The Capitalization of Energy Labels for Post-1950 Residences in the Netherlands: A Comparison between Housing Types

July 1, 2022

University of Groningen

Daan Schipper

S3470180

COLOFON

Title	The Capitalization of Energy Labels in the Netherlands: A Comparison between Housing Types
Version	Final Version
Author	Schipper, Daan
Student Number	S3470180
Supervisor	Daams, M.
Assessor	van der Vlist, A.J.
E-mail	d.schipper.2@student.rug.nl
Date	August 22, 2022

ABSTRACT

This research paper uses hedonic regression analysis to examine the relationship between energy labels and housing prices in the form of government-assessed WOZ-value. Additionally, interaction effects are used to infer whether the extent to which energy labels are capitalized into housing price differs between housing types. Data from CBS's 2015 WoonOnderzoek is used, containing detailed data on residences in the Netherlands. The final sample consists of 2,534 properties. It is concluded that properties with an A energy label are assessed at a value that is between 4.1% and 5.5% higher than properties with an energy label of E or worse. Flats and apartments are the only housing type for which evidence is found which suggests that the price effects of a poor energy label are more negative than they are for other housing types. This could be a result of a more general relation between building quality and energy labels, but the results found still shed light on the problem that is energy poverty and its prevalence in low-quality flats and apartments. Future research could use more detailed data to further examine this relationship, and could incorporate the effect of energy price shocks on capitalization and energy poverty.

Keywords: *Energy Efficiency, Energy Labels, Energy Poverty, Housing Price, Capitalization, Hedonic Model*

Table of Content

1. INTRODUCTION

1.1 MOTIVATION

The Dutch market prices of both gas and electricity have been on the rise recently, as gas prices rose to approximately the sevenfold of their prices one year ago (Nu.nl, 2021). This has sparked an increased amount of interest in energy prices and energy efficiency. It remains to be seen in what way these prices will develop, but the short-term impact of these energy price shocks will be felt by the more financially vulnerable groups of the Dutch population. This can lead such groups to experience a concept called energy poverty, in which they are limited in their use of energy because they are forced to save on their energy bills as a result of the high prices (Rooyakkers, 2021).

This situation sparks interest in the extent to which energy efficiency measures in a residence are capitalized into housing prices. Energy efficiency measures are improvements that can be made to real estate to increase how efficiently the building utilizes energy, to reduce the amount spent on energy utility bills and create a more sustainable environment. Such energy efficiency measures include insulation, certain types of windows, and efficient appliances. Energy efficiency measures require an initial capital investment, but the cost of this investment can be repaid to the homeowner through lower utility bills over the years and an increased capital value of the house, and thus a higher transaction price if the homeowner decides to sell (RFF, 2020).

Households with a lower income, however, may not have the funds to invest in energy efficiency, or they might not be in a situation where waiting for the made investments to be repaid through lower energy bills or an increased transaction price of the residence is financially feasible. This makes them more vulnerable to energy poverty, especially in the event of energy price shocks, such as the one happening right now, as their residences are not energy-efficient enough to keep utility bills at a reasonable level (Tardy, 2019). Furthermore, those with lower income may be more inclined to rent as opposed to owning a home, which decreases the freedom they have to make structural modifications to their place of residence.

Households in flats and apartments might be more vulnerable to energy poverty due to having less disposable income. This might make it harder for them to make energy efficiency improvements depending on their financial situation, and especially in the event of significant price shocks such as the one happening at the time of writing, the issue of high utility bills could become more severe.

For households that have the financial means to potentially invest in energy efficiency, knowledge of the capitalization of energy efficiency measures has great importance in regards to informing the decisions of these households, as they compare the costs of these additional measures to the annual savings in energy cost and the expected increase in the value of the residence. Research has previously been done on this topic (e.g. Kahn & Kok, 2014), but varying results and contexts lead to energy efficiency capitalization not yet being completely understood. Additionally, it seems relevant to analyse this situation in varying contexts, as differences might be expected in the degree of capitalization between different countries and situations.

