REM WORKING PAPER SERIES

On the use of Hedonic Regression Models to Measure the Effect of Energy Efficiency on Residential Property Transaction Prices: Evidence for Portugal and Selected Data Issues

Rui Evangelista, Esmeralda A. Ramalho, João Andrade e Silva

REM Working Paper 064-2019

January 2019

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

On the use of Hedonic Regression Models to Measure the Effect of Energy Efficiency on Residential Property Transaction Prices:

Evidence for Portugal and Selected Data Issues[◊]

RUI EVANGELISTA[*], ESMERALDA A. RAMALHO[+] and JOÃO ANDRADE E SILVA[Φ]

[*] Instituto Nacional de Estatística, rui.evangelista@ine.pt

[+] Universidade de Lisboa, Instituto Superior de Economia e Gestão & REM - CEMAPRE, eramalho@iseg.ulisboa.pt

[Φ] Universidade de Lisboa, Instituto Superior de Economia e Gestão & REM - CEMAPRE, joaoas@iseg.ulisboa.pt

Abstract

Using a unique dataset containing information of around 256 thousand residential property sales, this paper discloses a clear sales premium for most energy efficient dwellings, which is more pronounced for apartments (13%) than for houses (5 to 6%). Cross-country comparisons support the finding that energy efficiency price premiums are higher in the Portuguese residential market than in central and northern European markets. Results emphasize the relevance of data issues in hedonic regression models. They illustrate how the use of appraisal prices, explanatory variables with measurement errors, and the omission of variables associated with the quality of the properties, may seriously bias energy efficiency partial effect estimates. These findings provide valuable information not only to policy-makers, but also to researchers interested in this area.

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

JEL Classification: C51, C55, C81, Q41, R21, R30

[◊] The authors would like to thank the *Instituto Nacional de Estatística* (INE) and *Agência para a Energia (ADENE)*, who have made possible the access to the data used in this paper at the premises of *INE*. This research was supported in part by the CEMAPRE **UID/MULTI/00491/2019** project from the Foundation for Science and Technology, FCT/MCTES, through national funds.

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). Given the relevance of the residential sector in the total building stock, the implementation of policies aimed at increasing its energy performance, such as the introduction of energy efficiency labelling schemes, 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 (ADENE, 2015: 10, 19).

Energy labelling has been applied in Europe for many years, with household appliances being its earliest widespread area of application. 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 which, essentially, translates the energy performance of property units into an energy efficiency scale. In the particular case of Portugal, analysed in this paper, EPCs started to be issued for all new residential buildings with more than 1,000 square meters from July 2007 onwards, and are mandatory in all residential transactions since the beginning of 2009 and in all properties listed for sale or rent since December 2013.

One of the main objectives of this paper is to assess whether energy efficiency, as measured by the EPC rating system, has an impact on the transaction price of residential properties in Portugal. The relationship between energy efficiency and residential house prices has been typically defined in the framework of hedonic price models (Rosen, 1974), which require the availability of information on dwelling transaction prices as well as on energy efficiency attributes and other price determinants such as their size, age, and location quality. Recent applications of this methodology are Hyland et al. (2013), Fuerst et al. (2015), Ayala et al. (2016), Fesselmeyer (2018) and Fuerst and Warren-Myers (2018), who provide results for Ireland, England, Spain, Singapore and Australia, respectively. Although the majority of the papers point to the conclusion that increased energy efficiency entails a market price premium at the time of sale (e.g., Fuerst et al., 2016), 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 energy efficiency technology, price premiums can be reduced or even take the form of price discounts (Yoshida and Sugiura, 2015).

The present paper exploits a dataset that includes more than 256 thousand property transactions for Portugal, for the 2009-2013 period, for which information from the EPC information system, together with prices and dwelling characteristics taken from transfer and property taxes purposes, were gathered. The tax sources cover all dwelling transactions carried out in Portugal, since it is not possible to make a property transaction without proof of payment of transfer tax. Two key contributions are offered to the literature. First, it constitutes the first large-scale study for a southern European country in which the effect of EPC labels on residential transaction prices is assessed. Although some evidence is given for Spain in Ayala et al. (2016), their results are based on a small sample and on the owner appraisal property valuation instead of effective transaction prices. This paper also provides the results of a cross-country comparability study in which it is possible to investigate the issue of whether energy efficiency is valued in Portugal as it is in central and northern Europe countries. Second, the availability of an unusually large and rich dataset allows for the clarification of important data issues associated with the use of the hedonic price model, which has been employed as the workhouse in this area of research. In particular, given their practical importance in the estimation of energy efficiency partial effects, three data issues were considered: (i) the influence of large samples on the potential inflation of significance levels of relevant parameters, (ii) the impact of using error-prone measures of either the price (e.g., appraisals or list prices) or property characteristics (e.g., size variables), and (iii) the importance of the omission of new or rarely used potential price determinants in the model (e.g., visual prominence of the location).

This paper is organized as follows. Section 2 briefly reviews the literature on the impact of energy efficiency on residential property prices, focussing on data issues related to the measurement of crucial variables in the specification of the hedonic models. Section 3 describes the sources, variables and information available for research and presents some data descriptive statistics. Section 4 provides an estimation of the effect of energy efficiency in property prices in Portugal, including a robustness check and a cross-country comparison exercise. Section 5 investigates the effect of some measurement issues often present in hedonic models designed to capture the effect of the EPC rating system. Finally, the last section provides a summary of the main results.

