

The fifth hypothesis analyzes the degree to which the impact of energy efficiency on residential property prices is invariant to the different building technology construction periods identified for Portugal. Since thermal and energy standards have generally been improved from one building technology period to another, it is expected, at least on *a priori* grounds, that the coefficient estimates resulting from the above-mentioned interaction terms would show a negative directional magnitude with a decreasing tendency of its absolute value from less to more recent time periods. The later expectation is based on the idea that older energy efficient existing residential properties may be perceived as demanding higher maintenance costs than energy efficient properties built with more recent building technology.

As Table 4.6 shows, the majority of the coefficients for the interactions between most energy efficiency properties and the dummies controlling for the different building technology periods (9 out of 12) display a negative sign (i.e., signal a price discount). The *a priori* expectations formulated above are confirmed for houses, where the negative directional magnitude of the coefficients increases in absolute value from newer to older time construction periods. For instance, for existing houses, the coefficient estimates for the interaction terms between the dummy for most energy efficient residential units and the dummy signaling different building technology periods are -0.033, -0.054 and -0.158 for the houses built between the 1990-2006, 1960-1990 and before 1960 periods, respectively. In addition, one third of the coefficient estimates (4 out of 12) are statistically insignificant, an outcome that is more frequent for newer construction periods and for apartments. For the oldest construction period considered (i.e., properties constructed before 1960), all coefficients are statistically significant. Moreover, apartments display a positive sign instead of the expected negative sign, an outcome for which it is difficult to find a plausible explanation. Finally, it is worth mentioning that only one of the four models passes the robust specification test. Overall, it can be said that there is some evidence that the impact of energy efficiency on residential property prices is not the same for the properties built before 1960. However, the signal of the impact is not the same for apartments and houses, with the latter showing a positive impact and the former a negative one.

The sixth hypothesis investigated the degree to which the partial effect of energy efficiency on residential property prices varies in response to changes in the quality of construction works (e.g., thermal and acoustic insulation quality, type of building materials used in the construction of the residential property unit). Table 4.6 provides the results of the model that includes an interaction term between the variable identifying A and B rated properties and the dummy variables identifying the properties with higher quality in construction works. As the last lines of the table show, five out of the eight coefficients have a statistically relevant positive sign, thus suggesting the existence of an additional energy efficiency price premium associated with the higher quality of the construction works. Existing residential properties display two coefficients with negative signs. However, these coefficients are not statistically significant. Three out of the four models pass the model specification test. In addition, the inclusion of the new interaction terms did not substantially change the sign and the magnitude of the overall impact of energy efficiency on property prices, which was estimated with the model used for Hypothesis 3. As a conclusion, it can be said that there is some evidence, which is stronger for new than for existing dwellings, that support the idea that above than average construction characteristics increment the premium attached to energy efficiency.

4.4. Summary

The results provided in this chapter clearly support a model specification approach that takes into account the specificities of apartments and houses and new and existing properties. Accordingly, the investigation of the six working hypotheses regarding the influence of energy efficiency in the formation of residential property prices was carried out separately for existing apartments, new apartments, existing houses and new houses. Based on these models, the research confirms the existence of a significant price premium associated with higher energy efficiency labels in the Portuguese residential property sales market. More specifically, when compared with less efficient properties, A and B rated new and existing apartments receive a sales price premium of 13.1 percent and 12.5 percent over the 2009-2013 period, respectively. Houses obtain smaller price premiums, with new and existing houses receiving a 5.7 percent and 4.6 percent sales premium over the same period. The euro value attached to this price premium is sizeable at the point of means: considering, for instance, that the average transaction price of an existing apartment is 97,695 € (see Table 3.1), it corresponds to 12,212 €.

The difference in the evaluation of energy efficiency by apartments and houses is maintained even when a control variable identifying transactions in rural areas is included in the regressions and when each one of the years is analyzed individually. A possible explanation for the higher valuation for apartments may be rooted in the different uses of apartments and houses and on how this reflects buyers' preferences in relation to purchased dwelling characteristics. For instance, a buyer of a house may be more interested in acquiring a big plot of land than concerned with its annual energy efficiency performance. Another possible explanation may stems from physical or engineering considerations. As mentioned above, apartments are often concomitant to other buildings and are, for this reason, often less exposed to the external environment making them more likely to obtain better energy efficiency levels than houses. The fact that energy efficiency is valued differently across houses and apartments is interesting from a policy point of view. Since houses are an important share of the Portuguese housing stock, it may be necessary for policy makers to shape specific policies targeting this sub-market to achieve more energy efficiency standards in the housing sector as a whole. The results show some mild evidence of the existence of higher price premium for new properties, which is more evident for houses than apartments (see Chapter 5 for more on this). A possible explanation may rest on the idea that it might be more costly to maintain an existing house at high levels of energy efficiency and, as such, markets may discount this feature in price premiums.

The results taken from the different hedonic regressions also disclose the existence of an annual upward tendency in the energy efficiency price premium. This result is consistent with the findings of Hyland et al. (2013), which suggest that, for the case of Ireland, the value of certification was higher when market conditions were worse. As the time period under analysis broadly overlaps a situation under which the housing market suffers from illiquidity, uncertainty and credit constraints, buyers may have seen most efficient energy labels as an extra guaranty of value. Another explanation for the existence of an increasing tendency in the valuation of energy efficiency may also rest on a possible gradual incorporation of the benefits of the EPC label, which was first made obligatory to all residential market transactions in 2009.

Finally, the degree to which the impact of energy efficiency on residential property prices varies to different building technologies and to the quality of construction were also investigated. In relation to the first issue, it can be said that the results give an indication that the impact of energy efficiency on residential property prices is not the same for properties built before 1960, the oldest building technology period considered. However, the signal of the impact is not the same for the apartments and houses sub-markets, with the latter showing an *a priori* expected negative impact and the former a positive one for which there is no concrete explanation. In relation to construction quality, it can be said that there is some evidence, which is stronger for new properties, supporting the idea that above than average insulation and thermal property (as well as other construction) characteristics increment the price premium associated with energy efficiency.

Chapter 5 Soundness of hedonic regression energy efficiency partial effects

5.1. Introduction

The impact of energy efficiency on transaction prices of residential properties in Portugal, was estimated using a dataset that, given its size and richness of information, allows to explore the coherence and sensitivity of estimated OLS energy efficiency partial effects to different data and estimation contexts. In particular, this chapter tackles four different estimation issues that, although latent to many hedonic regression studies focusing on energy efficiency partial effects, have been barely researched in the literature.

The first issue addresses the sensibility of energy efficiency coefficients to the replacement of key variables in the hedonic specification by proxies that necessarily display some sort of measurement error. This is an important issue since researchers are often limited by the data they have available for research. The use of alternative measurements for transaction prices is a notable example, where sometimes conclusions have to be drawn from models that use list prices or appraisal values as a dependent variable (e.g., Stanley et al., 2013). To investigate this issue, five experiments were designed, using fiscal appraisals, number of bedrooms and simulated list prices as replacement variables in the hedonic price model specification.

The second issue revolves around the sensitivity of energy efficiency partial effects to the omission of variables that measure the quality of transacted properties. Five different omitted variable experiments were designed, which explored the availability of variables that were expected to be associated with energy efficiency performance (e.g., existence of central heating systems) or that were difficult to be found in similar hedonic studies (e.g., quality of the location, construction quality). The analysis of this issue involved the re-estimation of hedonic price models in these different omitted variable scenarios and the comparison of the magnitude of estimated energy efficiency coefficients across the different experiments.

The third covered issue explored the size of the available dataset to provide evidence on problems associated with the use of large samples. Since standard errors decrease as the size of the sample increases, significance levels of energy efficiency and other parameters may be inflated to a point in which standard *t* and other statistical tests become irrelevant (Ziliak and McCloskey, 2004). Apart from some notable exceptions (e.g., Lin et al., 2013; Zietz et al., 2008), this topic has not deserved much attention in the empirical literature associated with the use of hedonic price models. However, this is an important matter since with the dissemination of energy labels, it is expected that the problems stemming from the use of econometrics in large datasets become more relevant [36]. The present research, which is based on one of the largest datasets employed to study the relationship between energy efficiency and residential transaction prices, provides an opportunity to shed some light on the influence of different sample sizes on the statistical significance of estimated energy efficiency partial effects.

The last explored issue, looks into cross-country comparisons and to the coherence of results to the use of different energy efficiency measurement scales. Although the implementation of a common energy performance labelling scheme in Europe has enhanced the degree of comparability across countries, it is not possible to carry out direct comparisons of the magnitude of different energy efficiency estimates. This has essentially to do with the fact that, in spite of being based on the EPBD and the Energy Labelling Directive (Directive 2010/30/EU), the methodology underlying the implementation of the EPC scheme in each country is tailored to national contexts [37]. However, by introducing some changes in the hedonic price models, it is possible to increase the degree of comparability between studies and present, in this way, a qualitative cross-country assessment of the impact of energy efficiency on dwelling prices. Moreover, as most of the research on this subject focuses on countries located in the north of Europe, it was found relevant to assess the degree to which the results for a southern country such as Portugal were consistently higher or lower than those found for northern European markets.

[36] Fuerst et al. (2015), for instance, base their conclusions on a sample of more than 330 thousand observations. A review of the sample sizes used in studies analyzing the relationship between energy efficiency and residential property prices is available in Chapter 2.

[37] For a good overview of the main differences among different EPC schemes within the EU see Atanasiu and Constantinescu (2011).

This chapter contributes to the literature by providing evidence on issues that, although of practical importance in studies measuring the impact of energy efficiency on residential property prices, have not been much explored. The experiments used as benchmark models the specifications that best answered the question of whether or not energy efficiency is rewarded in the Portuguese residential property market (previous chapter's hypothesis three).

The empirical results provide interesting conclusions. First, they support the idea that the use of list prices as a replacement of transaction prices can lead to an overestimation of the impact of energy efficiency on transaction prices. Second, the experiments covering different omitted variables scenarios stressed the importance of incorporating variables measuring the quality of the location of the dwelling in the hedonic model, since its omission can lead to an overestimation of energy efficiency partial effects. Third, the significance of the explanatory variables appears not to be inflated by the use of the large dataset used in this work. Finally, the comparison across similar studies suggests that price premiums in the Portuguese market are higher than in markets located in the north of Europe.

This chapter is organized into five main sections. Section 5.2 addresses the impact of measurement errors on the estimation of energy efficiency partial effects. Section 5.3 investigates the likely effect of the omission of variables characterizing the quality of the dwelling on estimated energy efficiency coefficients. Section 5.4 looks into large sample problems and assesses the existence of inflated significance levels. Section 5.5 focuses on cross-country comparability. Finally, the last section provides a summary of the main findings.

5.2. Measurement errors

The degree to which energy efficiency partial effect estimates can be affected by the use of mismeasured proxy variables was investigated through the design of five experiments, in which either the dependent or independent variables were replaced by variables measured with errors. In particular, the experiments stem from results obtained using the log-linear hedonic function, which was already presented in (4.3):

$$\ln(p) = \beta_0 + \beta_1 . E + \sum_{k=2}^{K} \beta_k . x_k^* + u, \tag{5.1}$$

where, as defined before, p is the transaction price, β_0 is the intercept term, β_1 represents the partial effect of E, a dummy variable signaling most energy efficient properties, x_k^* refers to the k explanatory variable, β_k corresponds to the kth parameter of the hedonic model and u represents the error term. The experiments cover the situation in which p is replaced by an error-prone measure of the price of the property, $\dot{p}$, or in which x_k^* is substituted by a mismeasured variable $\dot{x_k}$. The focus of the experiments is the impact on β_1, which measures the effect of energy efficiency on the natural logarithm of transaction prices.

Fiscal appraisals and list prices were used as replacement variables of p in four experiments. While the former variable was available in the data, the latter had to be generated through a Monte Carlo data generation process. The data generation process involved the simulation of 1,000 vectors of prices, which were used in the re-estimation of the same number of regressions for each one of the four market segments considered. List prices, or LP, are defined as being set above transaction prices by a given percentage. In this data generation process, $LP = (1 + k).p$, where k is a proportion, and $(1 + k)$ represents an upward bias in list prices, which is commonly reported in the literature in which these and transaction prices are compared (see, *inter alia*, Fuerst and Shimizu, 2016). Defining measurement error as $e = LP - p$, and assuming that it follows a normal distribution, it is easily demonstrated (see Appendix VI) that $e \sim N(k.\mu_{p^*}, k.\sigma_{p^*})$, where μ_{p^*} and σ_{p^*} are the mean and standard deviation of observed transaction values. For the experiments involving list prices, a reference proportion was chosen to enable the development of different structures for the error term e. The chosen reference k, 0.06, is coherent with reported values in the literature [38].

The five experiments involving the replacement of p were defined in the following way.

Experiment 1: *Appraisals as a proxy of transaction values*

In this case, equation (5.1) was re-estimated with $\dot{p} = \dot{a}$, where $\dot{a}$ corresponds to the appraisal values that are carried out for fiscal purposes. The choice of fiscal appraisal values provides the means to investigate the effect of the inclusion of a variable that, although having a high correlation with transaction prices, is generally set below sales prices (see Table 5.1 below). This replacement also allows testing the quality of the chosen model specifications. Since the

[38] Fuerst and Shimizu (2016: 112), for instance, report an average transaction price that is 3.6 percent lower than the average asking price. For Portugal, it is possible to derive from Ramos et al. (2015b) a 5.2 percent difference between list prices and bank appraisal values (see footnote 12).

formula used for the evaluation of properties for fiscal purposes does not explicitly takes into account energy efficiency parameters, it is expected that the size and significance of β_1 diminish when this proxy is introduced.

Experiment 2: *List price with a random error component*

In this experiment e is defined as a random error component, $e \sim N(k.\mu_{p^*}, k.\sigma_{p^*})$, where k is the reference upward bias in simulated list prices.

Experiment 3: *List price with an error that is proportional to the transaction value*

This experiment explores the idea that the differences between list and transaction prices increase with transaction price levels (Carrillo, 2010). To simulate a positive correlation between price levels and errors, a different error structure was defined for each transaction price quintile of existing apartments, new apartments, existing houses and new houses. Two different structures were applied to each one of the four market segments. In the first one, $e \sim N(k_i.\mu_{p^*}, k_i.\sigma_{p^*})$, where $k_i = \{(0.01), (0.02), (0.05), (0.07), (0.09)\}$. In the second case, the k parameter for the two quintiles with the most expensive properties was increased by a factor of two, so that $k_i = \{(\dots), (\dots), (\dots), (0.14), (0.18)\}$.

Experiment 4: *List price with measurement errors proportional to property size*

This experiment explores the idea that measurement errors may be positively correlated with the size of transacted dwellings (Hayunga and Pace, 2017). To implement it, measurement errors for each one of the four market segments considered above were defined to vary in line with the quintiles of gross floor area of sold properties. In this case, the measurement error, e, was defined as $e \sim N\left(k_i.\mu_{gr^*}, k_i.\sigma_{gr^*}\right)$, where gr stands for gross floor area, and $k_i = \{(0.01), (0.02), (0.05), (0.07), (0.09)\}$.

Moreover, for the investigation of the impact of erroneously measured explanatory variables on energy efficiency partial effects, the following experiment was considered.

Experiment 5: *Replacement of the model's explanatory variables*

In this case, hedonic regression models were re-estimated with the number of bedrooms in place of area variables. More concretely, $\dot{x}_k$, the number of bedrooms, is taken as being a poor measurement of the dimension of properties, which was addressed in original models through area variables (e.g., gross floor area).

The choice of the number of bedrooms as a replacement of area variables was based on two reasons. Firstly, this situation covers a typical case in which researchers do not have access to size measures other than the number of rooms or divisions. Secondly, it represents a case in which important variables (e.g., gross floor area) are replaced by a substitute with less quality (i.e., number of bedrooms). In Portugal, the two most common ways of describing residential property typologies is through the number of divisions or the number of bedrooms. The former typically includes not only the number of bedrooms but also the number of living rooms. The latter typology is usually expressed with a T, followed by the number of bedrooms. The inspection of the information obtained from this field revealed that its quality varied, thus reflecting the fact that while the data containing a T unambiguously revealed a preference for the number of bedrooms classification, single numbers could either represent divisions or number of bedrooms. Moreover, and contrary to gross floor area, the number of bedrooms is not included in the formula that calculates the amount of property tax to be paid [39]. As it does not affect taxes, it is not under the spotlight of tax payers and, as a result of this, the accuracy of the variable is affected even further.

A summary of key data features of transaction prices, fiscal appraisal values, number of bedrooms and of the 1,000 Monte Carlo runs of simulated list prices is available in the next table. Summary statistics are in accordance with expected results. For instance, simulated list prices are generally above transaction prices, with a minority of cases (0.3% to 17%) representing sales in which properties were transacted for higher values than those listed at initial phases of the buying and selling process. This data feature is often encountered in the literature comparing list and transaction prices (see, *inter alia*, Horowitz, 1992). Conversely, fiscal appraisal values are 75 to 89 percent of the cases below transaction values. Table 5.1 also allows highlighting an interesting point that has to do with the two mismeasurement scenarios introduced in the third experiment. While the average of ListPr2 is 6 to 7 percent higher than average transaction prices, the average of ListPr3 is 11 to 13 percent higher than sales prices. This result reflects the influence of most expensive properties, since the only difference in the two sets of simulated prices rest on the application of a higher measurement error for the 40 percent most expensive homes.

[39] A good description of the formula used for the fiscal appraisal of properties is available in DGI (2011).

Table 5.1: Features of transaction and simulated prices, appraisals and no. of bedrooms

	Existing apartments			*New apartments*			*Existing houses*			*New houses*		
	Mean value (€)	*Relative difference over (1)*	*% of cases below (1)*	*Mean value (€)*	*Relative difference over (1)*	*% of cases below (1)*	*Mean value (€)*	*Relative difference over (1)*	*% of cases below (1)*	*Mean value (€)*	*Relative difference over (1)*	*% of cases below (1)*
(1) TransVal	*97,695*	-	-	*149,007*	-	-	*143,783*	-	-	*179,155*	-	-
(2) FiscAppVal	*76,752*	*-21.4%*	*74.8%*	*111,038*	*-25.5%*	*89.0%*	*95,137*	*-33.8%*	*84.5%*	*124,852*	*-30.3%*	*89.2%*
(3) ListPr1	*103,553*	*6.0%*	*7.9%*	*157,937*	*6.0%*	*5.7%*	*152,421*	*6.0%*	*17.0%*	*189,978*	*6.0%*	*10.8%*
(4) ListPr2	*103,844*	*6.3%*	*0.4%*	*158,122*	*6.1%*	*0.4%*	*153,457*	*6.7%*	*1.6%*	*190,735*	*6.5%*	*0.9%*
(5) ListPr3	*108,776*	*11.3%*	*0.5%*	*165,417*	*11.0%*	*0.3%*	*161,793*	*12.5%*	*1.7%*	*200,332*	*11.8%*	*0.9%*
(6) ListPr4	*103,216*	*5.7%*	*3.1%*	*157,358*	*5.6%*	*2.0%*	*152,627*	*6.2%*	*9.4%*	*189,777*	*5.9%*	*4.7%*
(7) NBedRooms	*2.29*	-	-	*2.33*	-	-	*3.45*	-	-	*3.27*	-	-

Notes: The statistics for list prices are taken from the one thousandth simulation. For list price scenarios, point estimates are the averages over the 1,000 simulations. ListPr1: random error; ListPr2: stratified error; ListPr3: stratified error, most expensive homes receiving, on average, twice as much measurement error as in ListPr2 scenario; ListPr4: error stratified according to five area intervals.

The results of the five experiments are summarized in the next table. For comparison purposes, the energy efficient partial effects taken from the benchmark model are also included in the table.

Table 5.2: Energy efficiency parameters in different variable replacement scenarios

		Market segments			
		Existing apartments	*New apartments*	*Existing houses*	*New houses*
Benchmark scenario		*0.1183*** *(.00192)*	*0.1229*** *(.00244)*	*0.0448*** *(.00613)*	*0.0553*** *(.00624)*
Experiment1:	*FiscAppVal*	*0.0386*** *(.00108)*	*0.0485*** *(.00163)*	*-0.0085** *(.00360)*	*-0.0010* *(.00418)*
Experiment2:	*ListPr1*	*0.1103*	*0.1128*	*0.0454*	*0.0539*
Experiment3:	*ListPr2*	*0.1264*	*0.1314*	*0.0487*	*0.0617*
	ListPr3	*0.1352*	*0.1389*	*0.0547*	*0.0692*
Experiment4:	*ListPr4*	*0.1107*	*0.1149*	*0.0436*	*0.0527*
Experiment5:	*Nbedrooms*	*0.1481*** *(.00215)*	*0.1479*** *(.00281)*	*0.0940*** *(.00702)*	*0.0842*** *(.00706)*

*Notes: For List Price scenarios: Point estimates correspond to averages over the 1,000 simulations. ListPr1: random error; ListPr2: stratified error; ListPr3: stratified error, most expensive homes receiving, on average, twice as much measurement error as in ListPr2 scenario; ListPr4: error stratified according to five area intervals. For Benchmark, Fiscal Appraisal and n.º of bedrooms scenarios: * p-value < 0.05, ** p-value <0.0001. Robust standard errors in brackets*

As the table shows, when the logarithm of fiscal appraisals is used as a dependent variable (first experiment), energy efficiency coefficients are substantially smaller than those found in the benchmark scenario. This is particularly evident for existing houses, where the energy efficiency coefficient is consistent with the existence of a small price discount (-0.0085), or for new houses, where the parameter is statistically insignificant. The energy efficiency coefficients also drop for apartments. For instance, the impact of energy efficiency for existing apartments drops from 0.1183 to 0.0386. When translated into a percentage change, this implies a decrease in energy price premium from 12.5 to 3.9 percent. This is an expected outcome as the formula used for the valuation of properties for fiscal purposes does not explicitly take into account energy efficiency parameters.

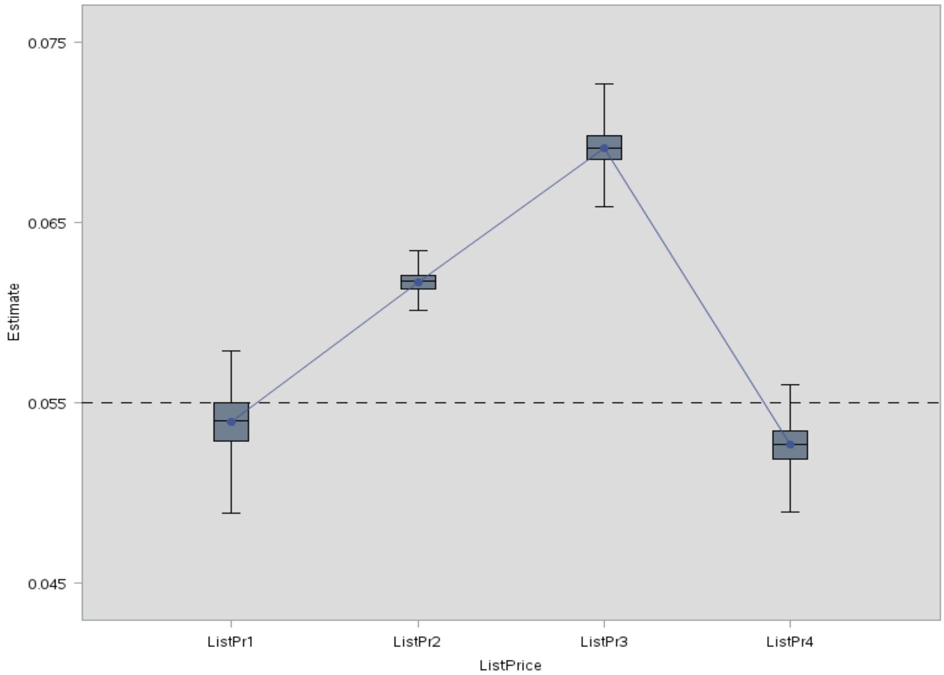
In relation to the experiments dealing with simulated list prices, it is possible to draw interesting conclusions. The first one is that there are no substantial differences between benchmark and the second experiment, where list prices differ from transaction prices by the inclusion of a random measurement error. In the case of existing houses, for example, the average coefficient obtained from the 1,000 runs is 0.0454, which is similar to the 0.0448 benchmark estimate. These results reinforce the idea that, when list prices differ from transaction prices by a random component, the OLS estimator for energy efficiency is not affected. This is an outcome that is in line with the literature dealing with this situation (see, *inter alia*, Berry, 1993: 51). A different picture emerges from the third experiment, where two stratified measurement error components were introduced. When list prices two and three are used, the OLS estimator provides higher energy efficiency partial effects than those given by the benchmark situation. For new apartments, for instance, the benchmark energy efficiency parameter estimate jumps from 0.1229 (13.1%) to 0.1314 (14.0%) and 0.1389 (14.9%) when list prices two and three are used, respectively. The results suggest that, when the dependent variable incorporates a measurement error that is positively correlated with transaction prices, the OLS estimator tends to overestimate the impact of energy efficiency on transaction prices. A possible explanation may rest on the fact that the price of a property may be positively correlated with its energy efficiency quality. The data provides some indications on this. For instance, while 53 percent of the transactions of the top 20 percent most expensive existing apartments had A and B rates, only 11 percent of the sales included in the group of the less 20 percent expensive properties were awarded with one of these energy efficiency rates [40].

To have a more complete picture of the impact of using mismeasured prices, an additional list price (ListPr4) was simulated in which errors were correlated with the size of the property (fourth experiment). The results presented in Table 5.2 show that, when this type of error was introduced, the coefficients were no longer overestimated in relation to the benchmark situation. A possible explanation may be the fact that area and energy efficiency are two characteristics that are not strongly correlated (e.g., one might have two houses with the same dimension and two completely different energy efficiency performances). As such, the introduction of this measurement error does not have a big influence in the estimates of energy efficiency partial effects.

[40] These results also raised the issue of knowing to what extend obtained energy efficiency price premiums were influenced by features, which, although present in most efficient and expensive properties, were not being taken into account in the models. The re-estimation of the hedonic models with the exclusion of the most efficient properties (i.e., A^+ and A rated dwellings) yielded similar price premiums (0.1158, 0.1063, 0.050 and 0.0486 for existing apartments, new apartments, existing houses and new houses, respectively) and provided, in this case, some evidence on price premium robustness since they are not overly influenced by the exclusion of most efficient (and also some of most expensive) properties.

A summary of the results of the experiments is given in the next figure, where the box-and-whisker plot (Tukey, 1977) displaying the 1,000 replications for new houses is shown [41]. As usual in these plots, the upper and lower hinges represent the 75th and 25th percentiles of the empirical distributions. The average of the 1,000 coefficient estimates are connected by a straight line. The dashed horizontal line signals the coefficient estimate used as the benchmark for new houses (0.055). As the figure shows, while the use of list prices two and three yield higher energy efficiency partial effects, list prices one and four provide results that cover the benchmark estimate.

Figure 5.1: Distribution of energy efficiency coefficients for new houses



The last experiment replaced area-related variables by the discrete variable on the number of bedrooms. As reported in the last row of Table 5.2, the re-estimation of the models using this specification yields the most different estimates from the benchmark situation. For instance, the impact for existing houses more than doubles, changing from 0.0448 to 0.0940. This is not a surprising outcome as this replacement subtracts explanatory power and flexibility from benchmark models. This happens not only because the variability of the number of bedrooms is much lower than the one found for area variables, but also because a non-linear relationship is eliminated from the model (i.e., the square root of area).