1.2 ACADEMIC RELEVANCE

One of the earliest instances of research into the capitalization of energy efficiency in the residential housing market was performed by Dinan & Miranowski (1989), who attempt to identify whether fuel savings as a result of energy efficiency investments such as insulation and solar applications are capitalized into housing prices. As part of their motivation, they discuss the existence of a possible 'energy efficiency gap', in which case the sub-optimal capitalization of fuel-saving measures would result in underinvestment in these measures.

To assess the value of energy-saving measures, the hedonic pricing method was used to identify the price effect of fuel cost-saving measurements such as insulation and solar applications. Energy efficiency was measured using a proxy variable and an instrumental variable approach was also used for this variable (which is adjusted fuel expenditures per heated square foot) to eliminate the correlation with the error term, as this would cause issues and bias in the analysis. It is concluded that the expected sales price of a house increases by $11.63 for a $1 decrease in the level of fuel expenditures necessary to maintain the house at 65°F in an average heating season. The present-day implications of the paper could be considered to be limited as it is dated and based on small sample size, but its importance, especially regarding methodology, remains significant.

Other research has found similar results, such as Kholodilin et al. (2017) who find that the 'Energy Performance' of a residence, which is measured as annual energy consumption in kilowatt-hours per square meter, is capitalized into housing prices in the German rental and owner-occupier market, and that willingness to pay is higher in the owner-occupier market, while Lyons et al. (2014) find that regarding the relation between energy performances of homes in Dublin, a 50-point improvement in Energy Performance Indicator is associated with a 1.5% higher list price.

Further recent research often uses Energy Performance Certificates or Labels, such as the Building Energy Rating as the independent variable measuring energy efficiency. Such studies include Aydin et al. (2020) who examined the effects of energy efficiency and Energy Performance Certificates (EPC) on residential sales prices. Using a hedonic model with data from the Dutch real estate market and a repeated sales approach, it concluces that a 10% increase in predicted energy efficiency is expected to lead to a 2.2% increase in the market value of the dwelling. No additional evidence which suggests that capitalization of energy efficiency is different for houses with an EPC was found.

Brounen & Kok (2011) find that energy labels create transparency in the energy efficiency of residences and that homebuyers are willing to pay a premium for more energy-efficient homes. Kahn & Kok (2014) find that a certified dwelling in California's housing market transacts with a premium of 2.1% compared to a non-certified home, which translates to a premium of $8400 on average. Additionally, annual energy savings for average consumers are expected to be around $720, which implies a payback period of around 12 years keeping the associated costs and savings in mind. Pontus & Cerin (2014) in similar research also confirm this relation, but in the Swedish real estate market, while Eichholtz et al. (2010) find significant evidence for this relation in commercial real estate.
While existing studies have clearly established this relationship between energy efficiency and housing price, housing types are often solely included as an independent variable that have an effect on housing prices in the analysis. The only instance of research that examines whether the extent of capitalization differs for different housing types was performed by Fuerst et al. (2015), who found that capitalization is higher for the 'flat' and 'terraced' housing types.

Different housing types are of different building qualities, with differences in structural characteristics as well, part of which can be explained by the different needs and possibilities of the residents of different housing types. These differences in structural quality and general building quality could lead residents of different building types to attach different values to certain levels of energy efficiency – which are proxied by energy labels. The effects of poor building quality cannot be perfectly distinguished from the effect of a poor energy label without a comprehensive building quality or depreciation variable, but the effects of poor energy labels for different housing types can still help in shedding light on the problem that is energy poverty, and the situations in which it prevails the most. Especially low-income housing with severely negative price effects of poor energy labels would show this problem, as it indicates a situation where the dwelling is cheap and of a low structural quality, it has a poor energy efficiency and thus high utility bills, and where investing into improving energy efficiency is costly, especially considering the low-income of a lot of these residents.

Whether capitalization is equal or different in these different housing types cannot be concluded from other research, as separate hedonic regressions or a regression including interaction effects would have to be performed for the different categories of housing. This research thus aims to add to the existing literature in the form of a case study of the Dutch residential market, contributing with an additional focus on potential differences in capitalization between housing types.