2. Energy efficiency, hedonic models and housing prices

The link between energy efficiency and housing prices is associated with the idea that markets are able to internalize the benefits of lower energy consumption patterns. For instance, as more energy efficient properties experience lower future utility bills, it is expected that they should display a market transaction premium when compared to less efficient properties; Dinan and Miranowski (1989). In reality, however, the relationship between energy efficiency and prices is far from straightforward. For example, in markets where energy efficiency standards are perceived as high, extra efficiency gains can be regarded as an imposition of additional technological maintenance costs and translated into market price discounts; Yoshida and Sugiura (2015). Therefore, the relationship between property prices and energy efficiency may well be insignificant or take up the form of a market price premium or discount.

2.1. Measuring the impact of energy efficiency through the use of hedonic price model

The hedonic price model has been widely used to measure the effect of property characteristics on their price. Nelson (1982), who summarises eight studies estimating the relationship between traffic noise and property values, is one early example. Chin and Chau (2003) and Malpezzi (2003) are two excellent reviews of the application of the hedonic price model to housing. For the specific case of the effect of energy efficiency on the value of housing, recent illustrations are given by Fesselmeyer (2018) and Fuerst and Warren-Myers (2018).

Central to the hedonic price model is the idea of the existence of a functional relationship between prices and attributes, which can be expressed as:

$$p^* = \beta_0 + \beta_1 . E + \sum_{k=2}^{K} \beta_k . x_k^* + u. \qquad (1)$$

In (1), p^* and E represent the price and the energy efficiency of the dwelling, respectively. Moreover, the x^* corresponds to the remaining housing attributes, * signals the fact that the attribute may be transformed and u is a term representing additional random factors, which are not measured by the k+1 variables included in the model. 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). In contrast, less exploited explanatory factors, which are available in our dataset, are the scenic quality of the location, building technology, and quality of construction works.

The main focus in this paper is to assess the statistical significance, sign and magnitude of the partial effect of E, as measured by the EPC rating system, on the transaction price of dwellings. As noted elsewhere (see, *inter alia*, Cropper el al., 1988), theory sheds little light on the selection of the appropriate functional form of (1) and the derivation of the hedonic function is essentially seen as an empirical matter. In the literature dealing with the impact of energy efficiency on prices, p^* typically assumes the form of a logarithmic transformation of p and the explanatory variable of interest E results from the transposition of a discrete measurement scale into one or more binary variables. This paper follows this approach. In this situation, the relative effect of E can be measured by $r = exp(\beta_1) - 1$, an estimator proposed by Halvorsen and Palmquist (1980).[2] In some studies, however, this effect is simply grasped by $\hat{\beta}_1$. Although this provides a reasonable approximation of r when E is included in the hedonic model as a continuous variable (see Megerdichian, 2018), it is not adequate to deal with the nine-level energy efficiency rating scale adopted in Portugal, which ranges from A^+, the most efficient level, to G, the lest efficient level.

Another issue related with the definition of the scale of measurement of the EPC label emerges from the fact that, despite the existence of a common energy performance framework in Europe, the methodology underlying the implementation of the EPC labelling scheme in each country is tailored to national contexts, which prevents direct comparisons of the magnitude of different energy efficiency estimates; for an overview of the different EPC schemes within the EU, see Atanasiu and Constantinescu (2011). However, by introducing some changes in the hedonic price models, it is possible to increase the degree of comparability between studies and present a qualitative cross-country assessment of the impact of energy efficiency on dwelling prices.

2.2. Data issues

While the use of the hedonic price model to estimate the relationship between market transaction prices and energy efficiency is well established in the literature, there are important empirical issues that remain seldom, if ever, assessed. This situation has, at least partially, to do with the fact that researchers are often confronted with the data they have at hand and are not able to conduct experiments involving different data contexts. Taking advantage of the quality of our database, we will discuss some of these questions.

---

[2] For a more detailed discussion of estimators of r see Kennedy (1981) and van Garderen and Shah (2002).

The first issue addressed in this paper are the consequences of using exceptionally large datasets. Since parameters' standard errors decrease as the size of the sample increases, significance levels of energy efficiency and other key variables may be inflated to a point in which standard *t* and other statistical tests become artificially relevant; see Ziliak and McCloskey (2004). Apart from some notable exceptions (e.g., Lin et al., 2013), this topic has not deserved much attention. However, this is an important matter since with the dissemination of energy labels, it is expected that the problems stemming from the use of large datasets in this area become more relevant; see for example Fuerst et al. (2015), who base their conclusions on a sample of more than 330 thousand observations.

A second data issue has to do with the sensitivity of energy efficiency partial effect estimates to the replacement of key variables in the hedonic specification by variables that necessarily display some sort of measurement error. A clear example is the use of surrogate, instead of actual, transaction prices in energy efficiency partial effect estimates. Ayala et al. (2016) and Hyland et al. (2013) are two examples where appraisals and list prices are applied as a proxy of market prices. This is usually a reasonable approach, as often the measurement error in the dependent variable is innocuous in ordinary least squares (OLS) coefficient estimation; see Wooldridge (2013: 318-20). However, if the differences between proxy and true transaction prices are correlated with some of the covariates included in the model (e.g., dwelling dimension), the OLS estimator of energy efficient partial effects is biased and inconsistent. In fact, the use of an inappropriate estimator may help explaining the existence in the literature of different energy efficiency price impact results. As the database used in the present paper includes both fiscal appraisal values and transactions prices, it is possible to highlight differences in coefficient estimates stemming from the use of proxy and actual price measures in the hedonic regression model. On the other hand, the consequences of having a property attribute measured with error, which are expected to cause the inconsistency of the coefficient estimators, as measurement error of the explanatory variables has in general severe destructive effects, are also illustrated.