[41] The box-and-whisker plots for the remaining three market segments display the same pattern and are not included in the text.

The empirical results presented in this section provide evidence on the quality of the chosen models and, in addition, provide two important messages for those interested in using hedonic price models to capture the influence of energy efficiency on residential property prices. First, they show that proxy prices in the left-hand side of the hedonic price model may distort the estimates of the impact of energy efficiency on transaction prices. As such, any assessment about the impact of energy efficiency on residential property prices using either appraisals or list prices should be seen with some care. Second, the experiments highlight the importance of size variables in hedonic price models. The use of proxy size variables, such as the number of bedrooms, can lead to the introduction of a sizeable bias in energy efficiency parameter estimates.

5.3. Omission of relevant variables characterizing the quality of properties

The omission of a variable affects the statistical properties of the OLS estimator if the missing variable, which is incorporated in the error term of the model, is correlated with included explanatory variables. In this situation, the estimators for energy efficiency and other explanatory variables on prices are biased and inconsistent. As we are in a multiple regression context, the direction and magnitude of the bias is generally not known. On the other hand, if the error term and the model's covariates are not correlated, then the OLS estimator for these coefficients is unbiased. Five empirical experiments, each of which with a different omitted variables scenario, were designed to investigate the impact of the omission of variables on the OLS estimator for energy efficiency. The choice of the variables to be omitted in each experiment rested on those quality attributes that were deemed to have a good correlation with energy efficiency and that are not often available in hedonic regression studies in this area. The experiments are described below.

Experiment 6: *Omission of the central heating and/or air conditioning attribute*

In this experiment, the variable identifying the existence or inexistence of central heating or air conditioning systems is excluded from the regressions (DcSystem). Since dwellings with these systems usually have good energy efficiency standards, it is expected that this variable is positively correlated with energy efficiency and that its omission would impact on coefficient estimates.

Experiment 7: *Omission of the scenic value of the location*

This experiment omits the variables identifying the visual prominence of the location (Dscenic2 and Dscenic3), an attribute that is not often available in hedonic studies. Properties with outstanding views, such as those with a seafront location, may display large panoramic windows, a feature that may increase energy consumption needs and has a negative impact on a dwelling's energy performance.

Experiments 8 and 9: *Omission of location and construction quality*

In this experiment, the variables identifying location quality (DbadLoc and DexcpLoc) and the construction quality of residential properties (DconstQ2 and DconstQ3) were ruled out from the hedonic models. There is an *a priori* expectation that the properties located in exceptionally good locations may display above than average building, thermal and insulation standards and that these, in turn, may translate into higher energy efficiency performances

Experiment 10: *Worst-case omitted variables scenario*

The last experiment refers to the situation in which all of the above-mentioned variables were omitted from the models. In addition, the worst-case scenario was tested in smaller samples. When relevant omitted variables are excluded, parameter estimates remain biased, even when the size of the samples increases. To check this assumption, a total of 1,000 samples of 500, 1,000, 2,500, 5,000 and 10,000 observations were randomly selected with replacement from the original dataset. Since the design of the simulation exercise covered four strata - existing apartments, existing houses, new houses and new apartments -, a total of 20,000 samples were drawn from the original dataset. The samples were drawn so that an equal number of observations were obtained for each year covered by the data.

The energy efficiency coefficients that were obtained for the five omitted variable scenarios are available in the next table. Full regression results are available in Appendix VII.

Table 5.3: Energy efficiency parameters in different omitted variable scenarios

	Benchmark model	Experiment 6	Experiment 7	Experiment 8	Experiment 9	Experiment 10	All omitted (averages over 1,000 replications)				
		Central heating and/or air conditioning	Visual quality	Location quality	construction quality	All omitted (full data)	n = 500	n = 1,000	n = 2,500	n = 5,000	n = 10,000
Existing apartments	0.118** (.00192)	0.125** (.00190)	0.118** (.00193)	0.131** (.00201)	0.120** (.00192)	0.145** (.00201)	0.147	0.147	0.146	0.146	0.146
New apartments	0.123** (.00244)	0.128** (.00244)	0.123** (.00245)	0.135** (.00255)	0.127** (.00245)	0.153** (.00259)	0.162	0.162	0.163	0.163	0.163
Existing houses	0.045** (.00613)	0.051** (.00608)	0.048** (.00616)	0.055** (.00640)	0.045** (.00613)	0.070** (.00642)	0.073	0.074	0.073	0.073	0.074
New houses	0.055** (.00624)	0.061** (.00624)	0.056** (.00625)	0.066** (.00641)	0.055** (.00624)	0.073** (.00645)	0.087	0.085	0.084	0.084	0.084

*Notes: Robust standard errors in parenthesis. ** p-value <0.0001. For the replications, energy efficiency point estimates refer to the averages over the 1,000 simulations.*

The seventh and ninth experiments suggest that the exclusion of the variables measuring the visual prominence of the location and the construction quality of a property do not impact much on estimated energy efficiency partial effects. A possible explanation for this low impact may rest on market and building technology specificities where, for instance, the scenic characteristics of the location may not be valued and, as such, are not correlated with the construction of panoramic windows. In contrast, the exclusion of the dummy variable controlling for the existence of central heating and/or air conditioning systems (sixth experiment) produced an upward shift on the level of energy efficiency coefficient estimates. For existing apartments, the valuation of energy efficiency increased from 0.118 to 0.125, a result that implies an energy efficiency price premium rise from 12.5 to 13.3 percent.

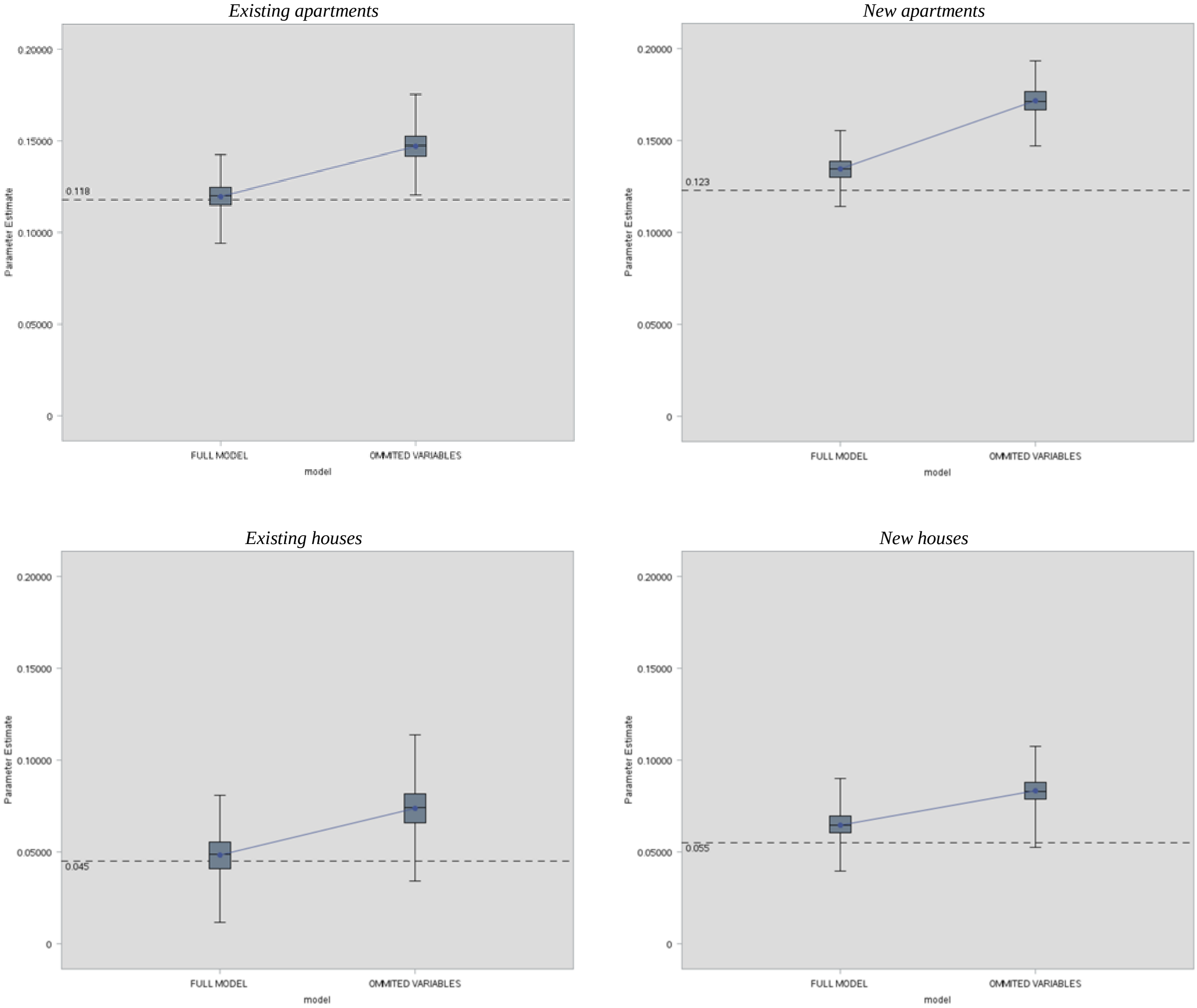
One of the largest differences between benchmark and experiment results was obtained for the omission of the quality of the location (eight experiment). When the dummy variables for this feature were not included, price premiums increased, highlighting the importance of the quality of the location in hedonic price models. The energy efficiency coefficient for new houses rises from 0.055 to 0.066, which involves a price premium increase from 5.7 to 6.8 percent [42]. As expected, the largest differences from the benchmark situation were obtained for the worst-case omitted variables scenario. For new apartments, worst-case scenario yields a 0.153 (13.1%) point estimate, which compares with the 0.123 (16.5%) coefficient given by the benchmark model. In terms of price premiums, this represents a difference of 3.4 percentage points, the largest obtained for all experiments. Finally, it is possible to see from table 5.3 that the upward shift given by the omission of relevant variables does not vary in function of the sample size. For instance, while the average coefficient estimate for existing houses over 1,000 runs of 500 observations is 0.073, for samples with 10,000 observations, this figure is 0.074.

The results suggest that, while the omission of individual variables from hedonic models may not change much the estimation of energy efficiency coefficients, the joint omission of key quality attributes can have a sizeable impact on estimates. The importance of this impact is highlighted through the comparison of the dispersion of the 1,000 energy efficiency coefficients, which were obtained for samples of 10,000 observations in benchmark (full

[42] In these situations, where a variable is omitted from the model, it may happen that its absence is compensated by changes in the coefficient estimate for energy efficiency and in other parameters. An analysis of the coefficient estimates (other than the one associated with energy efficiency) shows that coefficients remain relatively stable across the different omitted variable scenarios. See Appendix VII.

model) and worst-case (omitted variables) models. The next figure presents side by side the box-and-whiskers plots for the two situations. As usual in these plots, the representation of the data extends down to the minimum value and up to the maximum value. The dashed lines in each panel signal the price premiums that are obtained using the benchmark model specification and all data available for regression analysis.

Figure 5.2: Energy efficiency in benchmark and omitted variables cases



The overestimation of energy efficiency coefficients caused by the omission of chosen quality characteristics is easily pictured in Figure 5.2, where the benchmark and worst-case averages over the 1,000 runs are linked by a straight line. In the case of the simulation exercise involving the benchmark model, the dashed line is close or relatively close to the coefficient averages found for the 1,000 runs. However, this is not the case for the worst-case omitted variables model. In particular, for apartments, the upward shift in energy efficiency levels is such that the minimum of the box-and-whiskers plot lies above the dashed line representing the benchmark situation.

The results presented in this section provide an important message, as they warn about the consequences of leaving out from hedonic regression models variables that measure the quality of residential properties. If hedonic models do not include then, the impact of energy efficiency on residential transaction prices may be overestimated. More importantly, the upward shift can be of such magnitude that, as the case of apartments illustrates, the true impact of energy efficiency can be completely missed out.

5.4. Large samples and inflated significance levels

In the present study, which is based on a large dataset, it is extremely important to see if statistical significance is influenced by the number of transactions. To investigate this issue, it was carried out an experiment in which the hedonic regression models were rerun for a number of samples with different sizes. The experiment is described as follows.

Experiment 11: *Derivation of energy efficient coefficients for different sample sizes*

This experiment follows the same design of the tenth experiment, where 20,000 energy efficiency coefficients were calculated on the basis of 1,000 samples with sizes of 500, 1,000, 2,500, 5,000 and 10,000 observations, which were drawn for existing apartments, new apartments, existing houses and new houses. As in the tenth experiment, the samples were drawn so that an equal number of transactions were obtained for each year. As in other experiments, the benchmark specification was used as the basis for the regression work.

The results of the eleventh experiment are shown in Table 5.4. For reference, the parameter estimates that were derived for the benchmark model with all the data available for regression analysis are shown on the leftmost column of the table. The average values of energy efficiency parameter estimates are shown in the left part of the table. To provide information on the dispersion of the 1,000 point energy efficiency parameter estimates, standard deviations were calculated and provided underneath each average value. The counts of statistically significant positive and negative coefficients for each sample size experiment are also available from the rightmost section of the table.

Table 5.4: Energy efficiency parameters derived from different sample sizes

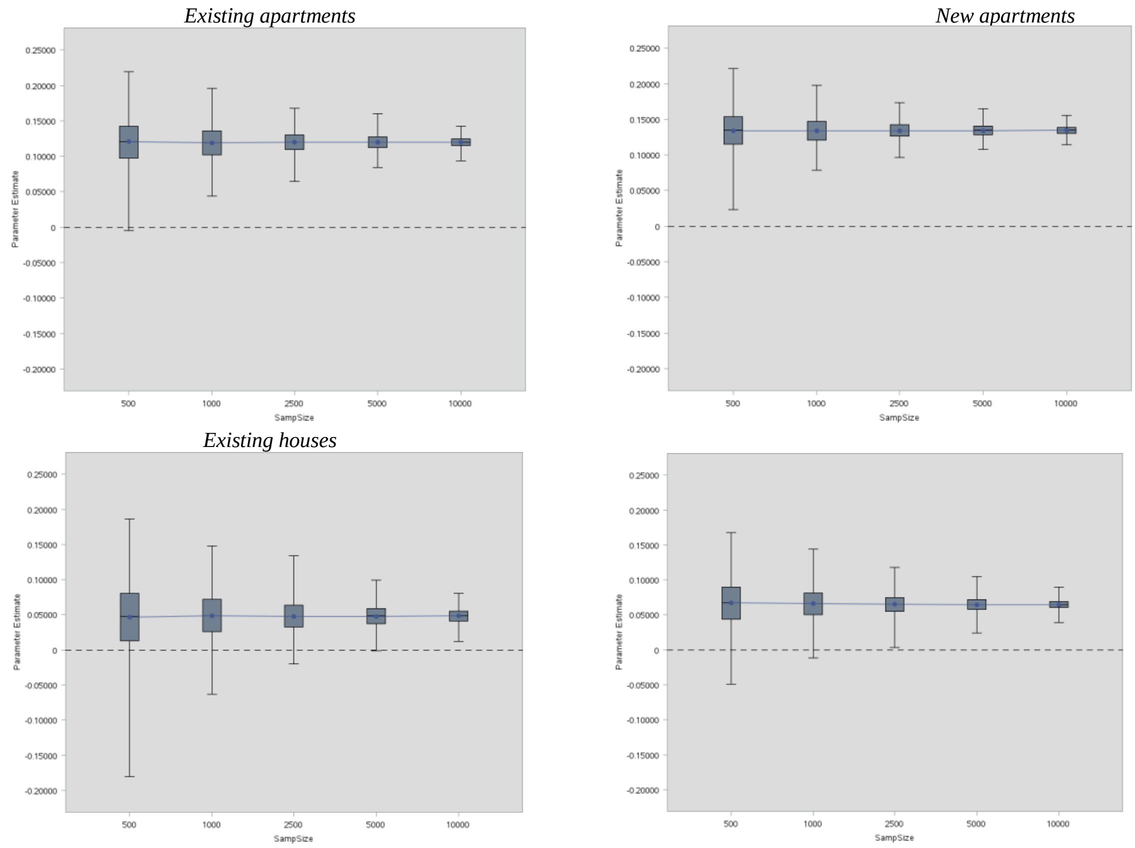
	Coefficient (benchmark estimate)	Coefficient estimate, average over 1,000 replications					Sign of parameter (no. of times, if statist. significant)									
							n = 500		n = 1,000		n = 2,500		n = 5,000		n = 10,000	
		n = 500	n = 1,000	n = 2,500	n = 5,000	n = 10,000	(+)	(-)	(+)	(-)	(+)	(-)	(+)	(-)	(+)	(-)
Existing apartments	*0.118*	*0.121* (0.034)	*0.119* (0.025)	*0.120* (0.015)	*0.120* (0.011)	*0.120* (0.008)	*947*	*0*	*996*	*0*	*1,000*	*0*	*1,000*	*0*	*1,000*	*0*
New Apartments	*0.123*	*0.134* (0.028)	*0.134* (0.019)	*0.134* (0.012)	*0.134* (0.009)	*0.135* (0.006)	*996*	*0*	*1,000*	*0*	*1,000*	*0*	*1,000*	*0*	*1,000*	*0*
Existing houses	*0.045*	*0.047* (0.052)	*0.049* (0.035)	*0.048* (0.023)	*0.048* (0.016)	*0.048* (0.011)	*169*	*4*	*272*	*0*	*572*	*0*	*849*	*0*	*982*	*0*
New houses	*0.055*	*0.067* (0.033)	*0.066* (0.024)	*0.065* (0.015)	*0.065* (0.010)	*0.065* (0.007)	*551*	*0*	*814*	*0*	*993*	*0*	*1,000*	*0*	*1,000*	*0*

Notes: For the replications, energy efficiency point estimates refer to the averages over the 1,000 simulations. Where relevant, standard deviation of simulated coefficients provided between brackets.

There are a few points from Table 5.4 that are worthwhile noting. First, with the exception of existing houses, the number of statistically significant energy efficiency coefficients is substantial, even for very small sample sizes [43]. This constitutes a strong indication that the statistical relevance of this characteristic is not inflated by the sample size. Second, the results also reveal that the cases of statistically significant coefficients with conflicting signs, which could suggest the existence of a model specification error, are the exception rather than the rule. These only appear for existing houses and for the smallest sample size considered. Third, average energy efficiency coefficients are very stable across the different sample sizes. For instance, the benchmark estimate for existing apartments is 0.118, which is similar to the 0.121 average found for the 1,000 rounds of samples with 500 observations. This result is remarkable, especially if it is taken into account that 500 observations represent less than 0.4 percent of the total transactions of this market segment. Finally, as expected, the spread of the coefficients substantially decreases across the different sample sizes. For new houses, for example, the spread for the estimates based on samples with 10,000 observations (0.007) is approximately one fifth of the one that is obtained for samples with 500 observations (0.033). The idea that the dispersion in coefficient estimates decreases as the size of the sample increases is clearly shown in the next figure, where the box-and-whiskers plots of the 20,000 energy efficiency coefficients are provided.

[43] Existing houses also perform reasonably well since energy efficient coefficients are statistically significant for more than half of the runs using sample sizes of 2,500 observations (i.e, for samples with 7.5 percent of all available existing houses transactions).

Figure 5.3: Distribution of energy efficiency coefficients for different sample sizes



The plots also stress the contrast between the valuation of energy efficiency for apartments and houses. While for apartments results clearly indicate the existence of a price premium (see top panels in Figure 5.3) for houses, the plots cover coefficient estimates that have either a zero or negative value. Price discounts are more common for existing houses and for smaller sample sizes. For samples of 500 observations, a total of 179 price discounts were found, with a maximum at -0.1802 (-16.5%) and a minimum very close to zero [44]. The negative coefficients are not concentrated on this maximum (i.e., coefficient estimates are evenly distributed). For the situations based on samples of 1,000 and 2,500 observations, a total of 85 and 16 price discounts were obtained for existing houses, respectively. New houses also presented some price discounts (21 and 5 cases for the 1,000 draws of 500 and 1,000 observations, respectively), with a maximum obtained for the smallest sample size (-4.8%). On the other hand, the runs for apartments produced only one price discount in a sample of transactions of existing dwellings [45]. These results point out for a significant difference

[44] The maximum is statistically insignificant. Of the 179 negative coefficients, only 4 are statistically significant.

[45] This coefficient estimate is statistically insignificant.

between apartments and houses and warn about the possibility of the existence of an energy efficiency price discount. This is more evident for existing houses, where this outcome appears for samples with 5,000 observations, which represent approximately 15 percent of all available transactions. These results also suggest that statistically significant price discounts may be obtained from hedonic studies, particularly if a small sample size is used to draw conclusions on the impact of energy efficiency on residential property prices. Interestingly, the plots do not picture a noticeable difference between new and existing residential properties.

The impact of using different sample sizes on the quality of regression results was also investigated through the percentages of *p*-values that display a value above the 0.05 threshold. Assuming that the parameter under analysis is, in reality, different from zero, these percentages would give an approximation of the probability of accepting a wrong decision or, what is the same, an estimation of its type II error, which is associated with the power of a statistical test [46]. This should ideally be small, even for small sample sizes. The next table provides these percentages for energy efficiency (DENERGYAB), existence of central heating and/or air conditioning equipment (DENERGYAB) and the scenic value of the location (DSCENIC3) coefficients [47].

Table 5.5: Percentage of statistically insignificant energy efficiency coefficients at the 5% level

	(*)	*Sample size considered in the replications*				
		n = 500	*n = 1,000*	*n = 2,500*	*n = 5,000*	*n = 10,000*
Energy efficiency	*(1)*	*5.3%*	*0.4%*	*0.0%*	*0.0%*	*0.0%*
	(2)	*0.4%*	*0.0%*	*0.0%*	*0.0%*	*0.0%*
	(3)	*82.3%*	*72.8%*	*42.8%*	*15.1%*	*1.8%*
	(4)	*44.9%*	*18.6%*	*0.07%*	*0.0%*	*0.0%*
Central heating and/or air conditioning equipment	*(1)*	*66.3%*	*45.5%*	*13.1%*	*1.2%*	*0.0%*
	(2)	*36.3%*	*9.4%*	*0.1%%*	*0.0%*	*0.0%*
	(3)	*71.5%*	*56.2%*	*20.4%*	*2.6%*	*0.0%*
	(4)	*54.8%*	*28.4%*	*2.9%*	*0.1%*	*0.0%*
Scenic value of the location	*(1)*	*44.8%*	*38.5%*	*16.7%*	*2.8%*	*0.0%*
	(2)	*60.7%*	*44.7%*	*14.5%*	*1.1%*	*0.0%*
	(3)	*56.8%*	*55.8%*	*36.1%*	*16.1%*	*2.1%*
	(4)	*69.2%*	*67.8%*	*44.7%*	*18.6%*	*2.3%*

Notes: (*)(1) Existing apartments; (2) New apartments; (3) Existing houses; (4) New houses. Percentage of p-values higher than 0.05 in 1,000 trials.

[46] The power of a test is equal to one minus type II error. Type I error is fixed (i.e., 0.05).
[47] For the sake of space, it was chosen not to include more coefficient results.

As shown in the table, the percentage of cases in which energy efficiency coefficients are statistically insignificant drops as the sample size increases. For draws of 5,000 observations, the percentage of statistically insignificant coefficients drops considerably. If seen as an approximation to type II error, these results imply that, the probability of rejecting the null hypothesis under an alternative hypothesis is very high, even at relatively small sample sizes. This reinforces the idea that the statistical significance of the variables is not inflated by sample sizes. The second important note is that the overwhelming majority of the parameters that do not have average *p*-values lower than 0.05 either represent situations of variables with a small variability in sampled data (e.g., regions in which transactions are scarce) or that, although not being statistically significant, do have an economic justification to remain in the model (e.g., dummy variables for the year, which reflect the situation in which the prices have not changed from one year to another).

The empirical results given in this section stress the stability of energy efficiency coefficient across different sample sizes and provide no indication that significance levels of energy efficiency partial effects have been inflated by the number of observations. The exercise on inflated significance levels supports the finding that the market makes a clear separation between houses and apartments in terms of energy efficiency price premiums. However, the idea that new residential properties reward energy efficiency differently from existing dwellings (Chapter 4), is not confirmed by the results presented in this chapter.

5.5. Cross-country comparisons

To carry out a cross-country qualitative assessment of the impact of energy efficiency on residential property prices, benchmark hedonic regression models were re-estimated with changes in the specification which increased the degree of comparability between the results obtained in this work and those found for Finland (Fuerst et al., 2016), England (Fuerst et al., 2015), Ireland (Hyland et al., 2013), the Netherlands (Brounen and Kok, 2011) and Portugal (Ramos et al.; 2015b) [48]. The cross-country qualitative experiment that was designed to investigate these issues is described as follows.

Experiment 12: *Cross-country impact of energy efficiency on transaction prices*

This experiment involved the inclusion of two main changes in the specification of the hedonic regression model, which increased the comparability across the different studies. The first one had to do with the explanatory variable measuring energy efficiency levels, which was modified so that it could replicate the energy measurement scales used in the regressions of comparison studies. The D rate was used as the reference class in all regressions. The second change had to do with the choice of the model's dependent variable, which in some studies is the natural logarithm of price (e.g., Hyland et al., 2013) and in others, the natural logarithm of price per square meter (e.g., Brounen and Kok, 2011). Since the design of the experiment covered five studies and four market segments, a total of twenty regressions were run.

With this experiment, it was possible not only to check the consistency of energy efficiency results in face of external information but also to investigate the degree to which the results obtained for a southern country such as Portugal were consistently higher or lower than those given for northern European markets. The energy efficiency partial effects of the twenty hedonic price models are shown in Table 5.6, which provides the impact of energy efficiency on prices when the logarithm of the price level was used, and in Table 5.7, which gives the impact for the logarithm of price per square meter. For reference, the results of the country studies are also included in the second leftmost column of the tables.

[48] The choice of these studies was naturally restricted to those applying an EPC label in which a discrete scale was used as the indication of energy performance standard.