1.3 RESEARCH PROBLEM STATEMENT

While existing studies have clearly established the effect of energy efficiency on housing price, they have generally not addressed that the capitalization of energy efficiency may not be the same across different categories of housing types. Homeowners in certain housing types might not attach the same value to a level of energy-saving in that specific housing type as homeowners in different housing types would attach to a similar level of energy-saving. Therefore, examining whether the level of capitalization remains consistent between building types or whether it differs could provide additional, more precise information regarding the level of value increase one can expect as a result of investment into additional energy efficiency measures.

The goal of the present research is to use detailed data on housing type, general housing characteristics, housing prices, and energy efficiency – in the form of energy labels – in statistical analysis to isolate the price effect of energy labels on housing prices and thus to determine the extent to which energy efficiency is capitalized into the sales price of houses. Interaction can then be used for the different housing types to compare the degree of capitalization and to then be able to clarify whether this degree of capitalization differs between the various categories, or whether no significant differences in capitalization can be found. This thesis aims to further examine the relation between energy efficiency and the sales price of residences, as well as to examine the potential differences for this relation when comparing residential categories. The following main question has been formulated to reach this aim:

To what extent does energy efficiency impact residential housing prices, and does this vary across housing types?

The capitalization of energy efficiency into residential real estate sales prices is best analysed through a hedonic regression model. This requires a detailed dataset including a large number of residential real estate prices, and data on the relevant characteristics of the house such as its age, size, and other factors indicative of the house’s qualities that are expected to impact sales price. As the focus of this research is the effect of energy efficiency measures on sales price, detailed variables on this need to be included such as the residence’s energy label. The CBS’ ‘WoonOnderzoek Nederland’ will be used for this purpose, as it is publicly available for research. Moreover, the data includes all of the required variables listed above, such as sales price, house characteristics, and energy label. The hedonic regression analysis that is to be performed will show to what extent energy efficiency is capitalized into housing prices.

After the degree of energy efficiency capitalization has been determined using hedonic regression analysis, the effect of housing type on the degree of capitalization will be examined. In the first analysis, housing type is assumed to be an independent variable affecting housing price, while energy efficiency independently affects housing price. To examine whether capitalization varies between housing types, a model using interaction variables for housing type and energy label will be used. The coefficients for these interaction variables can then be compared, allowing insight into the degree of capitalization and its differences across residential categories. Separate regressions for each housing type will also be performed to see whether this provides different results.

The remainder of this paper is organized as follows. Section 2 describes the theoretical background and section 3 describes the data and the exploratory analysis. Section 4 presents the results and discusses these results, and section 5 concludes.

2. BACKGROUND

2.1 CAPITALIZATION

Housing price capitalization is the extent to which housing prices change as a result of certain factors. These can be physical factors such as the size, level of maintenance and type of house, but non-physical factors such as local tax rates or neighbourhood cohesion can also affect housing prices. Capitalization can also be seen as “an accounting method in which a cost is included in the value of an asset and expensed over the useful life of that asset, rather than being expensed in the period the cost was originally incurred” (Investopedia, 2022).

This definition makes a lot of sense in the context of energy efficiency capitalization; both the cost of investments into energy efficiency and its benefits, which are lower utility bills and less environmental pressure, are taken into account when determining its value, and the cost of improving energy efficiency is directly capitalized through a higher capital value of the residence.

The level of capitalization depends on certain factors. Hilber (2017), for example, find that public and private investments into housing are capitalized at a higher degree in locations with strict geographical or regulatory supply constraints. Capitalization differs for individuals as not everyone attaches equal value to less environmental pressure, for example, and the premiums individuals are willing to pay for a higher level of energy efficiency thus vary, although generalizations can be made using certain statistical methods.