A final important data issue revolves around the sensitivity of energy efficiency partial effects to the omission of variables that measure the quality of transacted properties. While the omission of relevant covariates appears as the elephant in the room problem in hedonic regression applications, the literature in this area is not particularly prolific in tackling this topic. Stanley et al. (2016) highlight the importance of including controls for the age of the dwelling. Further examples include the omission of locational attributes and quality of the dwelling, noticed by Fuest et al. (2015) and the non-inclusion of hard-to-measure factors, such as buyer's predisposition to environmental issues (Brounen and Kok, 2011) and developer's reputation (Zheng et al., 2012). Despite the fact that variable omission only leads to the inconsistency of the parameter estimators in cases where the omitted variable is correlated with the included covariates (see Wooldridge, 2013: 172), housing attributes, which are hard to measure and often not available in datasets, such and the visual prominence of the location, may be correlated with both the EPC label and other covariates included in the hedonic model. By exploring the dataset available for regression analysis, this paper assesses the sensitivity of energy efficiency partial effect estimates to the omission of key and hard to measure variables.

3. Data and descriptive analysis

The dataset exploited in this paper combines energy efficiency data taken from the national supervision body responsible for the implementation and administration of the European EPC system in Portugal (ADENE), and transaction prices and dwelling attributes, obtained from the Portuguese Tax and Customs Authority. Transaction prices were obtained from real estate transfer tax (IMT) records, and property attributes were taken from local property tax (IMI) data.[3] The IMT source, which is available from 2009 onwards, covers the population of residential property sales since it is not possible to carry out a transaction without a proof of payment of this tax. Likewise, energy efficiency data covers all transactions from 2009, as EPC issuing became mandatory for all residential property transactions since the beginning of that year.

The matching of the IMT, IMI and ADENE data sources was carried out using the property cadastral register identification number, a unique identification code attributed to each property unit, as the linking key variable. The end-product of this merging process was an extremely rich dataset containing information on the prices, energy performance and other dwelling characteristics of 256,145 residential property transactions carried out from 2009 to 2013 in Portugal. This constitutes one of the largest datasets ever used to estimate the impact of energy efficiency on transaction prices. 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 latter data was used in Ramalho et al. (2017) in an empirical exercise involving the construction of hedonic price indexes with aggregate information, for which however the EPC rating was not available as a price determinant. Table 1 provides the summary statistics for a group of selected variables of the dataset. In

[3] The real estate transfer tax is designated as *Imposto Municipal sobre a Transmissão Onerosa de Imóveis* or simply as IMT. The local property tax is designated as *Imposto Municipal sobre Imóveis* (IMI).

addition to the totals, descriptive statistics for four market segments, existing apartments, new apartments, existing houses and new houses, are presented.

Table 1: Summary statistics of a group of selected variables

Variable description	*All (N=256,145)*		*Existing apartments (N=149,920)*		*New apartments (N=59,410)*		*Existing houses (N=33,282)*		*New houses (N=13,533)*	
	Mean	*Stdev*	*Mean*	*Stdev*	*Mean*	*Stdev*	*Mean*	*Stdev*	*Mean*	*Stdev*
Transaction value (€)	*119,888*	*98,131*	*97,695*	*68,876*	*149,007*	*93,688*	*143,783*	*150,717*	*179,155*	*145,671*
Fiscal Appraisal value (€)	*89,634*	*61,116*	*76,752*	*46,097*	*111,038*	*57,445*	*95,137*	*85,986*	*124,852*	*96,557*
Energy label A(.)	*.049*	*.215*	*.016*	*.125*	*.139*	*.346*	*.026*	*.160*	*.070*	*.255*
Energy label B	*.202*	*.402*	*.159*	*.366*	*.393*	*.489*	*.073*	*.259*	*.157*	*.364*
Energy label B^-	*.128*	*.334*	*.120*	*.325*	*.157*	*.364*	*.110*	*.313*	*.142*	*.349*
Energy label C	*.363*	*.481*	*.468*	*.499*	*.228*	*.419*	*.185*	*.388*	*.222*	*.415*
Energy label D	*.148*	*.355*	*.132*	*.338*	*.062*	*.241*	*.330*	*.470*	*.248*	*.432*
Energy labels E, F and G	*.111*	*.314*	*.105*	*.307*	*.021*	*.142*	*.277*	*.447*	*.161*	*.368*
Gross floor area (m2)	*110.6*	*50.7*	*96.0*	*34.0*	*113.1*	*38.0*	*148.3*	*79.7*	*168.6*	*67.4*
Dependent floor area (m2)	*31.1*	*39.1*	*18.6*	*21.1*	*36.1*	*26.3*	*60.2*	*66.7*	*75.4*	*64.4*
Uncovered land area (m2)	*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 (years)	*16.1*	*18.9*	*20.1*	*17.5*	*2.0*	*2.1*	*29.3*	*25.9*	*1.5*	*2.1*
Improv. or renewed property	*.054*	*.226*	*.015*	*.120*	*.089*	*.285*	*.45*	*.207*	*.359*	*.480*

Note: (.) A^+ and A ratings.