Table 5.6: Impact of energy efficiency on the log of price level

Label Classes	*Hyland et al. (2013)*	***Replication of energy label class scheme for...***				*Average* $^{(+)}$
		Exist. Apart.	*New Apart.*	*Exist. Houses*	*New Houses*	
A	*9.7%**	*23.6%**	*20.7%**	*3.0%*	*10.0%**	*14.3%*
B	*5.3%**	*13.5%**	*10.0%**	*5.2%**	*4.8%**	*8.4%*
C	*1.7%**	*1.5%**	*-1.4%**	*3.3%**	*1.8%**	*1.3%*
D		*No estimate (hold-out class)*				
E	*-0.4%*	*0.9%**	*-4.9%**	*-0.8%*	*-3.0%**	*-2.0%*
F/G	*-10.1%**	*-1.8%*	*-0.4%*	*-2.9%**	*-4.3%**	*-2.4%*
n	*15,060*	*149,920*	*59,410*	*33,282*	*13,533*	
$R^{2\,(.)}$	*0.65*	*0.68*	*0.74*	*0.67*	*0.75*	-
Reset, p-val.	-	*0.800*	*0.219*	*0,095*	*0.043*	

Label Classes	Fuerst *et al.(2016)*	***Replication of energy label class scheme for...***				*Average* $^{(+)}$
		Exist. Apart.	*New Apart.*	*Exist. Houses*	*New Houses*	
ABC	*1.3%**	*5.5%**	*7.0%**	*4.1%**	*4.0%**	*5.2%*
D		*No estimate (hold-out class)*				
E	*0.0%*	*1.0%*	*-4.7%**	*-0.8%*	*-3.0%**	*-1.9%*
FG	*0.0%*	*-1.9%*	*-0.5%**	*-2.9%**	*-4.5%**	*-1.5%*
n	*6,194*	*149,920*	*59,410*	*33,282*	*13,533*	
$R^{2\,(.)}$	*0.933*	*0.67*	*0.72*	*0.67*	*0.75*	-
Reset, p-val.	-	*0.030*	*0.066*	*0.092*	*0.041*	

Notes: (*) Based on a coefficient estimate that is statistically significant at the 5% level. (+) Arithmetic average of the four replications.
(.) The Adjusted R^2 is the correct measure for the comparison of models with the same dependent variable and different number of explanatory variables. However, for the sample dimensions considered, the difference between this measure and R^2 is negligible. As not all studies provide the Adjusted R^2, it was chosen to show in these tables the R^2.

Table 5.7: Impact of energy efficiency on the log of price per square meter

Label Classes	Fuerst et al. (2015)	Replication of energy label class scheme for... Exist. Apart.	New Apart.	Exist. Houses	New Houses	Average (+)
AB	*5.1%**	*13.4%**	*9.4%**	*5.2%**	*6.1%**	*8.5%*
C	*1.8%**	*0.8%**	*-3.5%**	*3.1%**	*1.9%**	*0.6%*
D			*No estimate (hold-out class)*			
E	*-0.7%**	*1.1%**	*-5.0%**	*-0.4%*	*-2.9%**	*-1.8%*
F	*-0.9%**	*-1.1%*	*-2.6%*	*-2.0%**	*-2.5%*	*-2.0%*
G	*-6.6%**	*-3.0%*	*4.8%*	*-4.2%**	*-9.6%**	*-3.0%*
n	*333,095*	*149,920*	*59,410*	*33,282*	*13,533*	-
Adj.$R^{2(.)}$	*0.693*	*0.559*	*0.625*	*0.477*	*0.575*	
Reset, p-val.	-	*0.000*	*0.000*	*0.000*	*0.000*	

Label Classes	Brounen and Kok (2011)	Replication of energy label class scheme for... Exist. Apart.	New Apart.	Exist. Houses	New Houses	Average (+)
ABC	*3.7%**	*4.5%**	*6.0%**	*4.7%**	*5.6%**	*5.2%**
DEFG			*No estimate (hold-out class)*			
n	*31,993*	*149,920*	*59,410*	*33,282*	*13,533*	-
$R^{2\,(.)}$	*0.527*	*0.550*	*0.610*	*0.478*	*0.575*	
Reset, p-val.	-	*0.000*	*0.000*	*0.000*	*0.000*	

Label Classes	Ramos et al. (2015b)	Replication of energy label class scheme for... Exist. Apart.	New Apart.	Exist. Houses	New Houses	Average (+)
ABC	*5.9%**	*4.9%**	*4.9%**	*4.2%**	*4.4%**	*4.6%*
D			*No estimate (hold-out class)*			
EFG	*-4.0%**	*0.9%**	*-4.3%**	*-1.3%**	*-3.4%**	*-2.0%*
n	*31,993*	*149,920*	*59,410*	*33,282*	*13,533*	-
$R^{2(.)}$	*0.527*	*0.550*	*0.610*	*0.478*	*0.575*	
Reset, p-val.	-	*0.000*	*0.000*	*0.000*	*0.000*	

Notes: (*) Based on a coefficient estimate that is statistically significant at the 5% level. (+) Arithmetic average of the four replications.
(.) The Adjusted R^2 is the correct measure for the comparison of models with the same dependent variable and different number of explanatory variables. However, for the sample dimensions considered, the difference between this measure and R^2 is negligible. As not all studies provide the Adjusted R^2 (Fuerst et al., 2015 give only this measure), it was chosen to show in these tables the R^2.

The results clearly identify a higher price premium for apartments, a finding that is invariant to the change of the energy efficiency scale and that is stable across all regressions. However, the existence of a higher energy efficiency price premium for new properties is not clear across the re-estimations. In Table 5.6, for instance, while the price premium of A, B and C-rated apartments increases from 5.5 to 7.0 percent when one moves from existing to new properties, it decreases from 4.1 to 4.0 percent in the case of houses. Energy efficiency is essentially rewarded by properties exhibiting A and B ratings. This can be seen in the top figure of Table 5.6., where C-rated even display a price discount (-1.4%), and A and B rates show price premiums lying between 23.6 and 3.0 percent. This result is further supported by the top figure of Table 5.7, where it is possible to see that the price premiums associated with A or B properties are always substantially higher than those attached to C rated properties.

Interestingly, the results suggest that energy efficiency seems to display higher price premiums in Portugal than in Ireland, Finland, England or the Netherlands. This is a finding that was also identified in Ramos et al. (2015b), who pointed the possible higher awareness of the EPC label and the existence of higher energy costs in Portugal as possible explanations for this situation. However, the magnitude of their price premiums is higher than those estimated in this thesis. For instance, while these authors estimate a 5.9 percent price premium for A, B and C rated residential properties, the highest price premiums provided by re-estimation of the hedonic price models using a dummy variable signaling A, B and C rates is 4.9 percent for apartments (see the bottom figure in Table 5.7). A possible explanation for this difference may rest in the fact that Ramos et al. (2015b) use list prices as a proxy of transaction prices. As it was already illustrated in Section 5.2, the use of list prices as the model's dependent variable may lead to the overestimation of the impact of energy efficiency on residential property transaction prices. Moreover, the overall fit of the regressions is in line with those found in similar studies. In addition, while the regressions using the logarithm of the price level as the dependent variable pass the Reset test at the 5 or 1 percent confidence levels, the specifications with the logarithm of the price per square meter are rejected by this specification test, a fact that reinforces the idea that the choice of the dependent variable was well done.

5.6. Summary

This chapter presents the results of 12 experiments that investigated the coherence and sensitivity of energy efficiency coefficients to different data and estimation contexts. In doing this, it was possible to draw four main conclusions, which are of relevance to all those interested in this research area.

First, the replacement of transaction prices by simulated list prices in hedonic price models has shown that, under reasonable assumptions about the difference between these two measurements, the use of a proxy variable as a regressand leads to the overestimation of the price premium associated with energy efficiency. Thus, any assessment about the impact of energy efficiency on residential property prices using list prices needs to be seen with some care. This is important since transaction prices are not always available to researchers. Second, the experiments using different omitted variables scenarios suggest that, while the omission of individual variables may not bias much the estimation of energy efficiency coefficients, the joint omission of key quality attributes can lead to a sizeable overestimation of the impact of energy efficiency on property prices. This provides an important message to researchers not only because some of the variables used are often not available for research, but also because the upward omitted variables shift can be of such magnitude that the correct impact of energy efficiency on prices can be completely missed out. Third, the exploration the large sample available for research has given important indications as to the soundness of the results. For instance, the number of statistically significant energy efficiency coefficients is substantial, even for draws of very small sample sizes, and the average partial effect estimates are very stable across the different sample sizes considered in the eleventh experiment. Finally, the qualitative comparison across different studies and markets suggests the existence of higher price premiums in the Portuguese market than in northern European countries. A greater EPC label awareness and the existence of higher energy costs in Portugal are possible explanations for this situation. Interestingly, the magnitude of the price premiums was found to be smaller than those estimated for Portugal by Ramos et al. (2015b) and also for Spain by Ayala et al. (2016), who report energy efficiency price premiums of 10.3 percent for the three most efficient energy efficiency rates, an outcome that could be associated with the use of proxy prices instead of transaction prices.

The results of the experiments stressed the suitableness of chosen hedonic price models. They responded in a coherent way to the different replacement and omitted variables scenarios and did not show evidence of inflated significance levels. Moreover, they confirmed the existence of a clear difference between apartments and houses, with the latter residential unit type showing lower price premiums than the former. However, the experiments did not find any substantial differences in the way energy efficiency is valued by new and existing dwellings and thus do not confirm the initial findings on this respect based on full OLS results (see Chapter 4).

Chapter 6 Conditional and unconditional quantile regression energy efficiency partial effects

6.1. Introduction

Instead of focusing on the effect of energy efficiency changes on the mean of residential property prices, researchers and policy makers may be interested in analyzing the impact of these changes on particular quantiles of the price distribution. Information by quantile may be crucial in the design of policy instruments directed to market segments with particular energy efficiency insufficiencies. For instance, it may be relevant to know that energy efficiency entails a price discount at the lowest quantiles of the price distribution. With this information at hand, policy makers could tailor appropriate incentives targeting low income purchasers (i.e., those that are more likely to purchase least expensive dwellings), and thus increase global energy efficiency standards in a cost-effective way.

The main objective of this chapter is to characterize the impact of a change in energy efficiency standards along the distribution of residential property prices. Since the seminal paper by Koenker and Basset (1978), conditional quantile regression (CQR) has been commonly used in various fields of applied work to provide evidence on this type of impacts. Examples include applications in many areas, ranging from demand analysis (Deaton, 1997) to finance (Bassett and Chen, 2001). An excellent recent account of the applicability of the method is Koenker (2017). Despite the pervasive use of CQR in empirical applications, this framework provides a narrower interpretation of quantile impact changes on the distribution of the dependent variable, as these are conditional on the values and the set of the model's chosen covariates and are often not interpretable in a policy or population context (see, *inter alia*, Borah and Basu, 2013). With OLS, conditional average partial effects can be interpreted as unconditional (or generalizable) partial effects through the law of iterated expectations.

Unfortunately, this equality does not hold for quantile regression analysis and results stemming from CQR and unconditional quantile regression (UQR) have a different interpretation. The difference between the two effects rests on the fact that the former results are provided conditionally on observed characteristics and the latter are not. For instance, a property with high energy efficiency may be located in the upper part of its conditional distribution (i.e., the distribution for the properties with the same area, age and other observable characteristics) and, at the same time, be in the middle or even lower part of the overall dwelling transaction price distribution.

Firpo et al. (2009) and Firpo (2007) constitute two prominent examples of UQR estimators. While the former approach is based on the notion of recentered influence functions, the latter is a reweighted version of the estimation procedure proposed by Koenker and Basset (1978) [49]. Notwithstanding its interest for policy issues, the UQR framework has not been widely used. There are at least two reasons that can explain this situation. First, for the lowest and highest quantiles, which are usually the most interesting to analyze, researchers need high quality data and a large number of observations in the neighborhood of the quantiles under study. As this is often not the case (e.g., in a typical bell curve, there is a higher density in the middle of the distribution of the outcome variable and, as such, there are not that many observations at the extremes), standard errors may become large and regression results for the lowest and highest quantiles may become unreliable. Second, the novelty of the UQR helps explain the reduced number of empirical applications based on this framework. Although quantile regression was introduced in the 1970's and the median regression has been known since the 18th century, UQR has only started to receive more attention after the influential work by Firpo et al. (2009).

The number of UQR studies has nevertheless been increasing over time. Wealth and labour economics, on the one hand, and health economics, on the other, are two of the areas in which it is possible to see some studies using the UQR framework. Fournier and Koske (2013), for instance, focus on the relationship between the level of public employment and earnings inequality in five different countries. Galego and Pereira (2014) apply UQR to analyze the determinants of regional wage gaps in Portugal. Maclean et al. (2013) study the heterogeneous response of smokers to state cigarette tax, and Jolliffe (2011) analyzes the

[49] The influence function is covered in the literature dealing with robust statistics and is used to assess the influence of removing or adding an observation on the value of a statistic (Borah and Basu, 2013).

relationship between income and corporal weight, as measured by the body mass index. Other areas include agricultural economics, where Mishra et al. (2015) explore the effects of off-farm income on food expenditures of rural Bangladeshi households, and environment economics, where Peeters et al. (2017) address the heterogeneity of the impact of soil pollution on farmland prices of in Belgium. Finally, Borah and Basu (2013) highlight the differences between conditional and unconditional quantile regression through a simulation study and an empirical application on the effects of a change in the determinants of prescribed medication adherence amongst Alzheimer's disease patients.

Despite a few studies applying CQR to study the importance of some attributes in the real estate market (e.g., Mak et al., 2010; Zhang and Yi, 2017), the impact of energy efficiency across the distribution of property transaction prices has never been studied before. This is the first study addressing this issue. Its originality lies not only in the use of the CQR to analyze energy efficiency partial effects across the price distribution, but essentially in the application of the UQR framework, which has never been used in this research context. The latter quantile regression results, for which Firpo's (2007) reweighting estimator was employed, offers generalizable results, which are of interest for policy makers and researchers in this area.

The findings support the idea that the impacts of greater energy efficiency are not uniformly positive across the unconditional quantile regressions for all dwelling types. For houses located at or below the 0.2th price quantile, there is even clear evidence for the existence of price discounts associated with greater energy efficiency standards. Moreover, although energy efficiency is positively rewarded across the entire price distribution for apartments, it is possible to observe a reduction in the magnitude of the price premiums at the highest quantiles of the distribution. These findings provide a more complete picture of the impact of energy efficiency on residential property prices, compared to the traditional OLS analysis, whose results are masked by the response at the mean of the price distribution. More importantly, by emphasizing the idea that the effect of energy efficiency is different across price segments, these findings provide an additional explanation as to the reason why price discounts, rather than price premiums, may appear in hedonic regression studies on this topic (e.g., Yoshida and Sugiura, 2015).

This chapter is organized into four main sections. Section 6.2 presents the conditional and unconditional regression frameworks. Section 6.3 provides descriptive statistics on the heterogeneity of energy efficiency and other dwelling attributes across residential property price quartiles. Section 6.4 presents and compares OLS, CQR and UQR results for existing apartments, existing houses, new apartments and new houses. Section 6.5 looks into the coherence of derived estimates. This is done not only through the few studies that use CQR in which it is possible to obtain coefficients for some key covariates of the hedonic price model (e.g., area), but also through the re-estimation of UQR results with different energy efficient scales and dependent variables that were used in Chapter 5 to increase cross-country comparability of results. Finally, the last section provides a summary of the main findings.

6.2. Conditional and unconditional quantile partial effects

Let $P = \ln(p)$ be the outcome variable of interest where, as defined before, $\ln(p)$ stands for the natural logarithm of a transaction price. The interest of the present analysis lies in understanding the effect of the binary variable E, which measures energy efficiency, on the continuous variable p^* at the τth quantile of its distribution, where $\tau \in (0,1)$ [50].

Following the seminal paper by Koenker and Basset (1978), which extends the classical OLS framework to quantile analysis, if a set of relevant explanatory variables $Z = \{E, X\}$ is observed, the conditional effect of E (and of other covariates) on P at the τth price quantile, $\frac{dQ_\tau(P|Z)}{dE}$, is defined as the argument that solves the following minimisation problem:

$$\hat{\beta}_{\tau,CQR} \equiv \arg\min_{\beta} \sum_{i=1}^{n} \rho_\tau.(p_i - z_i'\beta_\tau), \tag{6.1}$$

where $\rho_\tau = u_i.(\tau - \mathbb{I}\{u_i < 0\})$ corresponds to the reweighting (*alias* check) function of the residuals u_i, and $\mathbb{I}$ is an indicator function assuming the value 1 when the condition between brackets holds, and 0 otherwise. Unless $\tau = 0.5$, which gives rise to median or least absolute deviation (LAD) regression, the reweighting of residuals is done asymmetrically by the check function. The CQR is useful when the effect of a change in a covariate, as represented by the β in regression model, is not homogeneous or constant across the conditional quantiles of

[50] The main interest of quantile regression analysis is on residential property prices, not on P, its natural logarithm transformation. In practice, the two can be used interchangeably since the results stemming from quantile regression analysis are invariant to this monotonic transformation.

the distribution of the outcome variable of interest. In fact, the CQR estimator unveils potential differences in the magnitude and sign in quantile coefficients and thus constitutes an advantage over OLS regression analysis, which focuses on the mean response effect. A classic example of CQR usefulness is the interest in knowing whether the introduction of an additional year of compulsory schooling reduces income inequality in a population of workers.

However, researchers and policy makers are often not concerned with the overall impact of a change on a theoretical distribution of a variable of interest but rather on observing the impact of that change on individuals (i.e., on the unconditional distribution of the variable of interest). Unfortunately, unlike the analysis done at the mean, where conditional and unconditional means are the same, conditional quantile results are generally different from unconditional quantile partial effects. As mentioned before, the difference between the two effects stem from the fact that the latter results are provided conditionally on observed characteristics and the former are not. These differences change the interpretation of CQR and UQR results and generally imply that $\frac{dQ_\tau(P|Z)}{dE} \neq \frac{dQ_\tau(P)}{dE}$. A good description of the differences between conditional and unconditional quantile partial effects is available in Borah and Basu (2013). An excellent illustrative simulation example of why standard CQR results can be misleading and misinterpreted is provided in appendix 1 of Peeters et al. (2017).

The covariate effect on the τth unconditional quantile can be obtained using an approach proposed by Firpo et al. (2009), which consists of running a regression of the (recentered) influence function of the unconditional quantile on the explanatory variables. However, while the Firpo et al. (2009) estimator is suited to study the impact of marginal changes for continuously distributed covariates, it is not appropriate for contexts in which a researcher is interested in estimating the full unconditional quantile partial effect of a dichotomous covariate, such as E, on a continuous variable, such as P. This is acknowledged in Firpo et al. (2009: 962) and in Rothe (2012: 2271-2), which note that in presence of a dummy covariate, the proposed estimator would not give the full effect of a change from 0 to 1 in E, but rather an estimate of a small (marginal) change in the probability that this binary variable is equal to one.

In this context, the unconditional partial effect of a change in E at the τth price quantile would need to be calculated differently. Assuming that the data generating mechanism of E is based on the set of X exogenous observable characteristics [51], an asymptotically exact estimate of the quantile effect caused by a change (the treatment) in a dummy variable from 0 to 1 is provided by Firpo (2007), which uses propensity scores (Rosenbaum and Rubin, 1983), or $pr(x) = Pr(E = 1|X = x)$, the probability of E having an A or B EPC rate given a set of observed attributes, to reweight the sum of check functions in Koenker and Basset's (1978) minimisation procedure presented in (6.1) above. Thus, the effect of a change at the τth price quantile from 0 to 1 in E on P, $\hat{\Delta}_\tau \equiv \hat{p}_{1,\tau} - \hat{p}_{0,\tau}$, reflects the difference between the distribution for most energy efficient transacted properties and the distribution of less energy efficiency dwellings where, for $l = (0,1)$, the $p_{l,\tau}$ are obtained as:

$$\hat{p}_{l,\tau} \equiv \arg\min_{\beta} \textstyle\sum_{i=1}^{n} \omega_{l,i} . \rho_\tau . \left(p_i - \underbrace{z_i' \beta_\tau}_{\hat{p}_{l,\tau}} \right), \tag{6.2}$$

where $\widehat{\omega}_{1,i}$ and $\widehat{\omega}_{0,i}$, the estimated propensity-score weights, are calculated respectively as $\frac{E_i}{n.\widehat{pr}(x_i)}$ and $\frac{1-E_i}{n.(1-\widehat{pr}(x_i))}$, and n refers to the number of sales. The estimation of the unconditional quantile treatment effect in the case of a dichotomous variable follows a two-step approach where, in the first step, propensity scores are obtained nonparametrically and, in the second step, the $p_{j,\tau}$ are obtained through the minimisation of a reweighted sum of check functions.

Unfortunately, the SAS software (SAS Institute Inc, 2015), which was employed in the previous chapters, does not provide procedures for the implementation of UQR methods. In this context, STATA (StataCorp, 2017) was used and the *ivqte* command (Frölich and Melly, 2010), which provides results for the estimator proposed by Firpo (2007) [52]. Data analysis, CQR and UQR were applied separately for existing apartments, new apartments, existing houses and new houses. Likewise, the models that were used to test Chapter's 4 Hypothesis 3, the main research question addressed in this work, were used as benchmark specifications in quantile regression work.

[51] Frölich and Melly (2013) provide an UQR estimator in which the set of the X covariates are endogenous. Koenker and Basset's (1978) CQR estimator also assumes exogeneity.
[52] In StataCorp (2017), the Firpo et al. (2009) estimator is available through the *rifreg* command. This approach was not, as explained before, used to derive unconditional quantile results for the dichotomous variable measuring energy efficiency.

6.3. Preliminary evidence on heterogeneous energy efficiency effects

Table 6.1 presents the percentage of transacted residential properties with A or B EPC rates for the four covered market segments and by price quartile group. The data pertaining to each quartile group was found after ranking all observations according to the value of the transaction. The quartile groups have an equal or, when the division by four did not yield an integer, an almost equal number of transactions. The percentage of energy efficient properties is calculated simply by calculating the average of the dummy variable signaling A and B EPC rates.

Table 6.1. Most energy efficient properties, percentages by price quartile group

Transaction price quartile group	1st	2nd	3rd	4th	All
(1) Existing apartments	11.9	21.1	33.8	51.0	29.4
(2) New apartments	49.0	67.4	76.0	83.1	69.0
(3) Existing houses	13.8	13.2	22.2	34.4	20.9
(4) New houses	18.7	32.3	43.5	53.1	39.9

The table shows that the percentage of A and B energy efficient properties increases as one moves from less to more expensive price quartile groups. For instance, while only 11.9 percent of the first quartile group of existing apartment transactions bear an A or B EPC rate, as many as 51.0 percent of the transactions included in the quartile with most expensive existing apartment transactions had one of these rates attributed.

Evidence on the existence of differences along the price distribution is also observed when other variables are analyzed. Table 6.2 presents mean values by quartile price group on a selected group of residential property characteristics. As with the percentage of properties bearing A and B rates, the mean area and the percentages of properties displaying prominent visual attributes and bad location characteristics change monotonically as one moves from one quartile to another. For instance, the mean area for new apartments increases from 89.7 square meters, which is obtained for the 25 percent of less expensive properties, to 140.8 square meters, which was found for the top 25 percent of most expensive properties. Conversely, the number of properties associated with a bad location decreases as transaction prices increase (e.g., from 59.7 to 9.4 percent in the case of new houses).

Table 6.2. Selected property characteristics, mean values by price quartile group

Transaction price quartile group	(*)	1st	2nd	3rd	4th	All
Area, in square meters	(1)	77.6	88.4	97.5	120.5	96.0
	(2)	89.7	105.7	116.4	140.8	113.1
	(3)	94.5	128.3	161.3	209.1	148.3
	(4)	111.9	165.8	184.3	212.5	168.6
Visual quality dummy	(1)	0.2	0.3	0.3	1.5	0.6
	(2)	0.8	0.5	1.2	3.5	1.5
	(3)	0.1	0.2	0.3	2.0	0.6
	(4)	0.2	0.5	0.6	1.8	0.8
Bad location dummy	(1)	5.3	4.7	2.9	1.0	3.5
	(2)	12.2	6.2	2.6	1.3	5.6
	(3)	49.3	38.5	25.5	9.8	30.8
	(4)	59.7	49.8	26.6	9.4	36.4

Notes: (*) (1) Existing apartments; (2) New apartments; (3) Existing houses; (4) New houses.

These results highlight the fact that differences in the means of dwelling characteristics by quantile of transaction price are relevant. This asks for the application of quantile regression, as observed heterogeneity in dwelling attributes may give rise to parameter heterogeneity and different energy efficiency partial effects along the price distribution.

6.4. Estimation of conditional and unconditional regression coefficients

Table 6.3 provides a summary of energy efficiency partial effects for a group of selected price quantiles for existing apartments, new apartments, existing houses and new houses. For the sake of space, it was chosen to present the CQR (Koenker and Basset, 1978) and UQR (Firpo, 2007) only for the 0.1th, 0.5th and 0.9th price quantiles, a practice that is followed in similar studies (see, *inter alia*, Firpo et al., 2009; Peeters el at., 2017) [53]. This representation provides an overview of the entire range of quantile regression results and will be used throughout this chapter whenever similar comparisons are made. The OLS results are also provided in the table as a reference. The relative and percentage change effects of the impact of energy

[53] This is also a practice that is in accordance with the literature dealing with quantile regression analysis, which is often interested on the extremes of the distribution of interest (e.g., the most and the least well-off of the income distribution).

efficiency on transaction prices are estimated following Halvorsen and Palmquist's (1980) formula (see section 4.3.2).