Building characteristics are capitalized into housing price (e.g. Kholodin et al., 2017). The value of implicit characteristics of heterogenous goods such as housing can be estimated through a hedonic pricing method using an OLS Regression Model. While this method has been around for a long time, it was modernized for uses such as the current one by Rosen (1974), with the goal of attaching a value to specific characteristics of relatively heterogeneous goods in a market of pure competition. This makes the hedonic pricing method very fit for estimating the value of heterogeneous characteristics – such as energy efficiency – in the real estate sector.

2.2 ENERGY EFFICIENCY MEASURES

The degree of energy efficiency of a home is generally proxied by a constructed variable, as no universal measures of energy efficiency exist. Examples from the literature include Fuel Cost Savings (Dinan & Mirowski, 1986 & Kholodilin et al., 2017), which is the capital saved when heating to a specific temperature during the heating season as a result of improving energy efficiency measures, and forms of Energy Performance/Use (Lyons et al., 2014), which are more precise measurements of actual energy use, as higher energy efficiency is considered to lead to less absolute energy usage. However, the Energy Performance Certificate (EPC) or energy label for a residence is seemingly the most used in the academic literature on the energy efficiency of buildings. Energy Performance Certificates, often referred to as energy labels, give buildings a grade for their energy efficiency, generally using a scale where an A represents the most energy-efficient buildings, and where a G represents the least energy-efficient buildings. An EPC can be obtained by having a qualified assessor asses the building in question. Energy efficiency, and thus EPCs are capitalized into housing prices, but the added effect of an EPC and its information value remains questionable as there has been research that found that energy-efficient homes sell at a premium even without an EPC (Aydin et al., 2020 & Zhang et al., 2020). Even if this is the case, nonetheless, home energy ratings do provide benefits in regards to helping buyers identify properly priced homes (Gilmer, 1988).

One reason for the increased use of EPCs is the European Union's 2002 Energy Performance of Buildings Directive (EPBD), which was later recast in 2010 with enhanced requirements. The goal of the EPBD is to improve energy performance across the European Union through increased transparency, heightened minimum building codes, Energy Performance Certificates, and regular checks of climate-related systems within buildings (European Commission, 2020).

All member states of the European Union have since implemented the requirements associated with the EPBD into their legislation. Generally speaking, countries have their version of an EPC, and having an EPC set up is mandatory when selling or renting out residential properties, with certain types of buildings being exempt from this as exceptions. In the Netherlands, the Energy Label is mandatory for houses that are being sold or rented out. The energy performance of the residence is ranked with a label, with A++++ labels having the best energy performance and G having the worst energy performance. Additionally, the energy label includes information on how the residence could become even more energy-efficient, and on how dependent the residence is on natural gas use (Rijksoverheid, 2022). Energy labels are assumed to have an effect on transaction prices in the Dutch real estate market (Aydin et al., 2020)., assuming that the energy labels are transparent and readily available to potential buyers.

2.3 ENERGY EFFICIENCY GAP & UNCERTAINTY

Finding out to what extent energy efficiency is capitalized into housing price and how this differs for different housing prices should provide home-owners with a more accurate estimation of the returns they can expect in the form of a higher housing price when they decide to invest into energy efficiency. Such information not being available could limit investments into energy efficiency, as homeowners will experience a degree of uncertainty when estimating the returns on their investment. Not investing as a result of said uncertainty, while the investment would be economically viable if such information was viable, is one of the things that can lead to an energy efficiency gap.

In the theory of the energy efficiency gap, certain issues cause structural underinvestment into energy efficiency. Allcott & Greenstone (2012) describe it as 'a wedge between the cost-minimizing level of energy efficiency and the level actually realized' (p. 4). Though the magnitude of the gap is hard to estimate, the existence of the gap seems generally agreed upon. Dinan & Miranowski (1989) suggest that a possible reason for this underinvestment is homeowners who adopt a life cycle cost approach, but only consider their expected tenure in the home, and not the added value it brings to the residence after their tenure in the form of a higher sales price. This would lead to an undervaluation of the returns on investment in energy efficiency, especially if the expected remaining tenure in the homeowner's current residence is relatively short

Aydin et al. (2020) also discuss this energy efficiency gap, and state that a lack of transparency in energy efficiency can lead homeowners to inaccurately be able to assess the energy efficiency of a residence and the benefits regarding future energy savings, leading to underinvestment. Other factors that could explain the existence of an energy efficiency gap are stated to be hidden costs of investing in energy efficiency and principal-agent problems, while Kholodilin et al. (2017) add that heterogeneous incentives, especially for owner-occupants and landlords, can be a potential explanation.