The data reveals clear differences amongst the different dwelling types. As expected, new is generally more expensive than existing and houses are more expensive than apartments. For instance, while existing apartments' price present an average value of nearly 97,000 €, new houses provide a much higher average of approximately 180,000 €. Existing apartments stand out as the most common property to be sold in the dataset (59% of all observations), and new houses as the less frequently purchased property type (5% of all transactions). Naturally, price dispersion is much higher for houses than for apartments, suggesting that the former property type is more heterogeneous than the latter and that hedonic model specification may be more complex for houses than for apartments. The data also shows that appraisals carried out for fiscal purposes are generally defined at a lower level than transaction prices, with the average value for the former value being 21 to 34 percent lower than the latter value.

The most common energy 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, A and B properties were grouped together into a single dummy for modelling purposes; i.e., the variable E in (1) assumes a binary form, where the value 1 is attributed to all A and B rated properties and the value 0 in other cases; for alternative forms of aggregation see Section 4.2.

The differences across market segments are also explicit in other variables. In terms of dimension, gross floor area is higher for houses than for apartments and, interestingly, there is a clear difference between new and existing dwellings, with the latter property type being smaller than the former. Moreover, the summary statistics for age reveals that existing houses display an average age of 29 years and the set of transacted new houses show an average of 2 years. 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). The analysis of the data has also revealed a relatively high percentage of new houses that were completed in or before 1991. This is essentially explained by the existence of major improvements and renovations in this dwelling category, which account for 35.9 percent of the total of new house transactions (see the last line of Table 1). In these cases, the once before old property is put on the market as a new property and it is considered as such in the database. Given the smaller variability of the age variable for new dwellings, it may be argued that it may not be that important to explain the formation of price than it is for existing dwellings. However, since renovated dwellings may display vintage effects, it may be interesting to include information on the year in which they were completed in the hedonic models.

4. Hedonic models for property prices in Portugal

The key question investigated in this paper, in the framework of the hedonic price model, is of whether or not energy efficiency has an impact on transaction prices. This research question can be formulated as a hypothesis stated as follows:

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

It should be noted that this is underpinned by an inequality assumption, were the null was defined as $H_0\colon \beta_1 \leq 0$, to reflect the idea that energy efficiency has either no or negative impact on prices. 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.

The modelling approach builds on the market segments identified in the descriptive analysis of section 3. 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. On the other hand, special care was taken as to the inclusion in the models of all possible price-determining variables suitable to capture the impact of energy efficiency on residential property prices; see the comprehensive list housing attributes used in hedonic price models in Chin and Chau (2003). The complete list of all the explanatory variables used in our regressions is available in Appendix 1.

4.1. Regression results

Table 2 presents the results of the OLS partial effects and of the one-tailed tests used to investigate the validity of the hypothesis under study. The Ramsey's (1969) Reset specification test results were obtained using a robust to the presence of heteroskedasticity procedure, which is based on Lagrange-Multiplier (LM) statistics, developed by Wooldridge (1991). The complete OLS regression results are available in Appendix 2.

Table 2: 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 ($\beta_1 \leq 0$)*, p-value*	*<.0001*	*<.0001*	*<.0001*	*<.0001*
Number of obs. used in estimation	*149,920*	*59,410*	*33,282*	*13,533*
Regressions' adjusted R^2	*.676*	*.733*	*.670*	*.753*
Reset, p-value	*.567*	*.098*	*.083*	*.052*

As the bottom line of Table 2 show, the functional form adopted for the four models is not rejected by the specification test at the 5 percent level of significance. More importantly, the null of the one-tailed test is rejected for all segments, which suggests that the impact of energy efficiency on residential property prices is positive.

Interestingly, 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; see Fuerst et al. (2015) for another example where price premiums differ across dwelling types. A 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.

4.2. Robustness analysis and cross-country comparisons

To increase the cross-country comparability of our results and assess their robustness, the hedonic regression models of the previous section were re-estimated with different energy efficiency measurement scales. Specifically, they replicate for Portugal the energy efficiency scales used in the regression studies for three northern European countries: Ireland (Hyland et al., 2013), Finland (Fuerst et al., 2016) and the Netherlands (Brounen and Kok, 2011). In the last case, the re-estimation of the hedonic regression model also took into consideration the fact that the model's dependent variable used in Brounen and Kok's (2011) paper was the natural logarithm of price per square meter. Table 3 provides the results.