Table 6.3. Comparing OLS, UQR and CQR partial effect estimates

	(*)	OLS	CQR			UQR		
			0.1th	0.5th	0.9th	0.1th	0.5th	0.9th
Energy efficiency	(1)	0.118**	0.138**	0.117**	0.105**	0.131**	0.105**	0.042**
		(.00192)	*(.0038)*	*(.00223)*	*(.00291)*	*(.00729)*	*(.00496)*	*(.00972)*
	(2)	0.123**	0.132**	0.118**	0.112**	0.163**	0.116**	0.094**
		(.00244)	*(.00442)*	*(.00261)*	*(.0042)*	*(.00913)*	*(.00694)*	*(.01922)*
	(3)	0.045**	0.021	0.049	0.077**	-0.182**	0.087**	0.095**
		(.00613)	*(.01142)*	*(.0061)*	*(.00993)*	*(.03063)*	*(.01894)*	*(.02287)*
	(4)	0.055**	0.063**	0.060**	0.068**	-0.182**	0.052*	0.034
		(.00624)	*(.00901)*	*(.00598)*	*(.00949)*	*(.04485)*	*(.01536)*	*(.02971)*

*Notes: (*) (1) Existing apartments; (2) New apartments; (3) Existing houses; (4) New houses; * p-value <0.05; ** p-value <0.0001. Standard errors given in parenthesis. Robust standard errors are provided for OLS results. Bootstrapped standard errors, based on 400 samples drawn with replacement, are provided for CQR results. The bootstrap was performed through the use of the sqreg command available in STATA. Due to the computational burden of estimating nonparametric propensity scores, it was not possible to derive bootstrapped UQR standards errors. Reported UQR standard errors are based on the variance estimator proposed by Firpo (2007).*

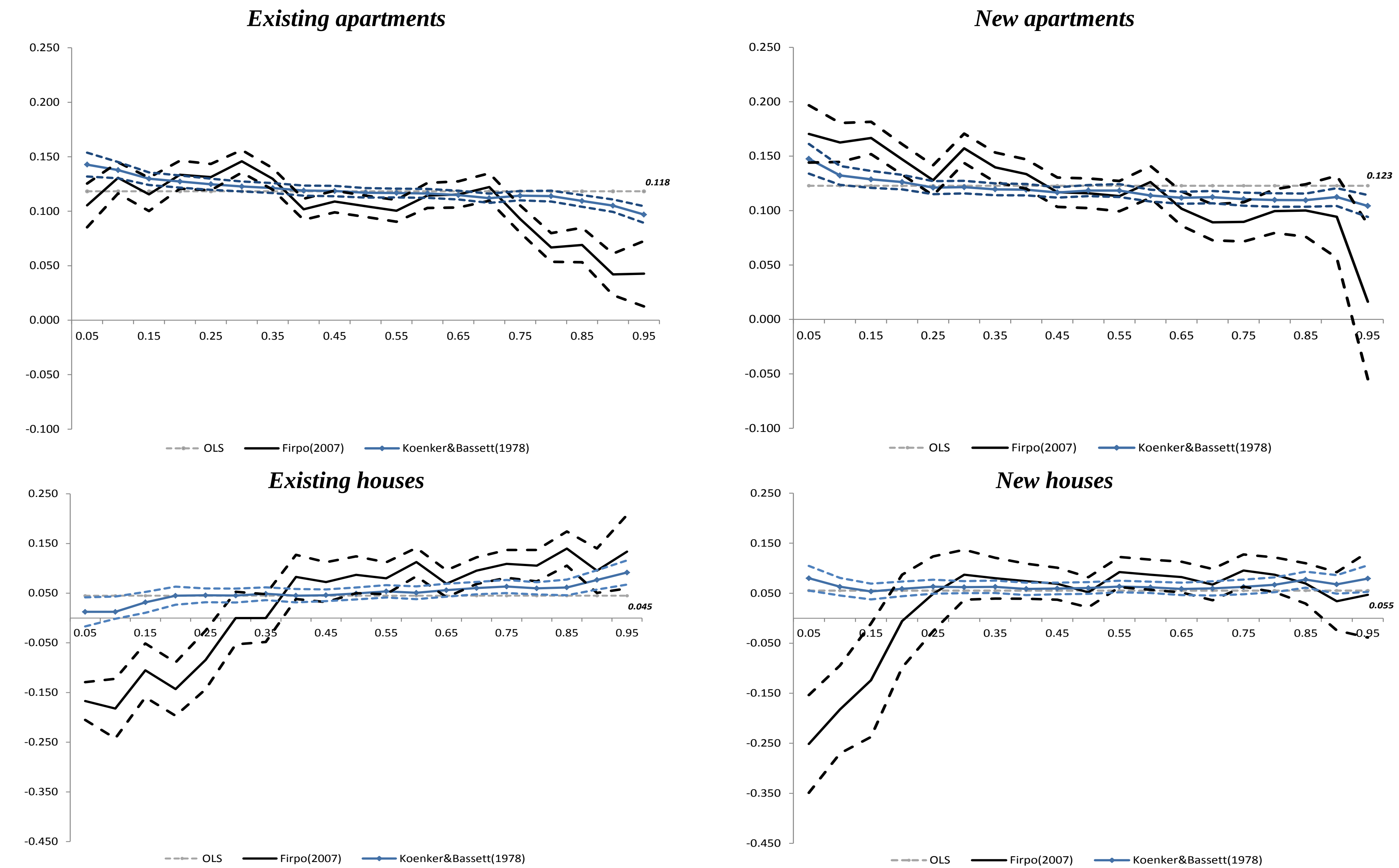
The table depicts differences between CQR and UQR results, not only in terms of magnitude of the partial effects (e.g., lower quantile for houses), but also in some cases in terms of signal (houses, at the lowest represented quantile). Quantile regression analysis confirms two key aspects of the nature of the relationship between energy efficiency and transaction prices, which were previously unveiled by OLS results. The first one is that energy efficiency is relevant as a price determining factor. As shown above, the variable measuring energy efficiency is statistically significant not only at the mean, but at the median, lower and higher quantiles of the conditional and unconditional price distribution. The second key aspect is the confirmation that apartments and houses do not reward energy efficiency in the same manner, with apartments receiving higher price premiums than houses. For instance, UQR results show that, while the price premium associated with a change from lower to higher energy efficiency at the median of the price distribution is 12.3 percent for new apartments, this percentage is only of 5.3 percent for new houses.

Additionally, there are important features in the relationship between residential property prices and energy efficiency that are disclosed by UQR analysis. The first is that energy efficiency partial effects appear to decrease monotonically across the quantiles for apartments. For instance, energy efficiency price premiums decrease from 14.0 percent at the 0.1th price quantile, to 11.1 and 4.3 percent at the 0.5th and 0.9th price quantiles of the unconditional price distribution of existing apartments, respectively (see Table 6.3). The second feature is that energy efficiency partial effects are not always positive across the distribution of transaction prices for houses. As Table 6.3 shows, statistically significant price discounts of 16.6 percent at the 0.1th price quantile are obtained for both new and existing houses. These are not captured by the least squares estimator, which seems to be essentially driven by the price premiums that exist at or above the median quartile. Interestingly, CQR coefficients are also positive at the 0.1th price quantile, a result that illustrates how dissimilar conditional and unconditional results can be. A more detailed view of UQR and CQR regression results is given in Figure 6.1, where energy efficiency quantile partial effects are depicted for 19 different quantiles (i.e., from the 0.05th to the 0.95th price quantiles). For ease of interpretation, OLS results (straight dashed lines) and the 95 percent confidence intervals of the UQR and CQR point estimates are also represented in the figure (by solid and dashed lines, respectively).

The first point that Figure 6.1 highlights is that CQR coefficients are less volatile than their UQR counterparts. This is an outcome that is also seen in studies that provide conditional and unconditional quantile partial effect estimates (see, *inter alia*, Peeters et al., 2017; Fournier and Koske, 2013). Moreover, and as it is possible to read from Firpo et al. (2009: 963), this can be seen as a standard result as it stems from the very nature of the two estimators. In fact, while CQR coefficients reflect a *within-group* impact from changing from less to more energy efficient standards - where *within-group* consists of the set of dwellings with the same values of all covariates (other than energy efficiency) used in quantile regression analysis -, UQR coefficients provide an overall impact change estimate, which reflects, in addition to *within-group* variation, *between-group* variation [54].

[54] In other words, while CQR partial effects are derived for a (theoretical) price distribution, where dwelling prices are conditional on the values of the set of the characteristics chosen for the specification of the hedonic price model, UQR results are built upon the heterogeneity associated with all the dwellings located at a given price quantile.

Figure 6.1: Conditional and unconditional energy efficiency partial effects



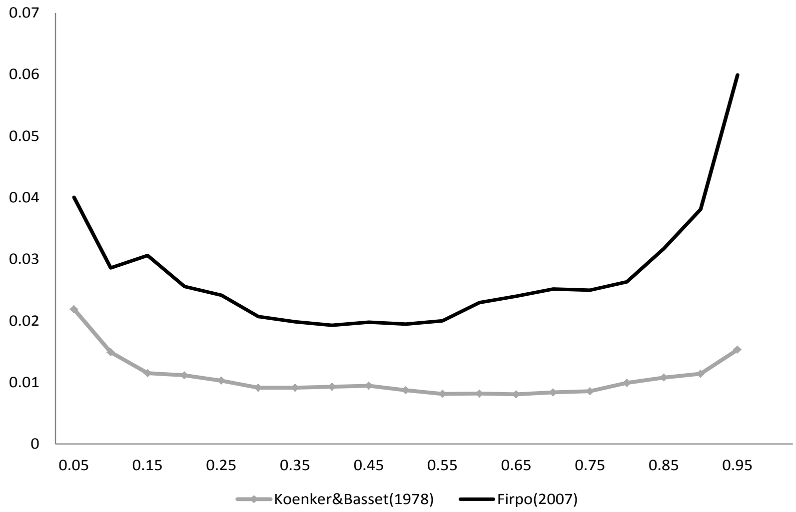
The second point worth noting from the figure above is the confirmation that energy efficiency partial effects for apartments tend to decrease from lower to higher quantiles of the conditional and unconditional price distributions. An explanation for the decreasing price premiums for apartments may be rooted in the idea that energy efficiency, which may be seen as a differentiating factor at lower quantiles of the sales distribution, is less valued by homebuyers interested in higher-priced apartments because good energy efficiency standards are more common in most expensive properties (see Table 6.1 for evidence on this). Houses, on the other hand, show a different behavior, which can be characterized by the existence of lower energy efficiency partial effects at the lower end of the price distribution. A possible explanation may be anchored in the idea that homebuyers may perceive high energy efficiency standards as additional future energy and maintenance costs, something that was already pointed out in Chapter 4. Following this reasoning, most expensive houses, which are likely to be purchased by high income homeowners, may reveal a lower price premium than least expensive homes because purchasers at the lower end of the distribution typically face more income constraints.

Lower price premiums may end up in price discounts, something that UQR results show at the lowest quantiles at the bottom panels of Figure 6.1. This leads to a third point that deserves to be underscored, which is the existence of price discounts associated with increases in energy efficiency standards. For new houses, for example, energy efficiency is clearly associated with price discounts below the 0.2th price quantile of the sales distribution. For existing houses, these persist at the 0.25th price quantile and energy efficiency coefficients are not statistically significant at the 0.3th and 0.35th price quantiles.

The amplitude of the confidence intervals for UQR point estimates, which are generally larger for the lower and upper quantiles of the sales distribution, is also a feature that stands out from Figure 6.1. For new dwellings, for instance, the confidence interval at the 0.95th quantile of the price distribution does not rule out the existence of a price discount associated with higher energy efficiency standards. Although less pronounced, this is also observed in CQR confidence intervals. The next figure illustrates the difference between the upper and lower bounds of CQR and UQR point estimate confidence intervals for existing apartments [55].

[55] Similar figures could be derived for the other market segments.

Figure 6.2: Upper and lower bounds confidence interval differences (existing apartments)



As it is evident from the picture, both CQR and UQR confidence interval differences follow an u-shaped curve, with the biggest differences corresponding to the upper and lower quantiles of the sales distribution. This characteristic reflects the much higher heterogeneity of less and most expensive transacted dwellings. This can be easily grasped if one looks at measures of variation of transacted prices. For existing apartments, the market segment represented in Figure 6.2, the group of the 25 percent of less expensive sales has a coefficient of variation of 20.7 percent. However, while this figure drops to 5.4 and 11.2 percent in the next two groups of transacted dwellings, the coefficient of variation for the top 25 percent of most expensive existing apartments jumps to 50.2 percent. Although it can be said that this data (dispersion) characteristic affects the width of confidence intervals (much more evident for UQR, which deals with the unconditional distribution of prices), it should not affect the quality of quantile regression point estimates.

Finally, it should be mentioned that Figure 6.1 reinforces the idea that, when apartments and houses are analyzed separately, there seems not to be an outstanding difference in the way new and existing dwelling markets value energy efficiency. In fact, the confidence intervals for new and existing dwellings overlap each other in many of the quantiles and provide, in this manner, clear evidence on the irrelevance of this variable for energy efficiency evaluation purposes. An example is given at the 0.75th price quantile, where energy efficiency entails a price premium for new apartments of 9.4 percent, a value that is close to its existing apartments' counterpart (9.7%) and that is in between its 95 percent confidence interval (8.4%, 11.1%).

As a conclusion, it can be said that that the UQR coefficients given in this section are important for energy efficiency policy purposes and for researchers in this area. They are relevant for policy purposes because low-priced houses are identified as the segment of the Portuguese residential property market that should be addressed with specific incentives aimed at overcoming energy efficiency price discounts. In addition, these results are also important for researchers because the idea that homebuyers of low-priced houses (or of other specific market segment) are only willing to tolerate extra energy efficiency gains when a price discount is provided, constitutes a plausible explanation as to the reason why some studies, which do not focus on the entire spectrum of the market, may present energy efficiency price discounts.

6.5. Coherence of results

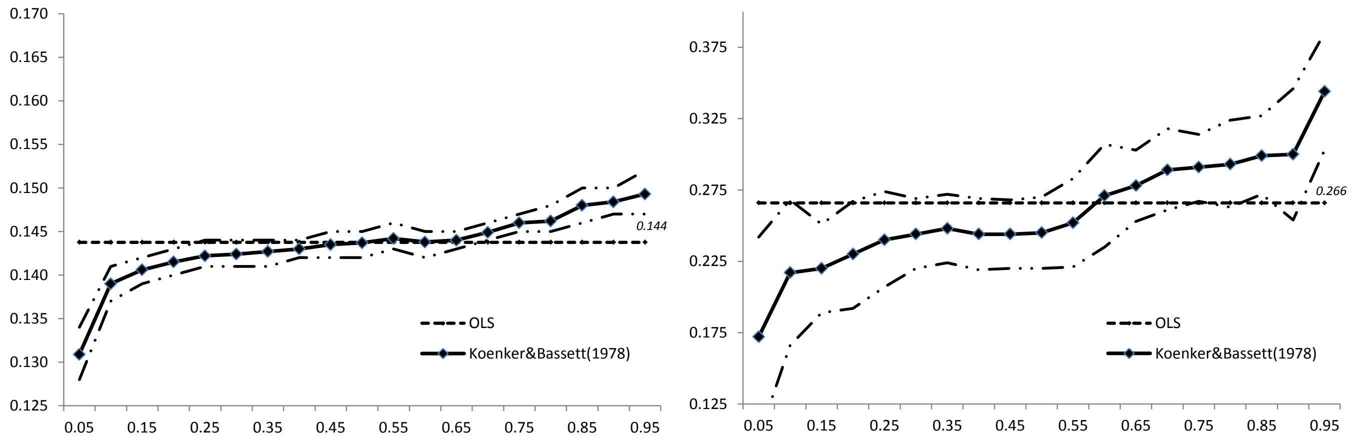
The soundness of the quantile regression results was assessed in two ways. The first one consisted in the comparison between the coefficient estimates of the variables that were common to this work and to the few studies employing conditional quantile regression analysis to the housing market. Although there are no papers applying UQR in this research context, it was possible to draw some tentative conclusions as to the reasonableness of the CQR valuation for some dwelling price determinants. In addition, the coherence of estimates was also investigated through a second exercise, which was based on Chapter's 5 cross-country qualitative assessment of the impact of energy efficiency on property prices. This last exercise involved the re-estimation of UQR coefficients using different energy efficiency measurement scales and dependent variables.

The few studies employing conditional quantile regression analysis to dwelling purchases suggest that there are substantial differences in the impact of attributes across the conditional distribution of house prices. In particular, size variables (lot size and square footage), were found in Zietz et al. (2008) to be more valued at the upper end of the sales price distribution. Area is also more valued at most expensive homes in Zhang and Yi (2017). A similar pattern for size was obtained in Mak et al. (2010). Moreover, age depreciation effects are reported by

Zietz et al. (2008) to be less pronounced for higher-priced properties. These authors acknowledge the existence of an effect in which age price discounts are smaller at the upper end than at the higher end of the sales distribution. For water front view, this study provides results that support the idea that this is a characteristic that is more valued for higher-priced properties. By the same token, Mak et al. (2010) find that obstructive views entail a higher discount at the upper end of the price distribution, thus providing evidence that homebuyers of higher-priced properties are more concerned with the visual prominence of the location than those purchasing cheaper properties.

Overall, these results are in the same line as those provided in this work. Figure 6.3, where CQR results per gross floor area (left panel) and scenic value of the location (right panel) are shown, illustrate this similarity. For the sake of completeness, CQR coefficient results (dot-dashed lines) are depicted together with 95 percent bootstrapped confidence intervals (based on 400 replications) and OLS coefficient results (dashed lines).

Figure 6.3: Selected OLS and CQR coefficients for existing apartments



The coherence of energy efficiency UQR partial effects to the use of different measurement scales was also tested. Table 6.4 provides OLS and UQR results for the 0.1th, 0.5th and 0.9th price quantiles for the energy efficient scales that were used in Hyland et al. (2013) and Fuerst et al. (2016). Although these studies apply the same dependent variable as the one that was used in this thesis, they are not directly comparable as price premiums or discounts are referenced to a different comparison base (i.e., use the D energy efficiency rate, instead of the

A or B EPC rates as reference). The results are only shown for existing apartments, which accounts for the largest group of transactions [56].

Table 6.4. Quantile estimates for different energy efficiency scales (existing apartments)

EPC rate	Hyland et al. (2013)				Fuerst et al. (2016)			
	OLS	0.1th	0.5th	0.9th	OLS	0.1th	0.5th	0.9th
A	0.211** *(00708)*	0.238** *(.03276)*	0.125 *(.1495)*	0.085 *(.08021)*	-	-	-	-
B	0.127** *(.00276)*	0.164** *(.00936)*	0.111** *(.01103)*	-0.104* *(.03484)*	-	-	-	-
C	0.015** *(.00249)*	0.072** *(.0074)*	0.027** *(.00574)*	-0.069** *(.01087)*	-	-	-	-
ABC	-	-	-	-	0.053** *(.00255)*	0.067** *(.00721)*	0.048** *(.00684)*	0.000 *(.01692)*
D	-	-	-	-	-	-	-	-
E	0.009* *(.00348)*	0.000 *(.01075)*	0.015 *(.0091)*	0.087* *(.01974)*	0.010* *(.00371)*	0.000 *(.01075)*	0.015 *(.00910)*	0.087** *(.01974)*
FG	-0.018 *(.00808)*	0.000 *(.04303)*	0.089* *(.0364)*	0.279** *(.06322)*	-0.019* *(.00952)*	0.000 *(.04303)*	0.089 *(.03640)*	0.279** *(.06322)*

*Notes: * p-value < 0.05; ** p-value <0.0001. Standard errors provided between brackets. Robust standard errors for OLS estimates. Reported UQR standard errors are based on the variance estimator proposed by Firpo (2007).*

On the whole, it can be said that Table 6.4 provides expected results across the quantiles. This is particularly evident for the top tier energy ratings where, for instance, A-rated properties yield higher price premiums at each quantile than those obtained for B-rated existing apartments. At the 0.1th quantile, while A-rated properties receive a price premium of 23.5 percent (coefficient of 0.238), the category immediately below obtains a 17.8 percent premium (coefficient of 0.164). In addition, the price premiums associated with A-rated properties at each quantile are higher than those that are obtained when A, B or C-rated properties are taken together (e.g., coefficients of 0.238 and of 0.067 at the 0.1th quantile).

Although the results for the rates that are common to the two rating systems (i.e., E and F or E EPC rates) are very similar for all the quantiles, the coherence of coefficient estimates for the less efficient dwellings is less evident than for most efficient properties. For the rates located below the comparison basis, unconditional quantile coefficient results generally show price premiums. However, the majority (7 out of 12) of these coefficients is not statistically significant, a result that is particularly evident for the 0.1th and 0.5th quantiles of the price

[56] Other market segments present similar results and their inclusion in this text would not add relevance to the analysis done on the basis of existing apartments data.

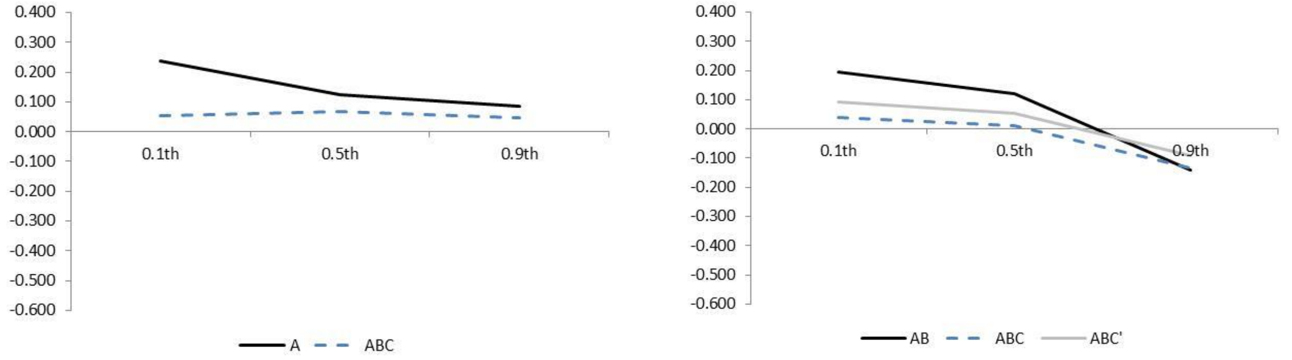
distribution. This suggests that energy efficiency is not taken into account by the market for the cheapest and less energy efficient residential properties. This finding is in line with Fuerst et al. (2016), who report the existence of a green signaling effect only in top-tier EPC rates as a consequence of the action of environmentally aware purchasers on the market. An explanation for the statistically significant price premiums for the simultaneously most expensive and less energy efficient properties may rest on the fact that these results are not derived on the basis of many observations. As shown in Chapter 3, only 1 percent of the existing apartments in the database bear an F or G EPC rate. These results stress the fact that, when more categories of energy efficiency levels are included in quantile regression models, the data requirements needed to produce reliable coefficient results increase a lot. This strengthens the decision taken at the outset of this thesis to concentrate on most energy efficient properties (i.e., A or B EPCs) and to use a single dummy to evaluate energy efficiency effects on dwelling transaction prices.

Figure 6.4 provides a comparison of coefficient estimates for all market segments focusing on the rates signaling most energy efficiency properties. In addition to the energy efficiency scales applied in Hyland et al. (2013) and Fuerst et al. (2016), which are signaled with the A and ABC labels on the left panels of Figure 6.4, those that were applied in Brounen and Kok (2011), Fuerst et al. (2015) and Ramos et al. (2015b) are also provided in the figure. The last three were grouped separately from the first two, since these authors use the logarithm of price per square meter in their models [57]. The results based on the energy efficiency scale used in Fuerst et al. (2015) are identified in the figure as AB and the ones referring to Brounen and Kok's (2011) and Ramos et al. (2015b) as ABC and ABC', respectively (all depicted on the right panels of the figure).

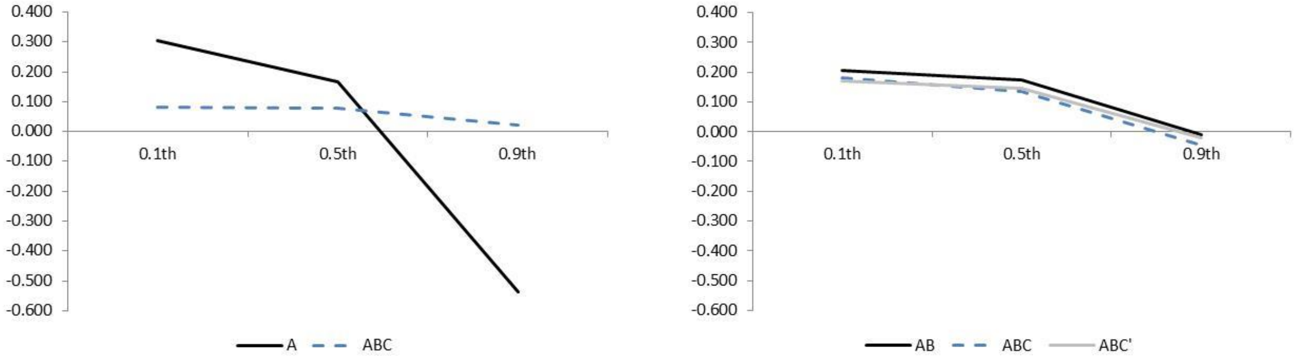
[57] At this point, it should be noted that the model that uses the logarithm of price per square meter as a dependent variable and the model that has the logarithm of price as a dependent variable and a transformation of area as an independent variable are not the same. This has to do with the type of transformation applied to area (its square root, not its logarithm), and the fact that area variables are always kept in the models as covariates, even when the model's dependent variable is the logarithm of the price per square meter.

Figure 6.4: Energy efficiency partial effects for different energy efficiency scales

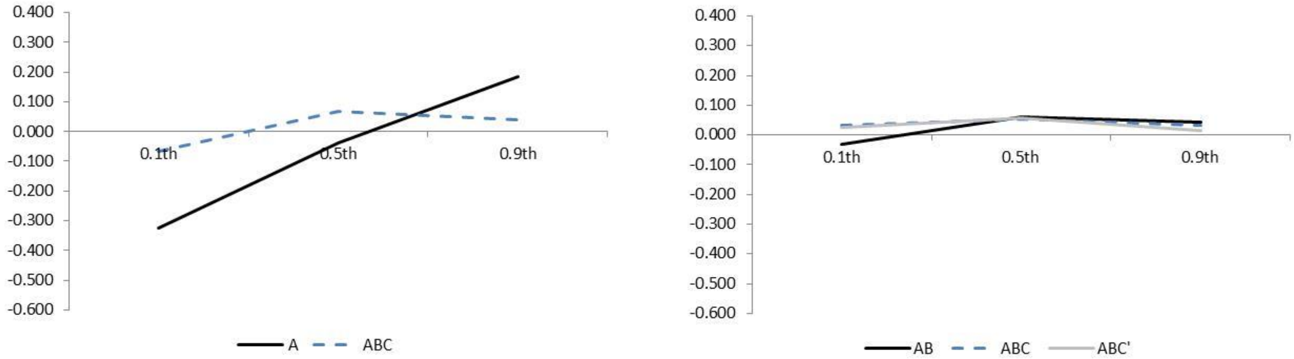
Existing Apartments



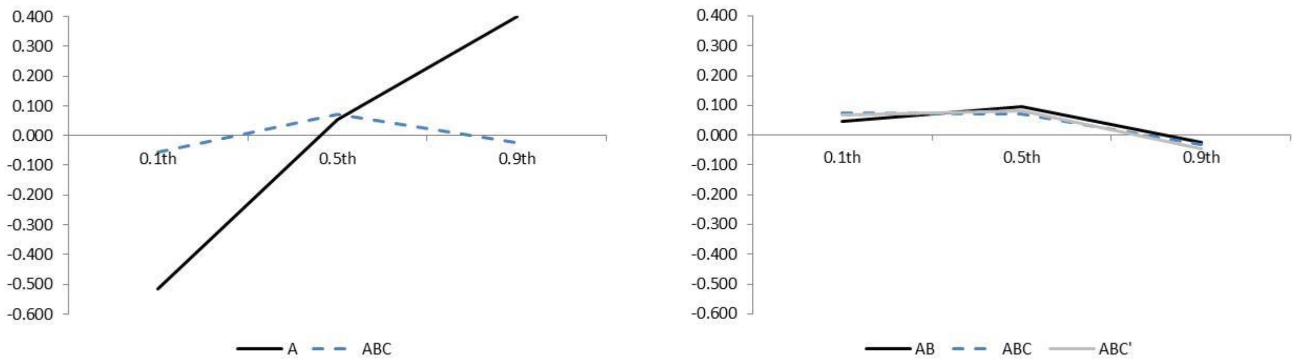
New Apartments



Existing Houses



New Houses



The results depicted by Figure 6.4 are in line with the idea, already pointed out in Section 6.4, that energy efficiency partial effects for apartments decrease from lower to higher quantiles of the unconditional price distribution (for the 0.9th quantile, there is evidence of price discounts). This situation is independent from the choice of the energy efficiency scale and of the model's dependent variable, and reinforces the idea that there may be some tendency regarding the reduction of price premiums as one moves from the lower to the higher spectrum of the price distribution. For houses, the evidence provided is mixed. However, most of the situations entailing a price discount are located at the beginning of the price distribution (0.1th quantile), something that was already portrayed in Figure 6.1.

By and large, it can be said that the two comparison exercises that were carried out to check the coherence of regression outputs confirm the quality of quantile regression coefficient estimates and of the conclusions drawn from them.

6.6. Summary

This is the first study to provide evidence on the response of residential transaction prices to changes in energy efficiency standards using an unconditional quantile framework. The findings support the idea that the impacts of greater energy efficiency are not uniformly positive across the unconditional quantile regressions for all dwelling types. In particular, for houses located at or below the 0.2th price quantile, there is clear evidence for the existence of statistically significant price discounts associated with greater energy efficiency standards. At the 0.1th price quantile, both new and existing houses show a statistically significant price premium of -16.6 percent. Moreover, although energy efficiency is positively rewarded across the entire price distribution for apartments, it is possible to observe a reduction in the magnitude of the price premiums at the highest quantiles of the distribution. Quality checks using different measurement scales and dependent variables support these main findings.