Gillingham & Palmer (2014) discuss the market failures that lead to the energy efficiency gap, and why the gap may be smaller than generally expected; it may be overestimated because of hidden costs, consumer heterogeneity, uncertainty, overestimated savings, and the rebound effect (p.20). Policy should address market failures and behavioural failures that lead to an energy efficiency gap, through the three primary types of energy efficiency policy: information strategies, economic incentives, and energy efficiency standards. Information strategies seem to be the most relevant, and economic incentives are the least relevant. The information strategies include programs such as the USA's Energy Guide Labelling Program for appliances, but also the Energy Star label, which is similar to European counterpart discussed above, and which can also be applied to homes and commercial buildings. Energy Labels are thus a way to increase transparency and provide more insight into the actual energy efficiency of a home, and an accurate energy labelling system should be able to play a role in reducing the severity of any possible existing energy efficiency gap.

Despite this, there are still certain issues that can be associated with the energy label. EPC assessor methods are not always consistent which can be a problem, though the importance of this consistency could be questioned. However, comparing methods and sharing the best approaches could be beneficial and lead to a better and more flexible approach (Semple & Jenkins, 2020). Additionally, the expected energy performance and the realized energy performance may not always match up. For labelled residential buildings in Switzerland, it was found that on average, buildings' actual energy consumption was 11% lower than theoretical energy performance, but this difference varies significantly for the different Energy Label Ratings. G-rated buildings' consumption was found to be 40.4% lower in actuality than theoretical, while A-rated buildings' consumption was 6.2% lower than theoretical, suggesting that buildings with a high Energy Label Rating may be more robust to this gap between actual and theoretical energy performance, and that the energy performance of buildings with a low energy label may be underestimated (Cozza et al., 2020).

2.4 HYPOTHESES

The degree of energy efficiency of a building is assumed to affect the transaction price of the building due to the lower utility bills and the more positive environmental impact associated with high energy efficiency. Lower utility bills and better sustainability will be reflected in the sales price of building, therefore the first hypothesis is formulated as follows:

H1: Energy Efficiency Label of residences will be capitalized into residential housing prices

The effect of housing types on residential sales prices has been documented in previous research. Hitherto, the impact of housing type on the degree of energy label capitalization is not prevalent in previous research, and has only been performed by Fuerst et al. (2015), who found that the capitalization of energy efficiency is higher for dwellings of the 'terraced' and 'flat' type than the capitalization is for dwellings of the 'detached' and 'semi-detached' type. As a result of differences in structural characteristics and building quality across housing types as well as differences in the value that residents of different housing types attach to a certain level of energy efficiency, differences in the extent of capitalization of energy labels are expected, and therefore, the second hypothesis is as follows:

H2: The extent to which labels are capitalized will be different for different housing types

3. METHODOLOGY & DATA

3.1 METHODOLOGY

A hedonic pricing model in the form of an OLS regression model will be used to estimate the impact of energy labels on housing prices. Hedonic regression allows researchers to estimate the effects that heterogenic characteristics of goods have on their price, and thus allows the implicit value of such characteristics to be judged.

The variable for housing price used in the model is the WOZ-value of the building as of 2014. WOZ-values of buildings are their estimated values as determined by municipalities, which estimate the value of properties using a model that compares the building in question with sales prices of recently sold comparable properties. Another option was the most recent sale price of the property, but this would lead to inaccurate results, especially for buildings that were sold long ago, as the sale price may no longer be representative of the building as certain characteristics such as energy efficiency may change over time. The only other option within the limits of the data was self-judged expected current sales price, which is likely more inaccurate than the WOZ-value as it is an estimation made at the moment of the survey by the owners, who are not experts and who are less likely to have spent time researching the transaction prices of comparable properties.