Table 3: Energy efficiency partial effects under different EPC measurement scales

	Hyland et al. (2013)(+)	*Sub-market*			
		Exist. Apart.	***New Apart.***	***Exist. Houses***	***New Houses***
A	*9.7%**	*23.6%**	*20.7%**	*3.0%*	*10.0%**
B	*5.3%**	*13.5%**	*10.0%**	*5.2%**	*4.8%**
C	*1.7%**	*1.5%**	*-1.4%**	*3.3%**	*1.8%**
E	*-0.4%*	*0.9%**	*-4.9%**	*-0.8%*	*-3.0%**
F/G	*-10.1%**	*-1.8%*	*-0.4%*	*-2.9%**	*-4.3%**
n	*15,060*	*149,920*	*59,410*	*33,282*	*13,533*
R^2 (.)	*.65*	*.68*	*.74*	*.67*	*.75*
Reset, p-value	*-*	*.800*	*.219*	*.095*	*.043*
	Fuerst *et al.*(2016) (+)				
ABC	*1.3%**	*5.5%**	*7.0%**	*4.1%**	*4.0%**
E	*0.0%*	*1.0%*	*-4.7%**	*-0.8%*	*-3.0%**
FG	*0.0%*	*-1.9%*	*-0.5%**	*-2.9%**	*-4.5%**
n	*6,194*	*149,920*	*59,410*	*33,282*	*13,533*
R^2 (.)	*.933*	*.67*	*.72*	*.67*	*.75*
Reset, p-value	*-*	*.030*	*.066*	*.092*	*.041*
	Brounen and Kok (2011)(-)				
ABC	*3.7%**	*4.5%**	*6.0%**	*4.7%**	*5.6%**
n	*31,993*	*149,920*	*59,410*	*33,282*	*13,533*
R^2 (.)	*.527*	*.550*	*.610*	*.478*	*.575*
Reset, p-value	*-*	*.000*	*.000*	*.000*	*.000*

*Notes: * p-value < 0.05. (+) Omitted class: D. (-) Omitted classes: D,E,F, and G. The dependent variable of the model is the natural logarithm of price per square meter. (.) 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.*

Energy efficiency partial effects are coherent across the different measurement scales and display a plausible pattern: e.g., are higher for A than B- rated properties and display price discounts for lower ratings. In addition, the results clearly identify a higher price premium for apartments. With the single exception of the minor difference between the price premiums associated with A, B and C-rated existing apartments and houses (4.5% against a 4.7% premium, respectively), this finding is invariant to the change of the energy efficiency scale and is stable across all regressions.

Energy efficiency is essentially rewarded by properties exhibiting most energy efficient (i.e., A and B) ratings. This is illustrated in the top figure of Table 3, where C-rated dwellings even display a price discount (-1.4%), and A and B properties show price premiums lying between 23.6 and 3.0 percent, respectively. Moreover, the existence of a higher energy efficiency price premium for new properties is not crystal-clear. For instance, while the price premium of A, B and C-rated apartments increases from 5.5 to 7.0 percent from existing to new properties, it basically remains stable (from 4.1 to 4.0 percent) when one moves from existing to new houses. Moreover, the overall fit of the regressions is in line with those found in similar studies. While the regressions using the logarithm of the price level as the dependent variable are not rejected by the Reset test at the 5 or 1 percent significance levels, the specification with the logarithm of the price per square meter is rejected by this specification test, a fact that reinforces the idea that the choice of the dependent variable was suitable for the Portuguese market.

Interestingly, the results suggest that energy efficiency is rewarded with higher price premiums in Portugal than in Ireland, Finland, and the Netherlands. This is particularly evident for top-rated energy efficiency properties, with A and B- ratings receiving price premiums that are always higher than those reported in the considered country studies. This is in line with Ramos et al. (2015), who reported a similar conclusion based on a much smaller sample of residential properties listed for sale. This outcome for Portugal could result from a greater awareness of the EPC label and/or the existence of higher energy costs.

5. Selected data issues

This section addresses some data issues often encountered in the measurement of energy efficiency by hedonic models. The first subsection considers smaller subsamples of the large scale dataset employed in this paper to check whether the individual statistical significance of the EPC dummy is inflated by the sample size. The remaining two sections assess the sensitivity of energy efficiency partial effects to the introduction of measurement errors in the dependent and explanatory variables and the omission of relevant information in the hedonic model outputs.

5.1. Large samples

The impact of using different sample sizes on the quality of regression results was investigated in an experiment in which the hedonic regression models were rerun for a number of samples with different sizes. In particular, 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 drawn for existing apartments, new apartments, existing houses and new houses.

The averages of energy efficiency parameter estimates over the 1,000 replications for each sample size are depicted in Table 4. Standard errors are provided underneath each average value in brackets as well as the counts of statistically significant positive coefficients, given in squared brackets. For reference, the parameter estimates of a benchmark model, which is the model proposed in section 4.1, are also provided.