These results provide a more complete picture of the impact of energy efficiency on residential property prices than the one provided by traditional OLS analysis, whose heterogeneity of responses are masked by the results obtained at the mean of the price distribution. In particular, the application of the UQR estimator on transactions data provided important insights not only to policy makers but also to those interested in unveiling the relationship between energy efficiency and transaction prices in the residential market. The use of UQR has identified low-priced houses as the market segment that should be addressed by specific policies aimed at increasing energy efficiency. The lower valuation observed for most expensive apartments may stem from the fact that these properties already have, in most cases, very high energy efficiency standards and are seen as a common feature rather than a price differentiation factor. Finally, provided results give a plausible explanation as to the reason why some studies, which do not focus on the entire spectrum of the market, present energy efficiency price discounts.

Chapter 7 Conclusions

This thesis investigates the degree to which energy efficiency is reflected in residential property prices in Portugal. Its originality stands on two key features. The first one is that it constitutes the first large-scale study on the impact of energy efficiency on residential property transaction prices for a European southern country. While only a few recent studies cover meridional markets (Ramos et al., 2015b; Ayala et al., 2016), they are all based on small samples and in poorer datasets, which do not have transaction prices and other important dwelling characteristics. The second original key feature is that it provides results of the application of conditional (Koenker and Basset, 1978) and unconditional (Firpo, 2007) quantile regression analysis, which have never been used in the estimation of energy efficiency partial effects.

The results stemming from the application of hedonic price models to transaction prices, dwelling characteristics and EPC information identified a 13 percent energy efficiency price premium for apartments and a 5 to 6 percent energy efficiency price premium for houses in the Portuguese residential property market. The euro value attached to these premiums is sizeable at the point of means. For instance, considering that the average transaction price of an existing apartment is 97,695 €, it corresponds to 12,212 €. However, while these results are valid for the mean of the distribution of transaction values, the market response to energy efficiency is far from being always positive across the distribution of dwelling transaction prices. In particular, the magnitude of price premiums for apartments is reduced at the highest quantiles, and the houses located at or below the 0.2th price quantile display statistically significant price discounts (e.g., -16.6 percent at the 0.1th quantile of the price distribution).

In addition to the finding that energy efficiency is reflected in residential property prices in Portugal, this research adds other important contributions to the literature in this research context. First, the empirical work carried out in Chapter 4 clearly suggests that hedonic price modelling should be done separately for each market context. Rather than trying to find a model for all data, researchers should develop model specification strategies addressing the specificities of residential property sub-markets. Second, Chapter's 5 results clearly suggest that the use of list prices in hedonic regression models may lead an overestimation of the price premium associated with energy efficiency. This is of paramount importance as researchers often do not have access to transaction prices and have to use proxy prices, which typically exhibit some sort of measurement error. In addition, the experiments using different omitted variable scenarios underline the idea that, while the omission of a single variable may not be too problematic, the joint omission of key quality attributes, such as those reflecting the quality of location, may seriously bias the estimate of the impact of energy efficiency on property prices. Third, by disclosing price discounts for the less expensive houses, the UQR results given in Chapter 6 offer a plausible explanation as to why some studies, which do not cover the entire spectrum of the market, may not exhibit energy efficiency price premiums. Moreover, in applying quantile regression analysis, its usefulness in delivering a more complete picture of the impact of energy efficiency than the one provided by OLS regression analysis was also emphasized.

With the exception of the finding of the existence of a higher valuation of energy efficiency for new dwellings, the quality of OLS hedonic regression outputs was confirmed through coherence tests, which included the comparison of results across different studies and energy efficiency measurement scales. In particular, cross-country comparisons supported the view that the Portuguese residential market displays higher price premiums than northern European markets. The use of different measurement scales was also applied in Chapter 6 within the UQR framework. This allowed not only to confirm the conclusions drawn from earlier quantile regression analysis, but also to illustrate the demanding data requirements associated with the estimation of reliable energy efficient quantile coefficients.

The results also provide extremely valuable contributions for policy makers in this area. More concretely, they show that, although both sub-markets are associated with overall mean price premiums, apartments value energy efficiency more than houses. In addition, the use of UQR demonstrates that energy efficiency partial effects are not uniformly positive across the price distribution, with less expensive houses showing clear price discounts. As mentioned above, this is particularly evident for the existing and new houses located at or below the 0.2th price quantile, which are those that are transacted at or below the 60,000 € and 90,000 €, respectively. As houses represent the majority of the housing stock in Portugal, these findings provide useful information to policy makers interested in tailoring policy measures and to use them in a cost-effective way.

One shortcoming of the present research is associated with the fact that, although the derived dataset is one of the largest ever used to study the relationship between energy efficiency and transaction prices, it was only possible to match 60 percent of all individual transactions with EPC data. As such, the answers provided in this thesis would naturally benefit from a higher matching rate and the availability of more data. This limitation is more evident for thinner market segments such as new houses, where the soundness of the results could benefit from data on more transactions. As the information that was used refers to an early stage of EPC implementation in Portugal, it is possible that some of the inaccuracies found in the EPC data are simply a symptom of the novelty of the label. Following this reasoning, it is reasonable to assume that its underlying quality will improve as the EPC data generating system develops into more mature stages. Another limitation of the present study has to do with the fact that, although extremely rich in terms of variables, the dataset did not have real list prices to empirically assess the impact of the use of a proxy of transaction prices in hedonic regression models. As described in Chapter 5, this limitation was overcome by the use of a Monte Carlo exercise in which list prices were simulated. Although providing valuable results, it would be interesting to see Chapter's 5 findings reassessed in a study in which transaction and list prices are available from the same buying-selling processes.

This work opens up new research directions. The first one has to do with the valuation of the EPC label from 2014 onwards, a period of time in which the residential property market has recovered from its recession years. In fact, the time period covered broadly overlaps the years in which the Portuguese residential property market was depressed. For this period, it was found that price premiums increased over the years, a situation that is in line with the literature supporting the idea that the value of certification is high when market conditions are worse (e.g., Hyland et al., 2013). As pointed out by Lourenço and Rodrigues (2017), residential property prices fell by an average of 4 percent per year from 2007 to 2013. However, from this year on, the market recovered, with prices rising on average 5 percent a year. In this context, it would be important for a sounder evaluation of the implementation of EPC label in Portugal to have an empirical assessment of the impact of energy efficiency in residential transaction prices from 2014 to more recent years.

The second research direction is the development of cross-country comparisons, which were explored in Chapters 5 and 6 of this thesis. In particular, the higher price premiums for Portugal should be further investigated. As pointed out earlier in this thesis, this outcome could stem from a greater awareness of the EPC label or from the existence of higher energy costs in this country. However, another line of investigation would be the analysis and comparison of results in similar real estate markets. The findings of Ayala et al. (2016), which estimate relatively high energy efficiency price premium for the Spanish market, shed some light on where to look for similarities. Thus, a possible reason for the existence of higher price premiums in the Iberian Peninsula could be based on supply side factors, such as overall building technology and average quality of construction materials, which are probably worse than in northern European countries. Chapter 5 already provides some tentative conclusions on the importance of supply side factors to explain the price premium associated with the implementation of greater energy efficiency standards. Finally, the application of conditional and unconditional regression techniques would definitely deserve further research in the near future. In particular, it would be very interesting to see if the impact of energy efficiency improvements is uniformly positive across the distribution of residential (and commercial) property prices in other countries. Following this line of though, it would be important to test whether other markets exhibit energy efficiency price discounts for the least expensive houses of the market, a situation that was unveiled by the application of the UQR method on Portuguese residential property data.

Bibliography

Agência para a Energia (ADENE) (2015). *Energy efficiency trends and policies in Portugal*. 2015. Lisboa: ADENE.

Amel, D. F., and Bernt, E. R. (1986). *Depreciation in the Swedish automobile market: An integration of hedonic and latent approaches*. Cambridge: Massachusetts Institute of Technology, Center for Energy Policy Research.

Anglin, P. M., and Gençay, R. (1996). Semiparametric estimation of a hedonic price function. *Journal of Applied Econometrics*, 11, 633-48.

Ankamah-Yeboah, I., and Rehdanz, K. (2014). Explaining the variation in the value of building energy efficiency certificates: A quantitative meta-analysis. *Working paper n.º 1949*. Kiel Institute.

Arguea, N., and Hsiao, C. (1993). Econometric issues of estimating hedonic price functions: With an application to the U.S. market for automobiles. *Journal of Econometrics*, 56(1-2), 243-67.

Atanasiu, B., and Constantinescu, T. (2011). A comparative analysis of the energy performance certificates schemes within the European Union: Implementing options and policy recommendations. *ECEEE Summer Study 2013 Proceedings*. European Council for an Energy-Efficient Economy (ECEEE).

Ayala, A., Galarraga, I. and Spadaro, J. V. (2016). The price of energy efficiency in the Spanish housing market. *Energy Policy*, 94, 16-24.

Banco de Portugal (n.d.). *Novas operações: empréstimos a particulares da UM, habitação*. Data available from https://www.bportugal.pt/Mobile/BPStat/DominiosEstatisticos.aspx.

Bajari, P., and Benkard, L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6), 1239-76.

Bassett, G., and Chen, H.-L. (2001). Quantile style: Return-based attribution using Regression quantiles. *Empirical Economics*, 26(1), 293-305.

Bernt, E. (1991). *The practice of econometrics: Classic and contemporary*. Reading, MA: Addison-Wesley Publishing Company.

Berry, W. D. (1993). *Understanding regression results*. Sage university papers series. Quantitative applications in the social sciences, no. 07-092.

Borah, B. J., and Basu, A. (2013). Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. *Health Economics,* 22, 1052-70.

Bover, O., and Izquierdo, M. (2003). Quality-adjusted prices: Hedonic methods and implications for national accounts. *Investigaciones Económicas,* 27(2), 199-238.

Box, G. E., and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26(2), 211-52.

Breusch, T. S., and Pagan, A. R. (1979). A Simple Test for Heteroskedasticity and Random Coefficient Variation. *Econometrica,* 47 (5), 1287-94.

Brounen, D., and Kok, N. (2011). On the economics of energy labels in the housing market. *Journal of Environmental Economics and Management,* 62(2), 166-79.

Brounen D., Kok, N., and Quigley, J.M. (2013). Energy literacy, awareness, and conservation behavior of residential households. *Energy Economics*, 38, 42-50.

Brown, M. A. (2001). Market failures and barriers as a basis for clean energy policies. *Energy Policy*, 29(14), 1197-207.

Bruegge, C., Carrión-Flores, C. and Pope, J.C (2016). Does the housing market value energy efficient homes? Evidence from the energy star program. *Regional Science and Urban Economics,* 57, 63-76.

Buildings Performance Institute Europe (BPIE) (2010). *Energy Performance Certificates across Europe: From design to implementation*. Available at http://bpie.eu/publication/energy-performance-certificates-from-design-to-implementation/.

Cajias, M., and Piazolo, D. (2013). Green performs better: Energy efficiency and financial return on buildings. *Journal of Corporate Real Estate,* 15(1), 53-72.

Cassel, E., and Mendelsohn, R. (1985). The choice of functional forms for hedonic price equations: Comment. *Journal of Urban Economics,* 18, 135-42.

Carrillo, P. E. (2010). *To sell or not to sell: List price, transaction price and marketing time in the housing market*. Available at https://ssrn.com/abstract=1762757.

Cropper, M. L., Deck, L. B., and McConnell, K. E. (1988). On the choice of functional form for hedonic price functions. *Review of Economics and Statistics,* 70(4), 668-75.

Cerin, P., Hassel, L. G., and Semenova, N. (2014). Energy performance and housing prices. *Sustainable Development,* 22(6), 404–19.

Collins, A., Scorcu, A., and Zanola, R, (2009). Reconsidering hedonic art price indexes. *Economic Letters,* 104(2), 57-60.

Combris P., Lecocq, S., and Visser, M (1997). Estimation of a hedonic price equation for Bordeaux wine: Does quality matter?. *Economic Journal*, 107(441), 390-402.

Colwell, P. F. (1993). Comment: Semiparametric estimates of the marginal price of floorspace. *Journal of Real Estate Finance and Economics*, 7, 73-5.

Colwell, P. F., and Dilmore, G. (1999). Who was first? An examination of an early hedonic study. *Land Economics*, 74(4), 620-6.

Court, A. T. (1939). Hedonic price indexes: With automotive examples. In *The Dynamics of automobile demand* (pp. 99-119). New York: General Motors Corporation.

Chin, T., and Chau, K. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing and its Applications*. 27(2), 145-65.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591-603.

Cropper, M. L., Deck, L. B., and McConnell, K. E. (1988). On the choice of functional form for hedonic price functions. *The Review of Economics and Statistics*, 70(4), 668-75.

Davis, T.P., McCord, J., McCord, M.J., and Haran, M. (2015). Modelling the effect of energy performance certificate rating on property value in the Belfast housing market. *International Journal of Housing Markets and Analysis*, 8(3), 292-317.

Deaton, A. (1997). *The analysis of household surveys*. Baltimore: Johns Hopkins.

Decreto-Lei n.º 38382/1951 de 7 de agosto. *Diário do Governo 1º Suplemento, Série I*. Ministério das Obras Públicas - Gabinete do Ministro. Lisboa.

Decreto-Lei n.º 40/1990 de 6 de fevereiro. *Diário da República n.º 67, 1 Série-A*. Ministério das Obras Públicas Transportes e Comunicações. Lisboa.

Decreto-Lei n.º 287/2003 de 12 de novembro. *Diário da República n.º 262/2003*, Série I-A. Ministério das Finanças. Lisboa.

Decreto-Lei n.º 78/2006 de 4 de abril. *Diário da República n.º 67/2006, Série I-A*. Ministério da Economia e da Inovação. Lisboa.

Decreto-Lei n.º 80/2006 de 4 de abril. *Diário da República n.º 67/2006, Série I-A*. Ministério das Obras Públicas, Transportes e Comunicações. Lisboa.

Decreto-Lei n.º 50/2008 de 19 de março. *Diário da República n.º 56/2008, Série I*. Ministério das Obras Públicas, Transportes e Comunicações. Lisboa.

Decreto-Lei n.º 118/2013 de 20 de agosto. *Diário da República n.º 159/2013*, Série I. Ministério da Economia e do Emprego. Lisboa.

Deng, Y., Li, Z., and Quigley, J. M. (2012). Economic returns to energy-efficient investments in the housing market: Evidence from Singapore. *Regional Science and Urban Economics*, 42(3), 506-15.

Deng, Y., and Wu, J. (2014). Economic returns to residential green building investment: The developers' perspective. *Regional Science and urban Economics,* 47, 35-44.

Dinan, T. M., and Mirarowski, J. A. (1989). Estimating the implicit price of energy efficiency improvements in the residential housing market: A hedonic approach. *Journal of Urban Economics,* 25, 52-67.

Direção-Geral dos Impostos (DGI) (2011). *Manual de Avaliação de Prédios Urbanos (Versão 5.0)*. Lisboa: DGI, Divisão de Serviços de Avaliações.

Directive 1992/75/EEC. *Official Journal* of the *European Union no L 297/16.* Brussels: Council of the European Union.

Directive 2010/30/EU. *Official Journal of the European Union no L 153/1.* Strasbourg: European Parliament and of the Council.

Directive 2010/31/EU. *Official Journal of the European Union no L 153/13.* Strasbourg: European Parliament and of the Council.

Eichholtz, P., Kok, N., and Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *American Economic Review,* 100(5), 2492-509.

Eichholtz, P., Kok, N., and Quigley, J. M. (2013). The economics of green building. *Review of Economics and Statistics,* 95(1), 50-63.

Ekeland, I., Heckman, J. J., and Nesheim, L. (2002). Identifying hedonic models. *American Economic Review*, 92(2), 304-9.

Epple, D. (1987). Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products, *Journal of Political Economy*. 95(1), 59-80.

European Commission (2011). *The economic adjustment programme for Portugal.* Occasional papers no. 79. Brussels: Directorate-General for Economic and Financial Affairs.

Eurostat (2013). *Handbook on residential property prices indices*. Luxembourg: Eurostat.

Fesselmeyer, E. (2018). The value of green certification in the Singapore housing market. *Economic Letters,* 163, 36-9.

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, 75(1), 259-76.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica,* 77, 953–73.

Fournier, J.-M., and Koske, I. (2013). Public employment and earnings inequality: An analysis based on conditional and unconditional quantile regressions. *Economics Letters*, 121(2), 263-6.

Frölich, M., and Melly, B. (2010). Estimation of quantile treatment effects with Stata. *Stata Journal*, 10(3), 423-57.

Frölich, M., and Melly, B. (2013). Unconditional Quantile Treatment Effects Under Endogeneity. *Journal of Business and Economic Statistics*, 31(3), 346-57.

Fuerst, F., and McAllister, P. (2011a). Green noise or green value? Measuring the effects of environmental certification on office values. *Real estate Economics*, 39(1), 45-69.

Fuerst, F., and McAllister, P. (2011b). The impact of energy performance certificates on the rental and capital values of commercial property assets. *Energy Policy*, 39(10), 6608-14.

Fuerst, F., McAllister, P., Nanda, A., and Wyatt, P. (2015). Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. *Energy Economics*, 48, 145-56.

Fuerst, F., Oikarinen, E., and Harjunen, O. (2016). Green signalling effects in the market for energy-efficient residential buildings. *Applied Energy*, 180, 560-71.

Fuerst, F., and Shimizu, C. (2016). Green Luxury Goods? The Economics of Eco-Labels in the Japanese Housing Market. *Journal of Japanese and International Economy*, 39, 108-22.

Galego, A., and Pereira, J. (2014). Inter-regional wage differentials in Portugal: An analysis across the wage distribution. *Regional Studies*, 48(9), 1529-46.

Giles, D.E.A. (1982). The interpretation of dummy variables in semilogarithmic equations: Unbiased estimation. *Economic Letters*, 10(1-2), 77-9.

Goodman, A. C. (1998). Andrew Court and the invention of hedonic price analysis. *Journal of Urban Economics*, 44(2), 291-8.

Goodman, A. C., and Thibodeau, T. G. (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, 6(1), 25-42.

Greening, L. A., Sanstad, A. H., and McMahon, J. E. (1997). Effects of appliance standards on product price and attributes: An hedonic pricing model. *Journal of Regulatory Economics*, 11(2), 181-94.

Griffits, W. E., Hill, R. C., and Judge, G. G. (1993). *Learning and practicing econometrics*. New York: John Willey & Sons.

Griliches, Z. (1961). Hedonic prices for automobiles: An econometric analysis of quality change. In: *The price statistics of the Federal Government* (pp. 137-196). New York: Columbia University Press for the National Bureau of Economic Research.

Halvorsen, R., and Palmquist, R. (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review, 70*(3), 474-5.

Halvorsen, R., and Pollakowski, H. O. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, 10(1), 37-49.

Hayunga, D. K., and Pace, R. K. (2017). List prices in the US housing market. *Journal of Real Estate Finance and Economics*, 55(2), 155-84.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1). 153-61.

Hill, R. J. (2013). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, 27(5), 879-914.

Högberg, L. (2013). The impact of energy performance on single-family home selling prices in Sweden. *Journal of European Real Estate Research*, 6(3), 242-61.

Horowitz, J. L. (1992). The role of the list price in housing markets: Theory and an econometric model. *Journal of Applied Econometrics*, 7, 115-29.

Howarth, R. B., Haddad, B. M., and Paton, B. (2000). The economics of energy efficiency: Insights from voluntary participation programs. *Energy Policy*, 28(6-7), 477-86.

Hyland, M., Lyons, R. C., and Seán, L. (2013). The value of domestic building energy efficiency: Evidence from Ireland. *Energy Economics*, 40, 943-52.

Instituto Nacional de Estatística (INE) (2011). *Composição das cidades estatísticas por subseções da BGRI 2011*. Lisboa: Instituto Nacional de Estatística.

Instituto Nacional de Estatística (INE) (2012*). Censos 2011, XV recenseamento geral da população: V recenseamento geral da habitação, Resultados definitivos*. Lisboa: Instituto Nacional de Estatística.

Instituto Nacional de Estatística (INE) (2015). House Price Index, 1st Quarter 2015. Press Release. Lisboa: Instituto Nacional de Estatística.

Instituto Nacional de Estatística (INE) (2017a). *Documento metodológico, Índice de Preços da Habitação*. Versão 2.0. Lisboa: Instituto Nacional de Estatística.

Instituto Nacional de Estatística (INE) (2017b). *Índice de Preços das Propriedades Comerciais. Documento Metodológico*. Versão 1.0. Lisboa: Instituto Nacional de Estatística.

International Energy Agency (IEA) (2017). *CO_2 Emissions from fuel Combustion: Highlights*. 2017. Paris: International Energy Agency.

Jensen, M.O., Hansen, A.R., and Kragh, J. (2016). Market response to the public display of energy performance rating at property sales. *Energy Policy*, 93, 229-35.

Jolliffe, D. (2011). Overweight and poor? On the relationship between income and the body mass index. *Economics and human biology*, 9, 342-55.

Kahn, M. E., and Kok, N. (2014). The capitalization of green labels in the California housing market. *Regional Science and Urban Economics,* 47, 25-34.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogaritmic equations. *American Economic Review*, 71(4), 801.

Kiel, K. A., and Zabel, J. E. (2008). Location, location, location: The 3L Approach to house price determination. *Journal of Housing Economics*, 17(2), 175-90.

Koenker, R. (2017). Quantile Regression: 40 Years On. *Annual Review of Economics,* 9(1), 155-76.

Koenker, R., and Basset, G. (1978). Regression quantiles. *Econometrica*, 46(1), 33-50.

Kuminoff, N. V., Parmeter, C. F., and Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?. *Journal of Environmental Economics and Management*, 60(3), 145-60.

Kuminoff, N. V., and Pope, J. C. (2012). A novel approach to identifying hedonic demand parameters. *Economic Letters*, 116(3), 374-6.

Laquatra, J. (1986). Housing market capitalization of thermal integrity. *Energy Economics*, 8(3), 134-8.

Laquatra, J., Dacquisto, D. J., Emrath, P., and Laitner, J. A. (2002). Housing market capitalization of energy efficiency revisited. *Summer Study on Energy Efficiency in Buildings*, Washington, D.C.: American Council for an Energy-Efficient Economy.

Lin M., Lucas, H.C., and Shmueli, G. (2013). Too big to fail: Large samples and the p-value problem. *Information Systems Research*, 24(4), 906-17.

Lourenço, R. F., and Rodrigues, P. M. M. (2017). Preços da habitação em Portugal - uma análise pós-crise. *Revista de Estudos Económicos,* 3(4), Banco de Portugal, 45-62.

Maclean, J. C., Webber, D. A., and Marti, J. (2014). An application of unconditional quantile regression to cigarette taxes. *Journal of Policy Analysis and Management*, 33, 188–210.

Mak, S., Choy, L., and Ho, W. (2010). Quantile regression estimates of Hong Kong real estate prices. *Urban Studies*, 47, 2461-72.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O. Gibbs (Ed.), *Housing economics and public policy: Essays in honour of Duncan Maclenn* (pp. 67-90). Blackwell.

Megerdichian, A. (2018). Further results on interpreting coefficients in regressions with a logarithmic dependent variable. Journal of Econometric Methods, 7(1), 1-5.

Mishra, A. K., Mottaleb, K. A., and Mohanty, S. (2015), Impact of off-farm income on food expenditures in rural Bangladesh: an unconditional quantile regression approach. *Agricultural Economics*, 46, 139-48.

Montgomery E., Shaw, K., and Benedict, M. E. (1992). Pension and wages: An hedonic price theory approach. *International Economic Review*, 33(1), 111-28.

Nelson, J. P. (1982). Highway noise and property values: A survey of recent evidence. *Journal of Transport Economics and Policy*, 16(2), 117-38.

Novan, K., Smith, A., and Zhou, T. (2017). *Residential building codes do save energy: Evidence from hourly smart-meter data*. University of California, Davis Working Paper. Available at https://e2e.haas.berkeley.edu/pdf/workingpapers/WP031.pdf.

Oaxaca, R. (1973). Male-female wage differentials in urban labour markets. *International Economics Review*, 14(3), 693-709.

Oliveira, R., Lopes, J., and Abreu, I. (2012). *O impacto de soluções técnicas sustentáveis no imposto municipal sobre imóveis*. In Universidade de Coimbra (Ed.). (pp. 210-222). Coimbra: Universidade de Coimbra.

Parmeter, C. F., Henderson, D. J., and Kumbhakar, S. C. (2007). Nonparametric estimation of a hedonic price function. *Journal of Applied Econometrics*, 22, 695-9.

Peeters, L., Schreurs, E., and Van Passel, S. (2017). Heterogeneous Impact of Soil Contamination on Farmland Prices in the Belgian Campine Region: Evidence from Unconditional Quantile Regressions. *Environmental and Resource Economics*, 66(1),135–68.

Portaria n.º 982/2004 de 4 de agosto. *Diário da República n.º 182, Série I-B*. Ministério das Finanças Lisboa.

Portaria n.º 1426/2004 de 25 de novembro. *Diário da República n.º 277, Série I-B*. Ministério das Finanças e da Administração Pública. Lisboa.

Portaria n.º 1022/2006 de 20 de setembro. *Diário da República n.º 190, Série I*. Ministério das Finanças e da Administração Pública. Lisboa.

Portaria n.º 1305/2007 de 4 de outubro. *Diário da República n.º192, Série I*. Ministério das Finanças e da Administração Pública Lisboa.

Portaria n.º 461/2007 de 5 de junho. *Diário da República n.º 108, Série II*. Ministérios da Administração Interna, do Ambiente, do Ordenamento do Território e doo Desenvolvimento Regional, da Economia e da Inovação e das Obras Públicas, Transportes e Comunicações . Lisboa.

Portaria n.º 1119/2009 de 30 de setembro, *Diário da República n.º 190, Série I*. Ministério das Finanças e da Administração Pública. Lisboa.

Portaria n.º 420-A/2015 de 31 de dezembro. *Diário da República n.º 255, 2º Suplemento, Série I*. Ministério das Finanças. Lisboa.

Ramalho, E.A., and Ramalho, J.J.S. (2010). *On the econometrics of hedonic housing price: A survey*. Mimeo.

Ramalho, E.A., Ramalho, J.J.S., and Evangelista, R. (2017). Combining micro and macro data in hedonic price indexes. *Statistical Methods and Applications*, 26(2), 317-32.

Ramos, A., Gago, A., Labandeira, X., and Linares, P, (2015a). The role of information for energy efficiency in the residential sector. *Energy Economics*, 52(S1), S17–S29.

Ramos, A., Pérez-Alonso, A., and Silva, S. (2015b). *Valuing energy performance certificates in the Portuguese residential sector*. Available at http://eforenergy.org/docpublicaciones/documentos-de-trabajo/WP02-2015.pdf.

Ramsey, J.B. (1969). Tests for specification errors in classical linear least-squares regression. *Journal of the Royal Statistical Society*, 31(2), 350-71.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.

Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

Rothe, C. (2012). Partial distributional policy effects. *Econometrica*, 80(5), 2269 – 301.

Santos, E., and Coimbra, C. (1995) Preços de automóveis e alterações de qualidade, Banco de Portugal. *Boletim Económico*, Setembro, 57-64.

SAS Institute INC (2015). *SAS Enterprise Guide 7.11*, Cary, NC.

Soriano, F. (2008). Full statistical report. In: *Energy efficiency rating and house price in the Act* (pp. 9-49). Camberra: Department of the Environment, Water, Heritage and the Arts.

Stanley, S., Lyons, R., and Lyons, S. (2016). The price effect of building energy ratings in the Dublin residential market, *Energy Efficiency*, 9(4), 875–85.

Stanley, T. D, and Jarrell, S. B. (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys*. 3(2), 161–70.

Stapleford, T. A. (2011). Reconceiving quality: Political economy and the rise of hedonic price indexes. *History of Political Economy*, 43(1), 310-28.

StataCorp (2017). *Stata Statistical Software: Release 15*, College Station, TX.

Stone, R. (1956). *Quantity and price indexes in National Accounts*. Paris: Organization for European Economic Cooperation.

Triplett, J. (2006). *Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products*. Paris: Organization for European Economic Cooperation.

Tukey, J. W. (1977). *Exploratory data Analysis*. Reading, MA: Addison-Wesley.

van Garderen, K. J., and Shah, C. (2002). Exact interpretation of dummy variables in semi-logarithmic equations. *Econometrics Journal,* 5(1), 149-59.

Waugh, F. V. (1928). Quality factors influencing vegetable prices. *Journal of Farm Economics,* 10(2), 185-96.

Wooldridge, J. M. (1991). On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. *Journal of Econometrics,* 47(1), 5-46.

Yoshida, J., and Sugiura, A. (2015). The effects of multiple green factors on condominium prices. *Journal of Real Estate Finance and Economics,* 50(3), 412-37.

Zhang, L., and Yi, Y. (2017). Quantile house price indices in Beijing. *Regional Science and Urban Economics,* 63, 85-96.

Zheng, S., Jing, W., Kahn, M. E., and Deng, Y. (2012). The nascent market for "green" real estate in Beijing. *European Economic Review,* 56(5), 974-84.

Zietz, J., Sirmans, S. and Smersh, G.T. (2008). The impact of inflation on home prices and the valuation of housing characteristics across the price distribution. *Journal of Housing Research,* 17(2), 119-37.

Ziliak, S. T. and McCloskey, D. N. (2004). Size matters: The standard error of regressions in the American Economic Review. *Journal of Socio-Economics*, 33(5), 527-46.

Appendices

I - Tax forms

Real Estate Transfer Tax (IMT) form:

ANTES DE PREENCHER LEIA ATENTAMENTE TODO O IMPRESSO E CONSULTE AS INSTRUÇÕES

MINISTÉRIO DAS FINANÇAS DIRECÇÃO-GERAL DOS IMPOSTOS	**IMPOSTO MUNICIPAL SOBRE AS TRANSMISSÕES ONEROSAS DE IMÓVEIS (IMT)** DECLARAÇÃO PARA LIQUIDAÇÃO (Modelo 1)	01 SERVIÇO DE FINANÇAS ONDE É APRESENTADA A DECLARAÇÃO Cod.:

I — IDENTIFICAÇÃO DO SUJEITO PASSIVO

02 NIF / NIPC: | 03 Nome:
04 Tel/Tlm: | 05 E-Mail: @
06 Domicílio Fiscal: País/Território/Região: | 07 Estado Civil: | 08 Regime Casamento: | 09 NIF Cônjuge:

II — IDENTIFICAÇÃO DO FACTO TRIBUTÁRIO

10 CÓDIGO: DESCRIÇÃO: | 11 COD. CADUCIDADE ISENÇÃO: | 12 DATA: / /

III — TITULAR(ES) DO(S) BENS OU DIREITOS TRANSMITIDOS (Preencher Anexo I, caso existam mais intervenientes)

13 NIF - NIPC	14 Nome	15 Est.civil	16 Reg. Casamento	17 NIF Cônjuge

IV — 18 Nº. Ordem do Bem: 1 — IDENTIFICAÇÃO DO BEM (Preencher Anexo II, caso existam mais bens)

19 Tipo (R/U): | 20 Município: | 21 Freguesia: | 22 Artigo: | 23 Fracção/Secção: | 24 Árv/Col:
25 LOCALIZAÇÃO:
Confrontações: 26 Norte: | 27 Nascente: | 28 Sul: | 29 Poente:
30 Área(m²): | 31 Destino do Bem: | 32 Ónus/Encargos: Vistas, Passagens, Águas | 33 Tipo Bem:
34 Data de Arrendamento: | 35 Valor da Renda / Pensão: . . , | 36 Valor das Benfeitorias / Prédios em Construção: . . , | 37 Valor partes integrantes: . . ,
IDENTIFICAÇÃO DO DIREITO | 38 Tipo de direito: | 39 Período: | 40 Idade/Duração do contrato:
41 Observações:

V — FACTO TRIBUTÁRIO (Preencher Anexo III, caso o número de linhas seja insuficiente)

42 Nº ORDEM DO BEM: | 43 NIF/NIPC: | 44 QUOTA/PARTE: | 45 VALOR DO ACTO OU CONTRATO: . . ,
ONUS/ENCARGOS: 46 CÓDIGO: | 47 VALOR: . . , | BENEFÍCIOS FISCAIS: 48 CÓDIGO: | 49 VALOR: . . ,
50 PREÇO PREVISTO: . . , | 51 PERMUTA: | 52 VALOR ABATIMENTOS: . . , | 53 Identificador(es) da(s) Liquidação(ões) IMT:

VI — PARTES SOCIAIS OU QUOTAS NAS SOCIEDADES EM NOME COLECTIVO, EM COMANDITA SIMPLES OU POR QUOTAS

54 NIPC SOCIEDADE:	55 % CAPITAL PREVIAMENTE DETIDA	56 % CAPITAL DETIDA PELO CÔNJUGE	57 % CAPITAL JÁ TRIBUTADA	58 NIF DO SÓCIO ALIENANTE:	59 % CAPITAL A ADQUIRIR

VII — DISCRIMINAÇÃO DE OUTROS VALORES QUE INTEGRAM O ACTO OU CONTRATO

60 VALOR DOS MÓVEIS EM CASO DE PERMUTA: . . , | 61 VALOR GLOBAL DO ACTO OU CONTRATO: . . ,

VIII — OUTROS ELEMENTOS REFERENTES AO FACTO TRIBUTÁRIO

62

IX — DOCUMENTOS ANEXOS À DECLARAÇÃO (Quant.)

63 Declaração de Inscrição / Actualização da Matriz
64 Anexo I | 66 Anexo III
65 Anexo II | 67 Outros documentos

X — ENCERRAMENTO DA DECLARAÇÃO

A declaração corresponde à verdade e não houve qualquer omissão
68 Local e Data: , / /
O Declarante (assinatura)
Se a declaração for apresentada por um representante, gestor de negócios ou pelo cabeça-de-casal indique:
Nome:
69 NIF:

XI — PARA USO EXCLUSIVO DO SERVIÇO DE FINANÇAS

70 N.º de Registo da Declaração:
O FUNCIONÁRIO
Data: / /
Nome:____________________
____
Rubrica
CARIMBO DE RECEPÇÃO

DOCUMENTO VÁLIDO PARA ENTREGA NOS SERVIÇOS DE FINANÇAS

Local Property Tax (IMI) form:

ANTES DE PREENCHER LEIA ATENTAMENTE TODO O IMPRESSO E CONSULTE AS INSTRUÇÕES

MINISTÉRIO DAS FINANÇAS **DIRECÇÃO GERAL DOS IMPOSTOS**	**IMPOSTO MUNICIPAL SOBRE IMÓVEIS (IMI)** **DECLARAÇÃO PARA INSCRIÇÃO OU ACTUALIZAÇÃO DE PRÉDIOS URBANOS NA MATRIZ** (Modelo 1)	**01** SERVIÇO DE FINANÇAS DA ÁREA DA SITUAÇÃO DO PRÉDIO Cod.:

I — TITULAR do PRÉDIO ou FRACÇÃO (Caso não seja único proprietário ou titular da propriedade plena, preencha este quadro indicando um dos titulares e o ANEXO I, no qual deve indicar todos os titulares)

Nome / Designação:

02 NIF / NIPC:

03 Tipo de Titular: ☐ 1-Único proprietário ☐ 2-Comproprietário ☐ 3-Usufrutuário ☐ 4-Superficiário

04 Domicílio Fiscal: ☐ País/Região/Território:

05 O prédio é bem comum do casal ? ☐ Sim ☐ Não 06 Se assinalou Sim indique o NIF do Cônjuge:

07 Tel/Tlm: E-Mail : @

II — MOTIVO DA ENTREGA DA DECLARAÇÃO (Indique com um X)

08		Prédio Novo	09		Prédio Melhorado/Modificado	10		Prédio Melhorado / Modificado /Reconstruído	11		Anexo à declaração de IMT ou Imp. do Selo
12		Prédio Omisso	13		Pedido de Avaliação	14		Mudança da Afectação do Prédio	15		1ª Transmissão na Vigência do IMI

III — IDENTIFICAÇÃO MATRICIAL (Indique o(s) artigo(s) em que o prédio se encontra inscrito na matriz)

Tipo U / R		Freguesia		Artigo		Fracção Autónoma / Secção		Árvore / Colonia
17	18		19		20		21	
22	23		24		25		26	
27	28		29		30		31	

IV — SITUAÇÃO DO PRÉDIO

32 Av./ Rua/ Praça: 33 N.º 34 Lote 35 Andar:

36 Lugar: 37 Código Postal: 38 Freguesia:

39 Descrito na Conserv. Do Registo Predial de: 40 Sob o registo n.º:

CONFRONTAÇÕES (Preencher caso não haja número de polícia)

41 Norte: 42 Sul:

43 Nascente: 44 Poente:

V — TIPO DE PRÉDIO A AVALIAR (Assinale com X o tipo de prédio e indique as áreas em m2)

45 Fracção Autónoma de Prédio em Regime de Propriedade Horizontal (Preencha os campos de 52, 54 a 57 e 59 a 61)

46 Terrenos para Construção (Preencha os campos 55 e 57 a 59)

47 Loteamento de Terrenos para Construção (Preencha o Anexo III)

48 Prédio em Propriedade Total sem Andares nem Divisões Susceptíveis de Utilização Independente (Preencha os campos de 52 a 55 e 57 a 60)

49 Prédio em Regime de Propriedade Horizontal (Preencha o Anexo II)

50 Prédio em Propriedade Total com Andares ou Divisões Susceptíveis de Utilização Independente (Preencha o Anexo II)

51 Outros: Indique o Tipo ☐ , preencha os campos de 53 a 55 e 57 a 60; e indique o Custo de Construção + Valor do Terreno: €

52 Afectação	53 N.º de Pisos	54 Tipologia / N.º de Divisões	55 Área Total do Terreno	56 Área do Terreno Integrante da Fracção	57 Área de Implantação do Prédio	58 Área Bruta de Construção	59 Área Bruta Dependente	60 Área Bruta Privativa	61 Permilagem da Fracção

1

ANTES DE PREENCHER LEIA ATENTAMENTE TODO O IMPRESSO E CONSULTE AS INSTRUÇÕES

62	Elementos de Qualidade e Conforto para Prédio, parte de Prédio ou Fracção destinados a **Habitação** (Assinale com X os elementos que o seu prédio possui)	63	Elementos de Qualidade e Conforto para Prédio, parte de Prédio ou Fracção destinados a **Habitação** (Assinale com X os elementos que o seu prédio NÃO possui)
	1 – Moradias unifamiliares		11 - Inexistência de cozinha
	2 – Localização em condomínio fechado		12 - Inexistência de instalações sanitárias
	3 – Garagem individual		13 - Inexistência de rede pública ou privada de água
	4 – Garagem colectiva		14 – Inexistência de rede pública ou privada de electricidade
	5 – Piscina individual		15 - Inexistência de rede pública ou privada de gás
	6 – Piscina colectiva		16 - Inexistência de rede pública ou privada de esgotos
	7 – Campo de ténis		17 - Inexistência de ruas pavimentadas
	8 – Outros equipamentos de lazer		18 - Existência de áreas inferiores às regulamentares (RGEU)
	9 – Sistema central de climatização		19 - Inexistência de elevador em edifícios com mais de 3 pisos
	10 – Elevadores em edifícios de menos de 4 pisos		
	27 – Utilização de técnicas ambientalmente sustentáveis, activas ou passivas		

64	Elementos de Qualidade e Conforto para Prédio, parte de Prédio ou Fracção destinados a **Comércio, Indústria e Serviços** (Assinale com X os elementos que o seu prédio possui)	65	Elementos de Qualidade e Conforto para Prédio, parte de Prédio ou Fracção destinados a **Comércio, Indústria e Serviços** (Assinale com X os elementos que o seu prédio NÃO possui)
	20 – Localização em centro comercial		12 - Inexistência de instalações sanitárias
	21 – Localização em edifícios destinados a escritórios		13 - Inexistência de rede pública ou privada de água
	22 – Existência de elevador(es) e/ou escada(s) rolante(s)		14 - Inexistência de rede pública ou privada de electricidade
	9 – Sistema central de climatização		16 - Inexistência de rede pública ou privada de esgotos
	27 – Utilização de técnicas ambientalmente sustentáveis, activas ou passivas		17 - Inexistência de ruas pavimentadas
			19 - Inexistência de elevador em edifícios com mais de 3 pisos

VI	OUTROS ELEMENTOS	
66	Data da Licença de Utilização :	___/___/_____
67	Data de Conclusão das Obras:	___/___/_____
68	Data de Passagem a Urbano:	___/___/_____
69	Data de Ocupação:	___/___/_____
70	Início da Construção da Obra:	___/___/_____
71	Idade do Prédio (n.º de anos):	

VII	DOCUMENTOS JUNTOS À DECLARAÇÃO (Quantidade)				
72		Licença de utilização	78		Planta(s) do(s) edifício(s)
73		Escritura de Propriedade Horizontal	79		Alvará de loteamento
74		Alvará de licença ou autorização de construção	80		Contrato(s) de arrendamento
75		Planta(s) de localização / croquis	81		Anexo I
76		Planta de Implantação do(s) Edifício(s)	82		Anexo II
77		Projecto ou viabilidade construtiva	83		Anexo III

VIII	ENCERRAMENTO DA DECLARAÇÃO
	A declaração corresponde à verdade e não houve qualquer omissão
84	Local e Data: , / /
	O Declarante (assinatura)
	Se a declaração for apresentada por um representante, gestor de negócios ou pelo cabeça de casal indique: Nome:
85	NIF:

IX	PARA USO EXCLUSIVO DO SERVIÇO DE FINANÇAS
86	NIP:
87	N.º de Registo da Declaração:

O FUNCIONÁRIO	CARIMBO DE RECEPÇÃO
Data: / / Nome:__________________ ______________________ Rubrica	

2

II - Energy Performance Certificate

Model used from January 2009 to November 2013:

Certificação Energética
e Ar Interior
EDIFÍCIOS

Nº CER 1234567/2007

CERTIFICADO DE DESEMPENHO ENERGÉTICO E DA QUALIDADE DO AR INTERIOR

TIPO DE EDIFÍCIO: EDIFÍCIO HABITAÇÃO UNIFAMILIAR / FRACÇÃO AUTÓNOMA DE EDIF. MULTIFAMILIAR

Morada / Situação:

Localidade ______ Freguesia ______

Concelho ______ Região ______

Data de emissão do certificado ______ Validade do certificado ______

Nome do perito qualif. ______ Número do perito qualif. ______

Imóvel descrito na ☐ Conservatória do Registo Predial de ______

sob o nº ☐ Art. matricial nº ______ Fracção autón. ______

Este certificado resulta de uma verificação efectuada ao edifício ou fracção autónoma, por um perito devidamente qualificado para o efeito, em relação aos requesitos previstos no Regulamento das Características de Comportamento Térmico dos Edifícios (RCCTE, Decreto-Lei 80/2006 de 4 de Abril), classificando o imóvel em relação ao respectivo desempenho energético. Este certificado permite identificar possíveis medidas de melhoria de desempenho aplicáveis à fracção autónoma ou edifício, suas partes e respectivos sistemas energéticos e ventilação, quer no que respeita ao desempenho energético e à qualidade do ar interior.

1. ETIQUETA DE DESEMPENHO ENERGÉTICO

INDICADORES DE DESEMPENHO

Necessidades anuais globais estimadas de energia primária para climatização e águas quentes ☐ kgep/m².ano

Valor limite máximo regulamentar para as necessidades anuais globais de energia primária para climatização e águas quentes (limite inferior da classe B⁻) ☐ kgep/m².ano

Emissões anuais de gases de efeito estufa associadas à energia primária para climatização e águas quentes ☐ toneladas de CO_2 equivalentes por ano

CLASSE ENERGÉTICA

2. DESAGREGAÇÃO DAS NECESSIDADES NOMINAIS DE ENERGIA ÚTIL

Necessidades nominais de energia útil para...	Valor estimado para as condições de conforto térmico de referência	Valor limite regulamentar para as necessidades anuais
Aquecimento	kWh/m².ano	kWh/m².ano
Arrefecimento	kWh/m².ano	kWh/m².ano
Preparação das águas quentes sanitárias	kWh/m².ano	kWh/m².ano

NOTAS EXPLICATIVAS

As necessidades nominais de energia útil correspondem a uma previsão da quantidade de energia que terá de ser consumida por m² de área útil do edifício ou fracção autónoma para manter o edifício nas condições de conforto térmico de referência e para preparação das águas quentes sanitárias necessárias aos ocupantes. Os valores foram calculados para condições convencionais de utilização, admitidas como idênticas para todos os edifícios, de forma a permitir comparações objectivas entre diferentes imóveis. Os consumos reais podem variar bastante dos indicados e dependem das atitudes e padrões de comportamento dos utilizadores.

As necessidades anuais globais de energia primária (estimadas e valor limite) resultam da conversão das necessidades nominais de energia útil em kilogramas equivalente de petróleo por unidade (kgep) de área útil do edifício, mediante aplicação de factores de conversão específicos para a(s) forma(s) de energia utilizada(s) (0,290 kgep/kWh para electricidade e 0,086 kgep/kWh para combustíveis sólido, líquido ou gasoso) e tendo em consideração a eficiência dos sistemas adoptados ou, na sua indefinição, sistemas convencionais de referência.

As emissões de CO_2 equivalente traduzem a quantidade anual estimada de gases de efeito de estufa que podem ser libertados em resultado da conversão de uma quantidade de energia primária igual às respectivas necessidades anuais globais estimadas para o edifício, usando o factor de conversão de 0,0012 toneladas equivalentes de CO_2 por kgep.

A classe energética resulta da razão entre as necessidades anuais globais estimadas e as máximas admissíveis de energia primária para aquecimento, arrefecimento e para preparação de águas quentes sanitárias no edifício ou fracção autónoma. O melhor desempenho corresponde à classe A⁺, seguida das classes A, B, B⁻, C e seguintes, até à classe G de pior desempenho. Os edifícios com licença ou autorização de construção posterior a 4 de Julho de 2006 apenas poderão ter classe energética igual ou superior a B⁻. Para mais informações sobre o desempenho energético, sobre a qualidade do ar interior e sobre a classificação energética de edifícios, consulte www.adene.pt

Entidade gestora ADENE AGÊNCIA PARA A ENERGIA — Entidades supervisoras do SCE Direcção Geral de Energia e Geologia — Agência Portuguesa do Ambiente — 1/4

Model used from December 2013 onwards:

Certificação Energética e Ar Interior EDIFÍCIOS

Certificado Energético

Edifício de Habitação

SCE1234567890
Válido até 01/12/2013

IDENTIFICAÇÃO POSTAL

Morada AVª FONTES PEREIRA DE MELO, Nº51 A 51-G
Localidade LISBOA
Freguesia S. SEBASTIÃO DA PEDREIRA
Concelho LISBOA — GPS 39.7329, -7.0000

IDENTIFICAÇÃO PREDIAL/FISCAL

5ª Conservatória do Registo Predial de LISBOA
Nº de Inscrição na Conservatória 816
Artigo Matricial nº 898 — Fração Autónoma K

INFORMAÇÃO ADICIONAL

Área interior útil de Pavimento 320 m²

Este certificado apresenta a classificação energética deste edifício ou fração. Esta classificação é calculada comparando o desempenho energético deste edifício nas condições atuais, com o desempenho que este obteria nas condições mínimas (com base em valores de referência) a que estão obrigados os edifícios novos. Obtenha mais informação sobre a certificação energética no site da ADENE em www.adene.pt.

INDICADORES DE DESEMPENHO

Determinam a classe energética do edifício e a eficiência na utilização de energia, incluindo o contributo de fontes renováveis. São apresentados comparativamente a um valor de referência e calculados em condições padrão.

Aquecimento Ambiente

A referência para este edifício	200 kWh/m².ano
O consumo deste edifício	70 kWh/m².ano
Renovável	50 %

65% MAIS eficiente que a referência

Arrefecimento Ambiente

A referência para este edifício	20 kWh/m².ano
O consumo deste edifício	21 kWh/m².ano
Renovável	50 %

5% MENOS eficiente que a referência

Água Quente Sanitária

A referência para este edifício	30 kWh/m².ano
O consumo deste edifício	30 kWh/m².ano
Renovável	0 %

0% IGUAL à referência

CLASSE ENERGÉTICA

Mais eficiente

- A+ 0% a 25%
- A 26% a 50%
- B 51% a 75%
- B- 76% a 100% — Mínimo: Edifícios Novos
- C 101% a 150% — Mínimo: Grandes Intervenções
- D 151% a 200%
- E 201% a 250%
- F Mais de 251%

Menos eficiente

D 173%

ENERGIA RENOVÁVEL

Contributo de energia renovável no consumo de energia deste edifício.

15%

EMISSÕES DE CO_2

Emissões de CO_2 estimadas devido ao consumo de energia.

0,8 toneladas/ano

Entidade Gestora: ADENE AGÊNCIA PARA A ENERGIA

Entidade Fiscalizadora: Direcção Geral de Energia e Geologia

1 de 6

III - Heteroskedasticity-robust tests

This appendix describes the procedure employed to compute the Chow (1960), Ramsey (1969) and Breusch and Pagan (1979) type tests that were used to support the derivation of the hedonic price models used in this thesis (Chapter 4). The tests are based on heteroskedasticity-robust LM statistics suggested by Wooldridge (1991). The steps for the compilation of the LM statistics follow Wooldridge (2002: 243-255). The statistical tests were programmed using the SAS software base version 9.4 (SAS Institute Inc, 2015).

The first step of the tests involves the estimation of the (restricted) model of interest by OLS. Let us define it as:

$$\ln(p) = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + u_1, \qquad \text{(III.1)}$$

where p is the transaction price, $\ln(p)$ its logarithm transformation, x_k^* refers to the k explanatory variable, β_k corresponds to the parameter of the hedonic model and u_1 represents the error term of the econometric model.

For the *Chow (1960) type test*, the restricted model (III.1) is tested against the following alternative:

$$\ln(p) = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + \gamma_0 . S + \sum_{k=1}^{K} \gamma_k . (x_k^* . S) + u_3, \qquad \text{(III.2)}$$

where S is the dummy variable used to identify the possible structural break (new versus existing and apartments versus houses). The null, or parameter stability hypothesis, is defined as $H_0 : \gamma_0 = \gamma_1 = \cdots = \gamma_k = 0$. In order to obtain a LM statistic it is necessary to take the following steps:

1. Run the additional *k+1* regressions, and obtain the residuals:

$$S = \delta_{0,0} + \sum_{k=1}^{K} \delta_{k,0} . x_{k,0}^* + u_{0,4}$$

$$\ldots \qquad \text{(III.3)}$$

$$S . x_k^* = \delta_{0,k} + \sum_{k=1}^{K} \delta_{k,k} . x_k^* + u_{k,4};$$

2. Build *k+1* new variables, $w_j = \hat{u}_1 . \hat{u}_{j,4};\ j = 0, \ldots, k;$

3. Run a regression with no constant term, where the dependent variable is a column with 1's:

$$a = \theta_0 . w_0 + \cdots + \theta_k . w_k + u_5; \tag{III.4}$$

4. Finally, the LM statistic for this test is equal to the sample size n subtracted by the sum of square residuals taken from (III.4):

$$TStat_3 = n - SSR. \tag{III.5}$$

Under the null of coefficient stability (over apartments and houses and new and existing dwellings), the statistic (III.5) asymptotically follows a Chi-squared distribution with *k+1* degrees of freedom. If the computed *p*-value is smaller than, say, 0.05, the assumption on coefficient stability over the new/existing and apartments/houses strata is rejected.

For the *Ramsey (1969) type test*, the restricted model (III.1) is tested against the following alternative model:

$$\ln(p) = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + \gamma_0 . \widehat{\ln(p)}^2 + u_6, \tag{III.6}$$

where $\widehat{\ln(p)}^2$ is the square of the predicted dependent variable, which has been estimated by (III.1). The null, or correct functional form, is defined as $H_0 \colon \gamma_0 = 0$. In order to obtain a robust LM statistic for this test, it is necessary to take the following steps:

1. Run the additional regression, and obtain the residuals:

$$\widehat{\ln(p)}^2 = \delta_0 + \sum_{k=1}^{K} \delta_k . x_k^* + u_7; \tag{III.7}$$

2. Build a new variable, $w = \hat{u}_1 . \hat{u}_7$;
3. Run the regression with no constant term, where the dependent variable is a column with 1's:

$$a = \theta . w + u_8; \tag{III.8}$$

4. Finally, the LM statistic for this test is equal to the sample size n subtracted by the sum of square residuals taken from (III.8):

$$TStat_3 = n - SSR. \tag{III.9}$$

Under the null (of a correct functional form), the statistic (III.9) asymptotically follows a Chi-squared distribution with one degree of freedom. If the computed *p*-value is smaller than 0.05, then the functional form is rejected.

Finally, for the heteroskedasticity *Breusch and Pagan (1979) type test*, the following equation needs to be run using the squared OLS residuals obtained from (III.1):

$$\hat{u}_1^2 = \beta_0 + \sum_{k=1}^{K} \beta_k . x_k^* + u_2, \qquad \text{(III.10)}$$

The LM statistic for this test is obtained by multiplying the sample size n by the R^2 of equation (III.10):

$$TStat_1 = n.R^2. \qquad \text{(III.11)}$$

Under the null (of a homeskedastic error term), the statistic (III.11) asymptotically follows a Chi-squared distribution with *k* degrees of freedom. If the computed *p*-value is smaller than 0.05, then the homeskedasticity hypothesis is rejected for (III.1).

IV - Variables used in regression analysis

This appendix provides the description of the variables that were used in hedonic regression models. The list of variables is organized in alphabetical order.