Table 4: Energy efficiency partial effects, different experiments

	Sub-market			
	Existing Apartments	***New Apartments***	***Existing Houses***	***New houses***
Benchmark estimate				
	*0.118***	*0.123***	*0.045***	*0.055***
	(.0019)	*(.0024)*	*(.0061)*	*(.0062)*
Parameter results, averages over 1,000 replications(+)				
n = 500	*0.121*	*0.134*	*0.047*	*0.067*
	(.034)	*(.028)*	*(.052)*	*(.033)*
	[947]	*[996]*	*[169] (*)*	*[551]*
n = 1,000	*0.119*	*0.134*	*0.049*	*0.066*
	(.025)	*(.019)*	*(.035)*	*(.024)*
	[996]	*[1,000]*	*[272]*	*[814]*
n = 2,500	*0.120*	*0.134*	*0.048*	*0.065*
	(.015)	*(.012)*	*(.023)*	*(.015)*
	[1,000]	*[1,000]*	*[572]*	*[993]*
n = 5,000	*0.120*	*0.134*	*0.048*	*0.065*
	(.011)	*(.009)*	*(.016)*	*(.010)*
	[1,000]	*[1,000]*	*[849]*	*[1,000]*
n = 10,000	*0.120*	*0.135*	*0.048*	*0.065*
	(.008)	*(.006)*	*(.011)*	*(.007)*
	[1,000]	*[1,000]*	*[982]*	*[1,000]*
Parameter estimates, measurement error scenario (.)				
Fiscal appraisal as dependent variable	*0.0386***	*0.0485***	*-0.0085**	*-0.0010*
	(.00108)	*(.00163)*	*(.00360)*	*(.00418)*
No. of bedrooms as size measurement	*0.1481***	*0.1479***	*0.0940***	*0.0842***
	(.00215)	*(.00281)*	*(.00702)*	*(.00706)*
Parameter estimates, omitted variable scenario (.)				
Central heating and/or air cond.	*0.125***	*0.128***	*0.051***	*0.061***
	(.00190)	*(.00244)*	*(.00608)*	*(.00624)*
Visual quality	*0.118***	*0.123***	*0.048***	*0.056***
	(.00193)	*(.00245)*	*(.00616)*	*(.00625)*
Location quality	*0.131***	*0.135***	*0.055***	*0.066***
	(.00201)	*(.00255)*	*(.0064)*	*(.00641)*
construction quality	*0.120***	*0.127***	*0.045***	*0.055***
	(.00192)	*(.00245)*	*(.00613)*	*(.00624)*
All omitted	*0.145***	*0.153***	*0.070***	*0.073***
	(.00201)	*(.00259)*	*(.00642)*	*(.00645)*

Notes: (+) The point estimates refer to the averages over the 1,000 simulations; standard deviation provided in parenthesis. The number of statistically significant positive coefficients is shown between square brackets. (*) This experiment yielded 4 statistically significant negative coefficient estimates.
*(.) Robust standard errors in parenthesis. * p-value < 0.05, ** p-value <0.0001.*

With the exception of existing houses for the three smallest sample sizes, the number of statistically significant energy efficiency coefficients is substantial, even in cases where the sample size is strongly reduced, that display an important increase in the variability of the coefficients (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 obtained for samples with 500 observations (0.033)). This provides evidence that in the proposed hedonic models of section 4.1 the statistical relevance of this characteristic is not inflated by the sample size. The results also reveal that the cases of statistically significant coefficients with conflicting signs are extremely rare, occurring only four times for existing houses and for the smallest sample size (n=500). In addition, it is also possible to see that the 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, which represent less than 0.4 percent of the total transactions of this market segment.

5.2. Measurement errors

This section investigates the degree to which energy efficiency partial effects, which were estimated by the benchmark models presented in section 4.1, are sensitive to the introduction of dependent and explanatory variables with measurement errors. To illustrate the former case, the logarithm of transaction prices is replaced by the logarithm of fiscal appraisal values, which is available in the dataset. The choice of fiscal appraisals allows 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 1. The illustration of the impact of using erroneously measured explanatory variable on energy efficiency partial effects is provided by the re-estimation of the hedonic models with the number of bedrooms replacing the area variables. The former variable is a poor measurement of the dimension of properties, which may have to be used in cases where researchers do not have access to better size measures. The results are displayed in the middle of Table 4.

The damaging effects of using either the logarithm of fiscal appraisals as a dependent variable or the number of bedrooms as a measurement for size are evident. In the former case, 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 exhibit a small price discount (-0.0085), or for new houses, where the parameter is statistically insignificant. These results also emphasize the importance of energy efficiency as a factor in explaining residential property market prices, as the fiscal evaluation of properties does not explicitly takes into account energy efficiency parameters (DGI, 2011). As such, it is expected that the size and significance of energy efficiency partial effects diminish or even becomes statistically irrelevant when this variable is considered as the model’s dependent variable.

When area variables are replaced by the number of bedrooms, the distortion in energy efficient partial effects is even more evident. The coefficient of energy efficiency for existing houses more than doubles that of the reference, changing from 0.0448 to 0.0940, which represents a change in price premiums from 4.6 to 9.9 percent. In the remaining segments the distortions in the coefficients are smaller, but even so larger than 20%.

The empirical results highlight the importance of the use of transaction prices and of good quality size variables in hedonic price models as the use of error-prone variables, such as appraisals and the number of bedrooms, can lead to the introduction of a sizeable bias in energy efficiency parameter estimates.

5.3. Omitted variables

To illustrate the sensitivity of energy efficiency partial effect estimates to the omission of relevant variables, the regression hedonic models were re-estimated without a selected group of dwelling characteristics. The choice of the variables to be omitted rested on those quality attributes that were *a priory* deemed to have a reasonable correlation with energy efficiency (i.e., existence of central heating and/or air conditioning) and that are not often available in hedonic regression studies in this area of research (i.e., the scenic value of the location, location and construction quality of sold residential properties). A total of five omitted variable scenarios was considered, the first four merely omitting a single variable and a fifth one omitting all the selected variables for this experiment; see the bottom of Table 4.