Explanatory variable	Variable description
DABSGAS	A dummy variable that assumes the value 1 when the residential property is not connected to public or private gas distribution networks. It is taken from the IMI quality and comfort element (*Inexistência de rede pública ou privada de gás*), which identifies this situation and that is used in the calculation of fiscal appraisal property values.
DABSLIFT	A dummy variable that assumes the value 1 when the residential unit is in a building with more than four floors and that does not have an elevator. It taken from the IMI quality and comfort element (*Inexistência de elevador em edifícios com mais de 3 pisos*), which identifies this feature and that is used in the calculation of fiscal appraisal property values.
DBADCONSERVATION	A dummy variable that assumes the value 1 when the residential unit has a deficient conservation condition. It is taken from the IMI quality and comfort element (*Estado deficiente de conservação*), which identifies the conservation condition of the building and that is used in the derivation of fiscal appraisal property values.
DBADLOC	A dummy variable that assumes the value 1 when the residential unit is located in an extremely bad location. It is taken from IMI's location coefficient (*Coeficiente de Localização*), which is used in the derivation of fiscal appraisal property values.
DBGAPRTXCPL	A dummy variable that assumes the value 1 for all apartments with more than 250 square meters, more than four bedrooms and that are located in an extremely good location.
DBIGAPRT	A dummy variable that assumes the value 1 for all apartments with more than 250 square meters and more than four bedrooms.
DCONSTPi	A set of four dummy variables identifying the building construction technology time period in which the residential unit was first completed (i.e., before 1960, from 1961 to 1990, from 1991 to 2006 and after 2006).
DCONSTQi	A set of three dummy variables identifying the construction quality of the residential unit (e.g., quality of the project, thermal insulation, acoustic insulation, quality of building materials used at latter construction works phases). It is taken from the IMI quality and comfort element (*Qualidade construtiva*), which identifies this dwelling attribute. It is a variable that is taken into account by appraisers in the calculation of property values for fiscal purposes.
DCSYSTEM	A dummy variable that assumes the value 1 when the residential unit includes a central heating and/or air-conditioning system. It is taken from the IMI and comfort quality element (*Sistema central de climatização*), which signals this dwelling feature. It is a variable that is taken into account by appraisers in the calculation of property values for fiscal purposes.
DDISTRCAP	A dummy variable that assumes the value 1 when the residential unit is located in a capital of a district. A district is a first-level administrative subdivision of Portugal, which divides the country's mainland into 18 sub-regions. For the construction of this dummy, the capitals of the *Madeira* and *Açores* islands were considered as their district capitals.

Explanatory variable	Variable description
DLX	A dummy variable that assumes the value 1 when the residential unit is located in *Lisboa*, the capital of Portugal.
DENERGYAB	A dummy variable that assumes the value 1 when the EPC of the residential unit is either A^{+}, A, B or B^{-} . It is derived from ADENE's records and signals all properties that have annual energy needs that are estimated to be the same of or lower than reference standard consumption values.
DEXCPLOC	A dummy variable that assumes the value 1 when the residential unit is located in an extremely good location. As DBADLOC, it is taken from IMI's location coefficient (*Coeficiente de Localização*), which is used in the derivation of fiscal appraisal property values.
DGRFLOORENOV	A dummy variable that assumes the value 1 when the residential unit is a renovated house with less than 120 square meters.
DIRREGAREA	A dummy variable that assumes the value 1 when the residential unit has non-standard areas, as defined by the Portuguese building code. It is taken from the IMI quality and comfort element identifying the existence of non-standard areas (*Existência de áreas inferiors às regulamentares*), which is used in the derivation of fiscal appraisal property values.
DMROOMS	A dummy variable that assumes the value 1 when an apartment has four or more bedrooms.
DPORTO	A dummy variable that assumes the value 1 when the residential unit is located in Porto, the second largest city in Portugal.
DPARKING	A dummy variable that assumes the value 1 when the residential unit has parking facilities. It is taken from the IMI quality and comfort elements (*Garagem individual* and *Garagem coletiva*), which identify the existence of individual and collective parking facilities and that are used in the calculation of fiscal appraisal property values.
DCOND	A dummy variable that assumes the value 1 when the residential unit is located in a private condominiums. It is taken from the IMI quality and comfort element (*Localização em condomínio fechado*), which identifies this situation and that is used in the calculation of fiscal appraisal values.
DPRIVPARK	A dummy variable that assumes the value 1 when the residential unit has individual parking facilities. It is derived from the IMI's quality and comfort element (*Garagem individual*), which identifies this situation and that is used in the calculation of fiscal appraisal values.

Explanatory variable	Variable description
DREGIONi	A set of seven dummy variables identifying the following geographical areas: (1) North, without the metropolitan area of Porto (DREGION1), (2) metropolitan area of Porto (DREGION2), (3) *Centro* region (DREGION3), (4) metropolitan area of *Lisbo*a (DREGION4), (5) *Alentejo* region (DREGION5), (6) *Algarve* (DREGION6), and (7) *Madeira* and *Açores* islands (DREGION7).
DRENOV	A dummy variable that assumes the value 1 when the residential unit has been improved or renewed (*Prédio melhorado/modificado/reconstruído*). It is taken from the IMI variable that identifies the situation in which an improved or renewed dwelling has generated a new IMI tax form.
DSCENICi	A set of three dummy variables identifying the quality of the landscape of the area in which the residential unit is located. It is taken from the IMI quality and comfort element (*Localização excepcional*), which identifies this feature and that is used in the calculation of fiscal appraisal property. This element should not be confused with IMI's location coefficient, as the former essentially measures the scenic value and the visual prominence of the location (e.g., if the residential unit has a seafront) and the latter the quality of public and private services and goods available in the area.
DSEA	A dummy variable that assumes the value 1 when a property is located in parish that has access to the sea.
DSMALLBEEDR	A dummy variable that assumes the value 1 for all house with less than three bedrooms.
DSWIMM	A dummy variable that assumes the value 1 when the residential unit has swimming facilities. It is taken from the IMI quality and comfort elements (*Piscina individual* and *Piscina coletiva*), which identify the existence of individual and collective swimming facilities and that is used in the calculation of fiscal appraisal property values.
Di	A set of five dummy variables identifying the year in which the transactions take place. The oldest year (2009) is identified by i = 1, and the more recent one (2013) by i = 5.
SQRTGRFA	The square root transformation of gross floor area (*Área bruta privativa*). The gross floor area corresponds to the sum of all covered areas, as measured from the outer perimeter of walls, which have the same use as the residential unit. It may include private balconies, attics and basements (as long as they are covered and used for residential purposes) and is taken from IMI's records.
SQRTDWELLTRANSA	The square root transformation of the number of complete years of a residential unit at transaction date. This variable was built combining IMT's transaction date, IMI's appraisal date and IMI's age of the dwelling at fiscal appraisal date, which is used in the calculation of fiscal appraisal property values.

Explanatory variable	Variable description
SQRTDEPFLOORA	The square root transformation of the dependent floor area of a residential unit *(Área bruta dependente)*. The dependent floor area corresponds to the sum of all covered areas, including those located outside of the residential unit, which provide accessory services to the main use of that same residential unit. Garages, attics and cellars constitute typical examples of dependent areas, This information is provided by IMI records and is used in the calculation of fiscal appraisal property values.
SQRTPLOTAREA	The square root transformation of the plot area of a residential unit. The plot area corresponds to the total uncovered land area, which is associated with an individual residential unit. This measure is net of the area in which the building of the residential unit sits on. Although much more common for houses, it is also possible to find apartments with positive plot areas (e.g., backyards). It is taken from IMI's records and it is used in the calculation of fiscal appraisal property values.

V - OLS parameter estimates

Existing apartments

Explanatory variables	Hypothesis 3		Hypothesis 4		Hypothesis 5		Hypothesis 6	
	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.
Constant term	10.096**	1426.2	10.103**	1414.1	10.090**	1265.7	10.097**	1425.1
DENERGYAB	0.118**	61.7	0.092**	25.5	0.130**	26.4	0.116**	58.5
D2010	-0.004	-1.8	-0.003	-1.2	-0.004	-1.6	-0.004	-1.8
D2011	-0.071**	-28.9	-0.081**	-27.3	-0.070**	-28.7	-0.071**	-28.8
D2012	-0.152**	-56.4	-0.164**	-51.3	-0.152**	-56.2	-0.152**	-56.4
D2013	-0.182**	-67.5	-0.200**	-62.1	-0.182**	-67.3	-0.182**	-67.5
DENERGYD2010	-	-	0.003	0.7	-	-	-	-
DENERGYD2011	-	-	0.039**	7.7	-	-	-	-
DENERGYD2012	-	-	0.045**	7.8	-	-	-	-
DENERGYD2013	-	-	0.065**	11.8	-	-	-	-
SQRTGRFA	0.144**	234.9	0.144**	235.0	0.144**	234.9	0.144**	234.6
SQRTDEPFLOORA	0.024**	49.0	0.024**	49.1	0.024**	49.0	0.024**	49.0
SQRTDWELLTRANSA	-0.030**	-25.6	-0.031	-26.1	-0.030**	-25.5	-0.030**	-25.6
DCSYSTEM	0.079**	24.1	0.079**	24.1	0.078**	23.7	0.078**	23.7
DABSLIFT	-0.071**	-22.4	-0.070**	-22.2	-0.071**	-22.6	-0.071**	-22.5
DCOND	0.058**	9.7	0.058**	9.8	0.058**	9.8	0.058**	9.8
DSWIMM	0.153**	32.2	0.151**	31.9	0.153**	32.1	0.153**	32.1
DPARKING	0.057**	25.2	0.058**	25.3	0.057**	25.2	0.058**	25.3
DCONSTP2	-0.113**	-32.5	-0.111**	-32.0	-0.106**	-21.1	-0.112**	-32.2
DCONSTP3	-0.155**	-28.2	-0.153**	-27.7	-0.145**	-22.3	-0.155**	-28.1
DCONSTP4	-0.144**	-16.1	-0.140**	-15.6	-0.147**	-15.1	-0.144**	-16.0
DENERGCONSTP2	-	-	-	-	-0.010	-1.8	-	-
DENERGCONSTP3	-	-	-	-	-0.039**	-5.8	-	-
DENERGCONSTP4	-	-	-	-	0.064**	4.6	-	-
DCONSTQ2	0.056**	16.5	0.057**	16.7	0.056**	16.4	0.040**	8.1
DCONSTQ3	0.139**	15.1	0.141**	15.4	0.138**	15.0	0.158**	6.2
DENERGCONSTQ2	-	-	-	-	-	-	0.033**	5.2
DENERGCONSTQ3	-	-	-	-	-	-	-0.021	-0.8
DREGION1	-0.365**	-120.5	-0.365**	-120.6	-0.366**	-120.6	-0.365**	-120.2
DREGION2	-0.296**	-125.1	-0.296**	-125.1	-0.297**	-125.3	-0.296**	-125.1
DREGION3	-0.252**	-98.6	-0.252**	-98.7	-0.252**	-98.7	-0.252**	-98.5
DREGION5	-0.039**	-5.9	-0.039**	-5.9	-0.038**	-5.8	-0.039**	-5.9
DREGION6	-0.010*	-2.9	-0.011*	-3.1	-0.010*	-2.9	-0.010*	-2.8
DREGION7	-0.005	-0.7	-0.005	-0.7	-0.005	-0.7	-0.005	-0.7
DSEA	0.113**	58.3	0.112**	58.3	0.113**	58.4	0.113**	58.3
DLX	0.349**	84.2	0.348**	84.2	0.348**	84	0.349**	84.2
DPORTO	0.331**	73.8	0.330**	73.7	0.330**	73.7	0.330**	73.6
DSCENIC2	0.100**	24.3	0.100**	24.3	0.100**	24.4	0.101**	24.5
DSCENIC3	0.266**	22.5	0.265**	22.4	0.266**	22.5	0.267**	22.5
DBADLOC	-0.171**	-39.9	-0.171**	-39.9	-0.171**	-39.8	-0.171**	-39.9
DEXCPLOC	0.316**	98.1	0.316**	98.2	0.316**	98.2	0.316**	98.1
Number of obs. used in estimation	149,920							
Regressions' adjusted R^2	0.6758		0.6763		0.6761		0.6759	
RESET type test								
LM test statistic	0.3282		2,9459		.0075		0.043	
p-value	0.567		0.086		0.931		0.836	

Notes:* p-value < 0.05; ** p-value <0.0001. A description of the variables is available in Appendix IV.

New apartments

Explanatory variables	*Hypothesis 3*		*Hypothesis 4*		*Hypothesis 5*		*Hypothesis 6*	
	Param. estimate	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*
Intercept	*10.182***	*1155.6*	*10.220***	*1121.1*	*10.184***	*1143*	*10.184***	*1151.8*
DENERGYAB	*0.123***	*50.4*	*0.065***	*15.8*	*0.122***	*44.4*	*0.119***	*45.3*
D2010	*0.002*	*1.0*	*-0.033***	*-6.7*	*0.003*	*1.1*	*0.002***	*1*
D2011	*-0.026***	*-8.7*	*-0.097***	*-16.0*	*-0.026***	*-8.6*	*-0.026***	*-8.6*
D2011	*-0.088***	*-23.2*	*-0.143***	*-21.2*	*-0.088***	*-23.2*	*-0.088***	*-23.2*
D2012	*-0.095***	*-22.6*	*-0.179***	*-22.7*	*-0.095***	*-22.6*	*-0.095***	*-22.6*
DENERGYD2010	-	-	*0.052***	*9.2*	-	-	-	-
DENERGYD2011	-	-	*0.103***	*14.8*	-	-	-	-
DENERGYD2012	-	-	*0.083***	*10.2*	-	-	-	-
DENERGYD2013	-	-	*0.124***	*13.5*	-	-	-	-
SQRTGRFA	*0.131***	*149.3*	*0.131***	*149.9*	*0.131***	*149*	*0.131***	*149.2*
SQRTDEPFLOORA	*0.033***	*52.4*	*0.033***	*52.6*	*0.033***	*52.4*	*0.033***	*52.3*
DMROOMS	*0.039***	*7.6*	*0.038***	*7.4*	*0.040***	*7.7*	*0.038***	*7.5*
DBIGAPRT	*0.186***	*6.9*	*0.186***	*6.9*	*0.187***	*7.0*	*0.186***	*6.9*
DBGAPRTXCPL	*0.160***	*3.2*	*0.159**	*3.2*	*0.151**	*3.1*	*0.161**	*3.2*
DCSYSTEM	*0.070***	*28.5*	*0.071***	*28.8*	*0.070***	*28.6*	*0.070***	*28.3*
DCOND	*0.065***	*11.5*	*0.066***	*11.6*	*0.065***	*11.5*	*0.065***	*11.5*
DSWIMM	*0.169***	*35.0*	*0.169***	*35.1*	*0.169***	*35.1*	*0.170***	*35.1*
DCONSTP2	*-0.099***	*-37.9*	*-0.099***	*-37.6*	*-0.099***	*-21.6*	*-0.099***	*-37.8*
DCONSTP3	*-0.341***	*-22.7*	*-0.336***	*-22.4*	*-0.353***	*-21.8*	*-0.342***	*-22.7*
DCONSTP4	*-0.373***	*-20.5*	*-0.366***	*-20.2*	*-0.405***	*-19.5*	*-0.374***	*-20.5*
DENERGCONSTP2	-	-	-	-	*-0.002*	*-0.3*	-	-
DENERGCONSTP3	-	-	-	-	*0.077*	*1.8*	-	-
DENERGCONSTP4	-	-	-	-	*0.152***	*3.9*	-	-
DCONSTQ2	*0.081***	*28.4*	*0.082***	*28.9*	*0.081***	*28.3*	*0.062***	*10.3*
DCONSTQ3	*0.137***	*21.9*	*0.136***	*21.8*	*0.136***	*21.9*	*0.128***	*7.1*
DENERGCONSTQ2	-	-	-	-	-	-	*0.025**	*3.7*
DENERGCONSTQ3	-	-	-	-	-	-	*0.011*	*0.6*
DREGION1	*-0.366***	*-101.0*	*-0.367***	*-101.5*	*-0.366***	*-101*	*-0.365***	*-101*
DREGION2	*-0.239***	*-71.8*	*-0.239***	*-71.9*	*-0.239***	*-71.8*	*-0.239***	*-71.8*
DREGION3	*-0.256***	*-82.7*	*-0.257***	*-83.0*	*-0.257***	*-82.7*	*-0.256***	*-82.5*
DREGION5	*-0.076***	*-10.1*	*-0.077***	*-10.3*	*-0.075***	*-10.0*	*-0.076***	*-10.1*
DREGION6	*-0.013**	*-3.1*	*-0.013**	*-3.1*	*-0.013**	*-3.3*	*-0.013**	*-3.1*
DREGION7	*-0.045***	*-6.3*	*-0.045***	*-6.4*	*-0.046***	*-6.3*	*-0.045***	*-6.2*
DSEA	*0.081***	*31.6*	*0.080***	*31.6*	*0.081***	*31.7*	*0.081***	*31.6*
DLX	*0.265***	*44.2*	*0.264***	*44.3*	*0.263***	*44.0*	*0.264***	*44.1*
DPORTO	*0.291***	*46.9*	*0.291***	*47.4*	*0.291***	*47.0*	*0.291***	*46.9*
DSCENIC2	*0.063***	*14.0*	*0.062***	*13.8*	*0.063***	*14.0*	*0.063***	*14.1*
DSCENIC3	*0.142***	*15.0*	*0.143***	*15.1*	*0.142***	*15.0*	*0.142***	*14.9*
DBADLOC	*-0.184***	*-36.7*	*-0.183***	*-36.4*	*-0.184***	*-36.7*	*-0.184***	*-36.7*
DEXCPLOC	*0.324***	*68.5*	*0.324***	*69.0*	*0.323***	*68.4*	*0.324***	*68.7*
Number of obs. used in estimation	*59,410*							
Regressions' adjusted R^2	*0.7334*		*0.7351*		*0.7336*		*0.7335*	
RESET type test								
LM test statistic	*2.731*		*4.3077*		*4.1533*		*2.3322*	
p-value	*0.098*		*0.038*		*0.042*		*0.127*	

Notes: p-value < 0.05; ** p-value <0.0001. A description of the variables is available in Appendix IV.*

Existing houses

Explanatory variables	*Hypothesis 3*		*Hypothesis 4*		*Hypothesis 5*		*Hypothesis 6*	
	Param. estimate	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*	*Param. estimate*	*Robust t-stat.*
Intercept	10.874**	602.5	10.881**	600.4	10.852**	580.5	10.874**	602.6
DENERGYAB	0.045**	7.3	0.004	0.4	0.106**	7.9	0.045**	7.1
D2010	0.007	1.0	0.003	0.4	0.006	1.0	0.006	1.0
D2011	-0.067**	-9.5	-0.079**	-9.9	-0.068**	-9.6	-0.067**	-9.5
D2012	-0.131**	-17.2	-0.144**	-16.8	-0.132**	-17.2	-0.131**	-17.1
D2013	-0.158**	-21	-0.178**	-21.4	-0.160**	-21.3	-0.158**	-21
DENERGYD2010	-	-	0.019	1.2	-	-	-	-
DENERGYD2011	-	-	0.055*	3.1	-	-	-	-
DENERGYD2012	-	-	0.060*	3.2	-	-	-	-
DENERGYD2013	-	-	0.102**	5.4	-	-	-	-
SQRTGRFA	0.082**	66.7	0.082**	66.7	0.082**	66.7	0.082**	66.7
SQRTDEPFLOORA	0.012**	18.7	0.012**	18.7	0.012**	18.8	0.012**	18.7
SQRTPLOTAREA	0.007**	28.6	0.007**	28.6	0.007**	29.0	0.007**	28.6
DIRREGAREA	-0.129**	-9.4	-0.128**	-9.4	-0.128**	-9.5	-0.128**	-9.4
DSMALLBEEDR	-0.083**	-13.5	-0.083**	-13.4	-0.083**	-13.4	-0.083**	-13.5
SQRTDWELLTRANSA	-0.015**	-7.6	-0.015**	-7.6	-0.015**	-7.6	-0.015**	-7.6
DCSYSTEM	0.095**	9.7	0.095**	9.8	0.085**	8.7	0.097**	9.9
DSWIMM	0.240**	24.5	0.240**	24.5	0.237**	24.2	0.239**	24.4
DPRIVPARK	0.089**	15.9	0.089**	15.9	0.088**	15.6	0.089**	15.9
DABSGAS	-0.089**	-17.1	-0.088**	-17.1	-0.087**	-16.9	-0.089**	-17.2
DCONSTP2	-0.088**	-10.3	-0.087**	-10.2	-0.073**	-7.1	-0.087**	-10.2
DCONSTP3	-0.167**	-14.1	-0.166**	-13.9	-0.147**	-11.4	-0.166**	-14.0
DCONSTP4	-0.310**	-19.1	-0.308**	-18.9	-0.273**	-15.9	-0.309**	-19.0
DENERGCONSTP2	-	-	-	-	-0.033*	-2.2	-	-
DENERGCONSTP3	-	-	-	-	-0.054*	-2.7	-	-
DENERGCONSTP4	-	-	-	-	-0.158**	-7.6	-	-
DCONSTQ2	0.024*	2.3	0.026*	2.4	0.025*	2.3	0.034*	2.7
DCONSTQ3	0.127*	3.8	0.127*	3.7	0.124*	3.7	0.040	0.8
DENERGCONSTQ2	-	-	-	-	-	-	-0.027	-1.4
DENERGCONSTQ3	-	-	-	-	-	-	0.163	2.6
DREGION1	-0.369**	-41.7	-0.369**	-41.6	-0.368**	-41.6	-0.370**	-41.7
DREGION2	-0.255**	-32.2	-0.254**	-32.2	-0.254**	-32.2	-0.255**	-32.3
DREGION3	-0.331**	-43.9	-0.331**	-43.9	-0.332**	-44	-0.331**	-43.9
DREGION5	-0.214**	-19.4	-0.214**	-19.5	-0.215**	-19.5	-0.214**	-19.4
DREGION6	0.020	1.9	0.020	1.9	0.022*	2.1	0.020**	1.9
DREGION7	-0.004	-0.2	-0.004	-0.2	-0.001	0.0	-0.004**	-0.2
DDISTRCAP	0.078**	8.6	0.078**	8.6	0.079**	8.7	0.078**	8.6
DSEA	0.159**	26.7	0.159**	26.8	0.159**	26.8	0.159**	26.7
DLX	0.299**	11.5	0.299**	11.5	0.303**	11.7	0.298**	11.5
DPORTO	0.278**	13.2	0.278**	13.2	0.284**	13.5	0.278**	13.2
DSCENIC 2	0.145**	9.8	0.144**	9.8	0.145**	9.8	0.146**	9.9
DSCENIC3	0.248**	8.0	0.245**	7.9	0.246**	7.9	0.247**	8.0
DBADLOC	-0.154**	-25.1	-0.155**	-25.1	-0.153**	-24.8	-0.154**	-25.0
DEXCPLOC	0.475**	33.8	0.475**	33.9	0.477**	33.8	0.476**	33.8
Number of obs. used in estimation	33,282							
Regressions' adjusted R^2	0.67		0.6703		0.6709		0.6701	
RESET type test								
LM test statistic	3.001		2.273		5.878		2.910	
p-value	0.0832		0.1317		0.0153		0.088	

Notes: p-value < 0.05; ** p-value <0.0001. A description of the variables is available in Appendix IV.*

New houses

Explanatory variables	Hypothesis 3		Hypothesis 4		Hypothesis 5		Hypothesis 6	
	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.	Param. estimate	Robust t-stat.
Intercept	10.899**	417.3	10.910**	414.8	10.896**	418.2	10.901**	416.7
DENERGYAB	0.055**	8.9	0.029*	2.6	0.078**	11.7	0.049**	7.2
D2010	0.008	1.2	0.002	0.2	0.007	1.0	0.009	1.2
D2011	-0.034**	-4.0	-0.042*	-3.8	-0.036**	-4.3	-0.033**	-3.9
D2012	-0.099**	-10.3	-0.125**	-10.2	-0.100**	-10.4	-0.099**	-10.3
D2013	-0.127**	-12.0	-0.153**	-11.2	-0.130**	-12.3	-0.128**	-12.0
DENERGYD2010	-	-	0.018	1.3	-	-	-	-
DENERGYD2011	-	-	0.022	1.3	-	-	-	-
DENERGYD2012	-	-	0.073*	3.8	-	-	-	-
DENERGYD2013	-	-	0.073*	3.4	-	-	-	-
SQRTGRFA	0.086**	43.1	0.086**	43.0	0.086**	42.8	0.086**	43.1
SQRTDEPFLOORA	0.012**	13.2	0.012**	13.2	0.012**	13.2	0.012**	13.2
SQRTPLOTAREA	0.006**	18.8	0.006**	18.8	0.006**	19.2	0.006**	18.9
DRENOV	-0.164**	-16.3	-0.164**	-16.2	-0.163**	-16.1	-0.165**	-16.3
DGRFLOORENOV	-0.084**	-5.6	-0.085**	-5.7	-0.085**	-5.7	-0.084**	-5.6
DIRREGAREA	-0.133**	-5.3	-0.133**	-5.2	-0.134**	-5.4	-0.133**	-5.3
DBADCONSERVATION	-0.147**	-4.6	-0.146**	-4.6	-0.132**	-4.2	-0.146**	-4.6
DCSYSTEM	0.071**	9.2	0.072**	9.3	0.069**	8.9	0.070**	9.1
DSWIMM	0.271**	25.1	0.270**	25.1	0.270**	25.0	0.271**	25.1
DPARKING	0.039**	5.1	0.039**	5.1	0.037**	4.8	0.039**	5.1
DABSGAS	-0.067**	-10.2	-0.068**	-10.3	-0.067**	-10.1	-0.067**	-10.2
DCOND	0.071**	5.0	0.069**	4.9	0.069**	4.8	0.071**	5.0
DCONSTP2	-0.055**	-8.0	-0.055**	-8.0	-0.049**	-5.6	-0.055**	-8.0
DCONSTP3	-0.161**	-10.6	-0.160**	-10.6	-0.145**	-9.3	-0.162**	-10.7
DCONSTP4	-0.309**	-17.8	-0.308**	-17.8	-0.274**	-15.1	-0.309**	-17.8
DENERGCONSTP2	-	-	-	-	-0.015	-1.1	-	-
DENERGCONSTP3	-	-	-	-	-0.112*	-2.8	-	-
DENERGCONSTP4	-	-	-	-	-0.158**	-4.8	-	-
DCONSTQ2	0.021*	2.4	0.021*	2.4	0.020*	2.3	0.001	0.1
DCONSTQ3	0.109**	5.3	0.106**	5.2	0.107**	5.3	0.063*	2.2
DENERGCONSTQ2	-	-	-	-	-	-	0.042*	2.6
DENERGCONSTQ3	-	-	-	-	-	-	0.084*	2.3
DREGION1	-0.400**	-36.9	-0.398**	-36.7	-0.399**	-36.8	-0.399**	-36.9
DREGION2	-0.233**	-21.6	-0.232**	-21.6	-0.233**	-21.7	-0.233**	-21.6
DREGION3	-0.325**	-34.2	-0.325**	-34.2	-0.327**	-34.4	-0.325**	-34.2
DREGION5	-0.211**	-13.7	-0.211**	-13.7	-0.212**	-13.8	-0.211**	-13.7
DREGION6	0.042*	3.4	0.043*	3.5	0.044*	3.5	0.043*	3.5
DREGION7	-0.057*	-2.6	-0.057*	-2.6	-0.056*	-2.5	-0.057*	-2.6
DDISTRCAP	0.117**	10.1	0.117**	10.1	0.118**	10.3	0.116**	10.1
DSEA	0.118**	15.8	0.118**	15.9	0.118**	15.9	0.117**	15.8
DLX	0.364**	7.6	0.364**	7.7	0.364**	7.5	0.366**	7.7
DPORTO	0.370**	8.2	0.368**	8.1	0.375**	8.3	0.363**	8.0
DSCENIC2	0.098**	5.3	0.096**	5.2	0.098**	5.3	0.098**	5.3
DSCENIC3	0.131*	3.4	0.129*	3.4	0.129*	3.4	0.130*	3.4
DBADLOC	-0.148**	-20.0	-0.148**	-20.0	-0.147**	-20.0	-0.148**	-20.1
DEXCPLOC	0.358**	18.1	0.358**	18.1	0.357**	18.0	0.357**	18.0
Number of obs. used in estimation	13,533							
Regressions' adjusted R^2	0.7531		0.7535		0.7542		0.7532	
RESET type test								
LM test statistic	3.7717		3.7084		6.7314		4.7109	
p-value	0.052		0.054		0.01		0.03	

Notes: * p-value < 0.05; ** p-value <0.0001. A description of the variables is available in Appendix IV.*

VI – Deriving error measurement parameters used in simulated list prices

This Appendix explains how the mean and standard deviation parameters of the measurement errors structures, which were used in the simulation of list prices (Chapter 5), were derived. Let us define LP, List Prices, as:

$$LP = p + e, \tag{VI.1}$$

where e, is the measurement error, which is assumed to be proportional to transaction prices (p).

$$LP = (1 + k).p. \tag{VI.2}$$

In (VI.2), k represents the proportion of upward bias that list prices are assumed to have in relation to transaction prices. Substituting (VI.2) into (VI.1), and making a simple rearrangement e, is equal to:

$$e = k.p \tag{VI.3}$$

Using this last expression, the mean and variance of e are calculated as:

$$\begin{aligned} \mu_e &= E[k.p] = k.\mu_p \\ \sigma_e^2 &= Var(k.p) = k^2.\sigma_p^2, \end{aligned} \tag{VI.4}$$

where μ_p and σ_p^2 are the mean and variance of transacted prices. Taking the square root of the variance, it is possible to have the standard deviation of the measurement error, σ_e, which is equal to:

$$\sigma_e = k.p. \tag{VI.4}$$

For reference, the next table provides the mean and standard deviation parameters that were used in the simulation of list prices. As mentioned in Chapter 5, the error measurement structures were applied to existing apartments, new apartments, existing houses and new houses. Except for the first case (i.e., ListPr1), all other simulated list prices are shown by transaction price quintile interval.