The largest differences from the benchmark situation were obtained for the all-variable omitted variables scenario, as expected. For new apartments, this 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 substantial difference of 3.4 percentage points, the largest obtained for all scenarios. However, the omission of a single dwelling characteristic also produced relevant differences between benchmark and omitted variable results. This is the case when the two dummy variables measuring the quality of the location were omitted, which produced an upward shift in estimated coefficients. For new houses, this omission leads to a coefficient change from 0.055 to 0.066, which involves a non-negligible price premium increase from 5.7 to 6.8 percent. Moreover, the exclusion of the dummy variable controlling for the existence of central heating and/or air conditioning systems also produced a noticeable 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. On the other hand, 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.

6. Conclusions and discussion

The findings confirm that energy efficiency is positively rewarded in the Portuguese residential sales market. In particular, the results provide evidence of a clear difference between the way energy efficiency is rewarded for apartments and houses, with the former dwelling category yielding higher price premiums than the latter dwelling category. 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 these price premiums is sizeable at the point of means. For instance, considering that the average transaction price of an existing apartment is 97,695 €, the sales price premium corresponds to 12,212 € for most energy efficiency properties. This information is important for policy makers. Since houses represent the majority of the housing stock in Portugal [4], this result may play a crucial role in the definition of policy measures aiming at increasing energy efficiency standards in a cost-effective way.

The qualitative comparison across different studies and markets suggests the existence of higher price premiums in the Portuguese market than in central and northern European countries. Interestingly, the magnitude of the price premiums was found to be smaller than those estimated for Spain by Ayala et al. (2016) and also for Portugal by Ramos et al. (2015), in a work using list prices and a smaller sample than the one that is used in this paper. 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. A greater EPC label awareness and the existence of higher energy costs in Portugal are, in addition, possible explanations for the existence of greater price premiums. Moreover, the empirical exercise using different sample sizes provided no indication that significance levels of energy efficiency partial effects have been inflated by the size of the dataset available for research.

This paper provides valuable information for researchers. First, it shows that the use of appraisals and other type of surrogate prices in the left-hand side of the hedonic price model may distort significantly energy efficiency partial effect estimates. As such, any assessment

[4] 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 (authors' own calculations based on Census data).

about the impact of energy efficiency on residential property prices using appraisals, list or other type of non-market price should be seen with some care. This is an important finding since transaction prices are not always available in datasets. Second, the experiments using different omitted variables scenarios warn about the consequences of leaving out from hedonic regression models variables that measure the quality of the dwelling. If the models do not include then, energy efficiency partial effect estimates may be significantly biased and could provide wrong signals to all those interested in this area of research.

In general, our results clearly illustrate how partial effects of the most energy efficient ratings (i.e., A and B) substantially differ across dwelling types, with apartments displaying price premiums that are more than the double of those found for houses. A future avenue for research, which involves the application of quantile regression analysis, is the investigation of whether these differences are maintained for less and more expensive properties or if, on the contrary, this relationship changes along the price distribution of houses and apartments.

References

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

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.

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.

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

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

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.

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.

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

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

Fuerst, F. and Warren-Myers, G. (2018). Does voluntary disclosure create a green lemon problem? Energy-efficiency ratings and house prices. *Energy Economics,* 74, 1-12.

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.

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

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

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) (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) (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.

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.

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.

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.

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.

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

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., Pérez-Alonso, A., and Silva, S. (2015). *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.

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.

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

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.

Wooldridge, J.M. (2013). *Introductory Econometrics: A Modern Approach*. Fifth Edition. South-Western: Mason, USA.

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.

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.

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.

Appendix 1: Variable description

Explanatory variable	Variable description
DABSGAS	Dummy variable = 1 when the residential property is not connected to public or private gas distribution networks
DABSLIFT	Dummy variable = 1 when the residential unit is in a building with more than four floors and that does not have an elevator
DBADCONSERVATION	Dummy variable = 1 when the residential unit has a deficient conservation condition
DBADLOC	Dummy variable = 1 when the residential unit is located in an extremely bad location
DBGAPRTXCPL	Dummy variable = 1 for all apartments with more than 250 square meters, more than four bedrooms and located in an extremely good location.
DBIGAPRT	Dummy variable = 1 for all apartments with more than 250 square meters and more than four bedrooms.
DCONSTPi	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	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)
DCSYSTEM	Dummy variable = 1 when the residential unit includes a central heating and/or air-conditioning system
DDISTRCAP	Dummy variable = 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.
DLX	Dummy variable = 1 when the residential unit is located in *Lisboa*, the capital of Portugal.
DENERGYAB	Dummy variable = 1 when the EPC of the residential unit is either A^+, A, B or B^-
DEXCPLOC	Dummy variable = 1 when the residential unit is located in an extremely good location
DGRFLOORENOV	Dummy variable = 1 when the residential unit is a renovated house with less than 120 square meters.
DIRREGAREA	Dummy variable = 1 when the residential unit has non-standard areas, as defined by the Portuguese building code
DMROOMS	Dummy variable = 1 when an apartment has four or more bedrooms.
DPORTO	Dummy variable =1 when the residential unit is located in Porto, the second largest city in Portugal.
DPARKING	Dummy variable =1 when the residential unit has parking facilities.
DCOND	Dummy variable = 1 when the residential unit is located in a private condominiums.
DPRIVPARK	Dummy variable = 1 when the residential unit has individual parking facilities.
DREGIONi	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	Dummy variable = 1 when the residential unit has been improved or renewed.
DSCENICi	Set of three dummy variables identifying the quality of the landscape of the area in which the residential unit is located. 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.

Explanatory variable	Variable description
DSEA	Dummy variable = 1 when a property is located in parish that has access to the sea.
DSMALLBEEDR	Dummy variable = 1 for all house with less than three bedrooms.
DSWIMM	Dummy variable = 1 when the residential unit has swimming facilities.
Di	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.
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,
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).

Appendix 2: Complete OLS results

Explanatory variables	*Existing apartments*		*New apartments*		*Existing houses*		*New houses*	
	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	10.182**	1155.6	10.874**	602.5	10.899**	417.3
DENERGYAB	0.118**	61.7	0.123**	50.4	0.045**	7.3	0.055**	8.9
D2010	-0.004	-1.8	0.002	1.0	0.007	1.0	0.008	1.2
D2011	-0.071**	-28.9	-0.026**	-8.7	-0.067**	-9.5	-0.034**	-4.0
D2012	-0.152**	-56.4	-0.088**	-23.2	-0.131**	-17.2	-0.099**	-10.3
D2013	-0.182**	-67.5	-0.095**	-22.6	-0.158**	-21.0	-0.127**	-12.0
SQRTGRFA	0.144**	234.9	0.131**	149.3	0.082**	66.7	0.086**	43.1
SQRTDEPFLOORA	0.024**	49.0	0.033**	52.4	0.012**	18.7	0.012**	13.2
SQRTDWELLTRANSA	-0.030**	-25.6	-	-	-0.015**	-7.6	-	-
SQRTPLOTAREA	-	-	-	-	0.007**	28.6	0.006**	18.8
DIRREGAREA	-	-	-	-	-0.129**	-9.4	-0.133**	-5.3
DMROOMS	-	-	0.039**	7.6	-	-	-	-
DBIGAPRT	-	-	0.186**	6.9	-	-	-	-
DBGAPRTXCPL	-	-	0.160**	3.2	-	-	-	-
DSMALLBEEDR	-	-	-	-	-0.083**	-13.5	-	-
DRENOV	-	-	-	-	-	-	-0.164**	-16.3
DGRFLOORENOV	-	-	-	-	-	-	-0.084**	-5.6
DBADCONSERVATION	-	-	-	-	-	-	-0.147**	-4.6
DCSYSTEM	0.079**	24.1	0.070**	28.5	0.095**	9.7	0.071**	9.2
DABSLIFT	-0.071**	-22.4	-	-	-	-	-	-
DCOND	0.058**	9.7	0.065**	11.5	-	-	0.071**	5.0
DSWIMM	0.153**	32.2	0.169**	35.0	0.240**	24.5	0.271**	25.1
DPARKING	0.057**	25.2	-	-	-	-	0.039**	5.1
DPRIVPARK	-	-	-	-	0.089**	15.9	-	-
DABSGAS	-0.113**	-32.5	-	-	-0.089**	-17.1	-0.067**	-10.2
DCONSTP2	-0.113**	-32.5	-0.099**	-37.9	-0.088**	-10.3	-0.055**	-8.0
DCONSTP3	-0.155**	-28.2	-0.341**	-22.7	-0.167**	-14.1	-0.161**	-10.6
DCONSTP4	-0.144**	-16.1	-0.373**	-20.5	-0.310**	-19.1	-0.309**	-17.8
DCONSTQ2	0.056**	16.5	0.081**	28.4	0.024*	2.3	0.021*	2.4
DCONSTQ3	0.139**	15.1	0.137**	21.9	0.127*	3.8	0.109**	5.3
DREGION1	-0.365**	-120.5	-0.366**	-101	-0.369**	-41.7	-0.400**	-36.9
DREGION2	-0.296**	-125.1	-0.239**	-71.8	-0.255**	-32.2	-0.233**	-21.6
DREGION3	-0.252**	-98.6	-0.256**	-82.7	-0.331**	-43.9	-0.325**	-34.2
DREGION5	-0.039**	-5.9	-0.076**	-10.1	-0.214**	-19.4	-0.211**	-13.7
DREGION6	-0.010*	-2.9	-0.013*	-3.1	0.02	1.9	0.042*	3.4
DREGION7	-0.005	-0.7	-0.045**	-6.3	-0.004	-0.2	-0.057*	-2.6
DDISTRCAP	-	-	-	-	0.078**	8.6	0.117**	10.1
DSEA	0.113**	58.3	0.081**	31.6	0.159**	26.7	0.118**	15.8
DLX	0.349**	84.2	0.265**	44.2	0.299**	11.5	0.364**	7.6
DPORTO	0.331**	73.8	0.291**	46.9	0.278**	13.2	0.370**	8.2
DSCENIC2	0.100**	24.3	0.063**	14.0	0.145**	9.8	0.098**	5.3
DSCENIC3	0.266**	22.5	0.142**	15.0	0.248**	8.0	0.131*	3.4
DBADLOC	-0.171**	-39.9	-0.184**	-36.7	-0.154**	-25.1	-0.148**	-20.0
DEXCPLOC	0.316**	98.1	0.324**	68.5	0.475**	33.8	0.358**	18.1
Number of obs		149,920		59,410		33,282		13,533
Regressions' adjusted R^2		0.6758		0.7334		0.67		0.7531
RESET type test								
LM test statistic		0.3282		2.731		3.001		3.7717
p-value		0.567		0.098		0.0832		0.052

Notes: p-value < 0.05; ** p-value <0.0001.*