Mean and standard deviation parameters used in list price simulations

List Price \ Quintiles		Existing apartments		New apartments		Existing houses		New houses	
		μ_e (€)	σ_e (€)	μ_e (€)	σ_e (€)	μ_e (€)	σ_e (€)	μ_e (€)	σ_e (€)
ListPr1	-	*5,862*	*4,133*	*8,940*	*5,621*	*8,627*	*9,043*	*10,749*	*8,740*
ListPr2	*1st*	*425*	*83*	*721*	*150*	*408*	*106*	*578*	*188*
	2nd	*1,258*	*98*	*2,081*	*144*	*1,514*	*199*	*2,269*	*259*
	3rd	*4,035*	*287*	*6,439*	*380*	*5,630*	*558*	*7,627*	*517*
	4th	*7,519*	*720*	*11,331*	*860*	*10,919*	*1,025*	*13,891*	*1,195*
	5th	*17,627*	*8,633*	*25,251*	*11,948*	*30,158*	*22,133*	*33,631*	*19,722*
ListPr3	*1st*								
	2nd					*(*)*			
	3rd								
	4th	*15,038*	*1,440*	*22,662*	*17,19*	*21,838*	*2,050*	*27,783*	*2,391*
	5th	*35,255*	*17,266*	*50,502*	*23,897*	*60,316*	*44,265*	*67,262*	*39,444*
ListPr4	*1st*	*672*	*315*	*999*	*404*	*688*	*639*	*828*	*527*
	2nd	*1,554*	*710*	*2,467*	*981*	*1,924*	*1,229*	*3,001*	*1515*
	3rd	*4,307*	*2,064*	*7,018*	*3,001*	*6,583*	*4,260*	*8,949*	*4,339*
	4th	*7,110*	*3,640*	*10,883*	*47,02*	*11,543*	*6,965*	*14,043*	*7,292*
	5th	*14,026*	*9,938*	*20,321*	*13,526*	*23,153*	*23,313*	*25,498*	*21,617*

Note: (*) As explained in Section 5.2, the first three quintiles in Experiment 3 (ListPr3) have the same error structure parameters as in Experiment 2 (LisrPr2). As such, there is no need to replicate them in the table.

All the simulation work was carried out in SAS (SAS Institute Inc, 2015).

VII - Full regression outputs for omitted variable scenarios

Existing apartments

		Omitted scenario/variables					All omitted (average over 1,000 replications)				
	Benchamrk model	*Central heating and/or air conditioning*	*Visual quality*	*Location quality*	*construction quality*	*All omitted*	*n = 500*	*n = 1,000*	*n = 2,500*	*n = 5,000*	*n = 10,000*
Constant term	10.096**	10.103**	10.088**	10.070**	10.101**	10.078**	10.071	10.071	10.069	10.069	10.069
DENERGYAB	0.118**	0.125**	0.118	0.131**	0.120**	0.145**	0.147	0.147	0.146	0.146	0.146
D2010	-0.004	-0.004	-0.003**	-0.005*	-0.004*	-0.006*	-0.006	-0.006	-0.007	-0.007	-0.007
D2011	-0.071**	-0.071**	-0.070**	-0.072**	-0.072**	-0.073**	-0.074	-0.074	-0.073	-0.073	-0.074
D2012	-0.152**	-0.151**	-0.152**	-0.154**	-0.153**	-0.154**	-0.155	-0.155	-0.154	-0.154	-0.155
D2013	-0.182**	-0.182**	-0.181**	-0.182**	-0.183**	-0.181**	-0.179	-0.179	-0.182	-0.182	-0.182
SQRTGRFA	0.144**	0.144**	0.144**	0.144**	0.144**	0.147**	0.148	0.148	0.148	0.147	0.147
SQRTDEPFLOORA	0.024**	0.024**	0.024**	0.025**	0.024**	0.026**	0.026	0.026	0.026	0.026	0.026
SQRTDWELLTRANSA	-0.030**	-0.031**	-0.030**	-0.025**	-0.031**	-0.027**	-0.029	-0.029	-0.027	-0.027	-0.027
DCSYSTEM	0.079**	-	0.085**	0.087**	0.091**	-	-	-	-	-	-
DABSLIFT	-0.071**	-0.071**	-0.072**	-0.091**	-0.071**	-0.094**	-0.095	-0.095	-0.094	-0.095	-0.095
DCOND	0.058**	0.064**	0.055**	0.060**	0.061**	0.076**	0.077	0.077	0.078	0.076	0.078
DSWIMM	0.153**	0.153**	0.157**	0.184**	0.153**	0.192**	0.193	0.193	0.191	0.193	0.192
DPARKING	0.057**	0.060**	0.056**	0.068**	0.058**	0.073**	0.073	0.073	0.074	0.074	0.074
DCONSTP2	-0.113**	-0.119**	-0.111**	-0.105**	-0.117**	-0.123**	-0.124	-0.124	-0.128	-0.128	-0.128
DCONSTP3	-0.155**	-0.162**	-0.153**	-0.141**	-0.160**	-0.158**	-0.156	-0.156	-0.163	-0.163	-0.164
DCONSTP4	-0.144**	-0.150**	-0.145**	-0.101**	-0.147**	-0.116**	-0.114	-0.114	-0.122	-0.122	-0.123
DCONSTQ2	0.056**	0.067**	0.090**	0.064**	-	-	-	-	-	-	-
DCONSTQ3	0.139**	0.168**	0.209**	0.145**	-	-	-	-	-	-	-
DREGION1	-0.365**	-0.364**	-0.362**	-0.421**	-0.362**	-0.406**	-0.398	-0.398	-0.398	-0.399	-0.399
DREGION2	-0.296**	-0.294**	-0.296**	-0.327**	-0.297**	-0.329**	-0.324	-0.324	-0.323	-0.324	-0.324
DREGION3	-0.252**	-0.251**	-0.252**	-0.276**	-0.253**	-0.280**	-0.276	-0.276	-0.276	-0.276	-0.275
DREGION5	-0.039*	-0.041**	-0.043**	-0.070**	-0.038**	-0.083**	-0.085	-0.085	-0.087	-0.085	-0.084
DREGION6	-0.01	-0.015**	-0.014**	0.023**	-0.011*	0.009*	0.006	0.006	0.006	0.006	0.006
DREGION7	-0.005**	-0.014	-0.006	-0.032**	-0.008	-0.054**	-0.055	-0.055	-0.054	-0.055	-0.055
DSEA	0.113**	0.114**	0.116**	0.136**	0.112**	0.145**	0.148	0.148	0.148	0.148	0.147
DLX	0.349**	0.352**	0.344**	0.554**	0.348**	0.561**	0.567	0.567	0.566	0.567	0.566
DPORTO	0.331**	0.330**	0.327**	0.360**	0.344**	0.386**	0.386	0.386	0.384	0.386	0.387
DSCENIC2	0.100**	0.103**	-	0.125**	0.125**	-	-	-	-	-	-
DSCENIC3	0.266**	0.276**	-	0.321**	0.297**	-	-	-	-	-	-
DBADLOC	-0.171**	-0.175**	-0.165**	-	-0.174**	-	-	-	-	-	-
DEXCPLOC	0.316**	0.317**	0.325**	-	0.316**	-	-	-	-	-	-

*Note: A description of the variables is available in Appendix IV. * p-value < 0.05; ** p-value <0.0001.*

New apartments

	Omitted scenario/variables						All omitted (average over 1,000 replications)				
	Benchmark model	*Central heating and/or air conditioning*	*Visual quality*	*Location quality*	*construction quality*	*All omitted*	*n = 500*	*n = 1,000*	*n = 2,500*	*n = 5,000*	*n = 10,000*
Intercept	*10.182***	*10.188***	*10.181***	*10.198***	*10.173***	*10.187*	*10.164*	*10.162*	*10.163*	*10.163*	*10.163*
DENERGYAB	*0.123***	*0.128***	*0.123***	*0.135***	*0.127***	*0.153*	*0.162*	*0.162*	*0.163*	*0.163*	*0.163*
D2010	*0.002*	*0.003*	*0.002*	*0.003*	*0.000*	*0.001*	*-0.001*	*-0.001*	*0.000*	*0.000*	*0.000*
D2011	*-0.026***	*-0.025***	*-0.027***	*-0.022***	*-0.028***	*-0.023*	*-0.025*	*-0.024*	*-0.023*	*-0.023*	*-0.024*
D2012	*-0.088***	*-0.087***	*-0.088***	*-0.086***	*-0.091***	*-0.090*	*-0.089*	*-0.089*	*-0.089*	*-0.089*	*-0.090*
D2013	*-0.095***	*-0.092***	*-0.094***	*-0.091***	*-0.099***	*-0.089*	*-0.090*	*-0.089*	*-0.089*	*-0.089*	*-0.089*
SQRTGRFA	*0.131***	*0.131***	*0.131***	*0.127***	*0.132***	*0.130*	*0.131*	*0.131*	*0.131*	*0.131*	*0.131*
SQRTDEPFLOORA	*0.033***	*0.034***	*0.033***	*0.036***	*0.034***	*0.038*	*0.038*	*0.038*	*0.038*	*0.038*	*0.038*
DMROOMS	*0.039***	*0.041***	*0.043***	*0.054***	*0.038***	*0.069*	*0.056*	*0.056*	*0.057*	*0.058*	*0.058*
DBIGAPRT	*0.186***	*0.183***	*0.196***	*0.210***	*0.181***	*0.215*	*0.190*	*0.195*	*0.199*	*0.197*	*0.191*
DBGAPRTXCPL	*0.160***	*0.175**	*0.152**	*0.188**	*0.156**	*0.191*	*0.063*	*0.114*	*0.180*	*0.194*	*0.197*
DCSYSTEM	*0.070***	-	*0.073***	*0.086***	*0.086***	-	-	-	-	-	-
DCOND	*0.065***	*0.074***	*0.061***	*0.071***	*0.069***	*0.083*	*0.076*	*0.073*	*0.074*	*0.074*	*0.073*
DSWIMM	*0.169***	*0.168***	*0.177***	*0.182***	*0.168***	*0.197*	*0.206*	*0.207*	*0.207*	*0.207*	*0.207*
DCONSTP2	*-0.099***	*0.093***	*0.098***	*0.077***	*-0.102***	*-0.090*	*-0.097*	*-0.097*	*-0.097*	*-0.098*	*-0.098*
DCONSTP3	*-0.341***	*0.162***	*0.180***	*0.146***	*-0.347***	*-0.341*	*-0.334*	*-0.334*	*-0.333*	*-0.332*	*-0.335*
DCONSTP4	*-0.373***	*-0.102***	*-0.098***	*-0.085***	*-0.375***	*-0.351*	*-0.367*	*-0.377*	*-0.376*	*-0.378*	*-0.378*
DCONSTQ2	*0.081***	*-0.350***	*-0.339***	*-0.324***	-	-	-	-	-	-	-
DCONSTQ3	*0.137***	*-0.390***	*-0.373***	*-0.319***	-	-	-	-	-	-	-
DREGION1	*-0.366***	*-0.355***	*-0.362***	*-0.424***	*-0.367***	*-0.402*	*-0.394*	*-0.395*	*-0.395*	*-0.395*	*-0.394*
DREGION2	*-0.239***	*-0.229***	*-0.236***	*-0.268***	*-0.238***	*-0.243*	*-0.248*	*-0.247*	*-0.246*	*-0.246*	*-0.245*
DREGION3	*-0.256***	*-0.254***	*-0.255***	*-0.281***	*-0.260***	*-0.281*	*-0.277*	*-0.278*	*-0.278*	*-0.278*	*-0.278*
DREGION5	*-0.076***	*-0.084***	*-0.079***	*-0.094***	*-0.070***	*-0.106*	*-0.107*	*-0.110*	*-0.110*	*-0.109*	*-0.109*
DREGION6	*-0.013***	*-0.021***	*-0.014**	*-0.002*	*-0.017***	*-0.026*	*-0.038*	*-0.040*	*-0.038*	*-0.038*	*-0.037*
DREGION7	*-0.045***	*-0.056***	*-0.043***	*-0.079***	*-0.050***	*-0.104*	*-0.100*	*-0.100*	*-0.100*	*-0.098*	*-0.097*
DSEA	*0.081***	*0.079***	*0.083***	*0.111***	*0.080***	*0.117*	*0.124*	*0.124*	*0.123*	*0.123*	*0.123*
DLX	*0.265***	*0.274***	*0.258***	*0.504***	*0.267***	*0.529*	*0.554*	*0.556*	*0.556*	*0.556*	*0.557*
DPORTO	*0.291***	*0.282***	*0.288***	*0.301***	*0.321***	*0.353*	*0.382*	*0.380*	*0.378*	*0.378*	*0.377*
DSCENIC2	*0.063***	*0.068***	-	*0.091***	*0.113***	-	-	-	-	-	-
DSCENIC3	*0.142***	*0.148***	-	*0.172***	*0.206***	-	-	-	-	-	-
DBADLOC	*-0.184***	*-0.191***	*-0.181***	-	*-0.188***	-	-	-	-	-	-
DEXCPLOC	*0.324***	*0.330***	*0.332***	-	*0.321***	-	-	-	-	-	-

*Note: A description of the variables is available in Appendix IV. * p-value < 0.05; ** p-value <0.0001.*

Existing houses, omitted variable scenario results

		Omitted scenario/variables					All omitted (average over 1,000 replications)				
	Benckmark model	*Central heating and/or air conditioning*	*Visual quality*	*Location quality*	*construction quality*	*All omitted*	*n = 500*	*n = 1,000*	*n = 2,500*	*n = 5,000*	*n = 10,000*
Intercept	10.874**	10.877**	10.868**	10.804**	10.876**	10.804**	10.805	10.803	10.802	10.803	10.802
DENERGYAB	0.045**	0.051**	0.048**	0.055**	0.045**	0.070**	0.073	0.074	0.073	0.073	0.074
D2010	0.007	0.007	0.006	0.009	0.006	0.007	0.008	0.008	0.009	0.008	0.008
D2011	-0.067**	-0.067**	-0.067**	-0.070**	-0.067**	-0.068**	-0.068	-0.070	-0.069	-0.069	-0.068
D2012	-0.131**	-0.132**	-0.132**	-0.130**	-0.131**	-0.131**	-0.131	-0.131	-0.130	-0.131	-0.131
D2013	-0.158**	-0.158**	-0.159**	-0.152**	-0.158**	-0.154**	-0.152	-0.155	-0.154	-0.154	-0.154
SQRTGRFA	0.082**	0.082**	0.082**	0.086**	0.082**	0.088**	0.088	0.088	0.088	0.088	0.088
SQRTDEPFLOORA	0.012**	0.012**	0.012**	0.011**	0.012**	0.012**	0.012	0.012	0.012	0.012	0.012
SQRTPLOTAREA	0.007**	0.007**	0.007**	0.007**	0.007**	0.007**	0.007	0.007	0.007	0.007	0.007
DIRREGAREA	-0.129**	-0.128**	-0.128**	-0.135**	-0.128**	-0.133**	-0.128	-0.129	-0.130	-0.129	-0.128
DSMALLBEEDR	-0.083**	-0.083**	-0.082**	-0.075**	-0.083**	-0.072**	-0.071	-0.072	-0.071	-0.070	-0.070
SQRTDWELLTRANSA	-0.015**	-0.015**	-0.014**	-0.014**	-0.015**	-0.014**	-0.015	-0.014	-0.014	-0.014	-0.014
DCSYSTEM	0.095**	-	0.100**	0.095**	0.104**	-	-	-	-	-	-
DSWIMM	0.240**	0.246**	0.248**	0.289**	0.242**	0.319**	0.319	0.317	0.316	0.316	0.317
DPRIVPARK	0.089**	0.092**	0.089**	0.100**	0.090**	0.105**	0.103	0.105	0.105	0.105	0.105
DABSGAS	-0.089**	-0.089**	-0.090**	-0.118**	-0.090**	-0.125**	-0.125	-0.124	-0.125	-0.125	-0.125
DCONSTP2	-0.088**	0.041**	0.061**	0.027**	-0.091**	-0.096**	-0.096	-0.099	-0.100	-0.100	-0.101
DCONSTP3	-0.167**	0.166**	0.219**	0.095**	-0.170**	-0.151**	-0.152	-0.157	-0.158	-0.159	-0.160
DCONSTP4	-0.310**	-0.095**	-0.087**	-0.079**	-0.313**	-0.299**	-0.298	-0.300	-0.300	-0.300	-0.302
DCONSTQ2	0.024*	-0.176**	-0.164**	-0.135*	-	-	-	-	-	-	-
DCONSTQ3	0.127*	-0.318**	-0.308**	-0.284*	-	-	-	-	-	-	-
DREGION1	-0.369**	-0.366**	-0.373**	-0.459**	-0.367**	-0.453**	-0.443	-0.443	-0.445	-0.445	-0.445
DREGION2	-0.255**	-0.253**	-0.260**	-0.284**	-0.254**	-0.292**	-0.287	-0.285	-0.286	-0.286	-0.286
DREGION3	-0.331**	-0.330**	-0.336**	-0.411**	-0.330**	-0.415**	-0.413	-0.412	-0.414	-0.414	-0.413
DREGION5	-0.214**	-0.214**	-0.217**	-0.300**	-0.213**	-0.305**	-0.302	-0.299	-0.300	-0.301	-0.301
DREGION6	0.020	0.017	0.012	0.045**	0.022*	0.034*	0.034	0.037	0.038	0.038	0.039
DREGION7	-0.004	-0.008	-0.009	-0.077*	-0.003**	-0.093**	-0.094	-0.091	-0.088	-0.087	-0.085
DDISTRCAP	0.078**	0.083**	0.078**	0.151**	0.078**	0.158**	0.152	0.154	0.155	0.155	0.155
DSEA	0.159**	0.160**	0.163**	0.216**	0.159**	0.227**	0.229	0.231	0.230	0.230	0.230
DLX	0.299**	0.294**	0.287**	0.490**	0.300**	0.489**	0.495	0.497	0.495	0.494	0.494
DPORTO	0.278**	0.274**	0.285**	0.282**	0.279**	0.298**	0.301	0.300	0.296	0.299	0.298
DSCENIC2	0.145**	0.146**	-	0.222**	0.154**	-	-	-	-	-	-
DSCENIC3	0.248**	0.257**	-	0.318**	0.277**	-	-	-	-	-	-
DBADLOC	-0.154**	-0.155**	-0.152**	-	-0.154**	-	-	-	-	-	-
DEXCPLOC	0.475**	0.475**	0.500**	-	0.475**	-	-	-	-	-	-

*Note: A description of the variables is available in Appendix IV. * p-value < 0.05; ** p-value <0.0001.*

New houses, omitted variable scenario results

		Omitted scenario/variables					All omitted (average over 1,000 replications)				
	Benchmark model	*Central heating and/or air conditioning*	*Visual quality*	*Location quality*	*construction quality*	*All omitted*	*n = 500*	*n = 1,000*	*n = 2,500*	*n = 5,000*	*n = 10,000*
Intercept	10.899**	10.886**	10.895**	10.848**	10.896**	10.817**	10.808	10.812	10.808	10.807	10.807
DENERGYAB	0.055**	0.061**	0.056**	0.066**	0.055**	0.073**	0.087	0.085	0.084	0.084	0.084
D2010	0.008	0.008	0.009	0.005	0.009	0.005	0.006	0.007	0.006	0.006	0.006
D2011	-0.034**	-0.034**	-0.032*	-0.042**	-0.034**	-0.040**	-0.041	-0.040	-0.041	-0.040	-0.041
D2012	-0.099**	-0.099**	-0.098**	-0.097**	-0.099**	-0.096**	-0.095	-0.096	-0.096	-0.096	-0.096
D2013	-0.127**	-0.128**	-0.125**	-0.126**	-0.126**	-0.122**	-0.124	-0.122	-0.123	-0.123	-0.124
SQRTGRFA	0.086**	0.088**	0.087**	0.089**	0.086**	0.093**	0.093	0.093	0.093	0.093	0.093
SQRTDEPFLOORA	0.012**	0.012**	0.012**	0.012**	0.012**	0.013**	0.012	0.012	0.012	0.012	0.012
SQRTPLOTAREA	0.006**	0.006**	0.006**	0.006**	0.006**	0.005**	0.005	0.005	0.005	0.005	0.005
DRENOV	-0.164**	-0.166**	-0.165**	-0.168**	-0.165**	-0.174**	-0.174	-0.174	-0.175	-0.175	-0.175
DGRFLOORENOV	-0.084**	-0.080**	-0.083**	-0.075**	-0.083**	-0.063**	-0.080	-0.080	-0.077	-0.076	-0.076
DIRREGAREA	-0.133**	-0.131**	-0.132**	-0.137**	-0.133**	-0.133**	-0.138	-0.136	-0.133	-0.132	-0.132
DBADCONSERVATION	-0.147**	-0.149**	-0.145**	-0.152**	-0.147**	-0.152**	-0.142	-0.140	-0.141	-0.140	-0.142
DCSYSTEM	0.071**	-	0.073**	0.075**	0.078**	-	-	-	-	-	-
DSWIMM	0.271**	0.274**	0.273**	0.300**	0.272**	0.314**	0.320	0.321	0.321	0.321	0.321
DPARKING	0.039**	0.043**	0.037**	0.044**	0.040**	0.047**	0.047	0.048	0.049	0.049	0.049
DABSGAS	-0.067**	-0.068**	-0.067**	-0.095**	-0.069**	-0.100**	-0.096	-0.097	-0.097	-0.097	-0.097
DCOND	0.071**	0.076**	0.081**	0.054*	0.069**	0.073**	0.076	0.076	0.077	0.077	0.077
DCONSTP2	-0.055**	-0.058**	-0.054**	-0.055**	-0.056**	-0.057**	-0.054	-0.054	-0.053	-0.054	-0.054
DCONSTP3	-0.161**	-0.166**	-0.160**	-0.146**	-0.162**	-0.153**	-0.135	-0.134	-0.136	-0.135	-0.135
DCONSTP4	-0.309**	-0.312**	-0.308**	-0.299**	-0.310**	-0.305**	-0.283	-0.285	-0.287	-0.287	-0.289
DCONSTQ2	0.021*	0.034*	0.042**	0.015	-	-	-	-	-	-	-
DCONSTQ3	0.109**	0.141**	0.149**	0.075*	-	-	-	-	-	-	-
DREGION1	-0.400**	-0.395**	-0.404**	-0.482**	-0.394**	-0.468**	-0.461	-0.462	-0.461	-0.462	-0.461
DREGION2	-0.233**	-0.229**	-0.237**	-0.255**	-0.231**	-0.253**	-0.261	-0.261	-0.262	-0.262	-0.261
DREGION3	-0.325**	-0.326**	-0.329**	-0.393**	-0.323**	-0.395**	-0.406	-0.405	-0.404	-0.405	-0.405
DREGION5	-0.211**	-0.214**	-0.212**	-0.281**	-0.210**	-0.284**	-0.292	-0.293	-0.292	-0.292	-0.291
DREGION6	0.042*	0.036*	0.037*	0.061**	0.044*	0.052**	0.047	0.047	0.048	0.047	0.047
DREGION7	-0.057*	-0.065*	-0.061*	-0.105**	-0.056*	-0.119**	-0.120	-0.115	-0.114	-0.115	-0.113
DDISTRCAP	0.117**	0.125**	0.119**	0.176**	0.116**	0.189**	0.183	0.179	0.181	0.181	0.180
DSEA	0.118**	0.119**	0.120**	0.164**	0.118**	0.172**	0.181	0.180	0.181	0.181	0.181
DLX	0.364**	0.370**	0.341**	0.597**	0.365**	0.590**	0.447	0.530	0.553	0.550	0.550
DPORTO	0.370**	0.365**	0.376**	0.406**	0.382**	0.443**	0.446	0.451	0.445	0.443	0.443
DSCENIC2	0.098**	0.100**	-	0.122**	0.112**	-	-	-	-	-	-
DSCENIC3	0.131	0.136*	-	0.193**	0.175**	-	-	-	-	-	-
DBADLOC	-0.148**	-0.148**	-0.146**	-	-0.146**	-	-	-	-	-	-
DEXCPLOC	0.358**	0.359**	0.375**	-	0.355**	-	-	-	-	-	-

*Note: A description of the variables is available in Appendix IV. * p-value < 0.05; ** p-value <0.0001.*

UNIVERSIDADE DE ÉVORA

UNIVERSIDADE DE ÉVORA
INSTITUTO DE INVESTIGAÇÃO
E FORMAÇÃO AVANÇADA

Contactos:
Universidade de Évora
Instituto de Investigação e Formação Avançada - IIFA
Palácio do Vimioso | Largo Marquês de Marialva, Apart. 94
7002-554 Évora | Portugal
Tel: (+351) 266 706 581
Fax: (+351) 266 744 677
email: iifa@uevora.pt