The energy label of residences is used as a measure of energy efficiency. There is a definitive version and a provisional version of the variable. Despite the provisional version having significantly more observations, the definitive version is used, as it is assumed to reflect the energy efficiency of residences significantly more accurately than the provisional version. A model using an alternative measure of energy efficiency was also specified: yearly costs of energy and water use per m2. A similar variable has been used in previous research on the effect of energy efficiency on housing prices (Lyons et al., 2014). This model was found not to improve the explanatory power as opposed to using labels and is thus omitted from further analyses.

Housing Type is considered to be the type/category of residence that a household resides in. Most research using hedonic regression on housing price considers housing type to have a significant effect on housing price, and this is confirmed by studies previously discussed such as Brounen & Kok (2010). The semi-detached category turns out to have the most favourable effect on housing prices in their research, with apartments transacting with the lowest average price.

Hitherto, the degree to which housing type affects the capitalization of residential energy efficiency into housing price is not often discussed in academic literature. Fuerst et al. (2015) do research this, they conclude that the capitalization of energy efficiency is higher for dwellings of the 'terraced' and 'flat' type than it is for dwellings of the 'detached' and 'semi-detached' type. It is concluded that energy efficiency labels have a significant impact on housing prices in the study area of England.

Control variables are included in the model to attempt to eliminate external factors that are outside of the model influencing the results and thus causing bias in the coefficients of other independent variables, which would reduce the reliability and interpretability of the results. Such control variables include the year the residence was built in 10-year categories, whether the residence has a balcony and/or garage, and the size of the residence. The control variables used are based on previous literature (e.g. Lyons et al., 2014 & Cerin et al., 2014). Location fixed effects are also included in the form of COROP-regions, which are 40 areas that the Netherlands is divided into. This form of locational fixed effects was used as it is the most precise locational variable included in the dataset.

A base model will be specified which includes the energy label as the only independent variable. The second model will add housing characteristics, and the third model will add locational fixed effects. Model 4 will add a polynomial for the surface area of the property, which was found to significantly increase explanatory power. In model specifications 2, 3, and 4, housing type is considered to be an independent variable that influences the WOZ-value of residences. In model 5, interactions between housing type and energy label are added to the regression analysis to be able to determine whether the price effects of energy labels are different or consistent for different energy labels.

The WOZ-value variable was transformed into the natural logarithm which is standard in hedonic regression with housing price as the dependent variable (e.g. Kahn & Kok, 2014). This improves interpretability, as the effect of the energy labels and other variables in the model can be interpreted as a percentage change instead of an absolute increase in value in the form of a certain amount of euros. Transforming the dependent variable into a natural logarithm also improves the functional form and fit of the model. Buildings with more than 20 rooms in them were excluded from the analysis. Additionally, the WOZ-value variable was winsorized at the 1% level, meaning the top and bottom 1% of the observations were excluded from the analysis, as significant outliers will influence results and cause bias.

Houses that were built before 1950 were excluded from the analysis, as the data was inconsistent and created some bias in the coefficients. For a complete overview of all data modifications performed, consult the do-file of the Stata analysis found in Appendix 2. Several models are specified to show the effects of the key independent variable when other control variables are not included. The regression formula for model 4 is as follows:

$$lnWOZ_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \beta_3 X3_i + \beta_4 X4_i + \beta_5 X5_i + \beta_6 {X5_i}^2 + \beta_7 X6_i + \sum_{c=1}^{c} \varphi_c X_i + \sum_{z=1}^{40} \gamma_1 Z_1 + \varepsilon_i \quad (1)$$

In Equation 1, the left-hand side represents the natural logarithm of the WOZ-value of the residence *i*. On the right-hand side, β_0 represents the constant. X1 represents the energy label of residence *I*, and β_1 represents the coefficient associated with this energy label. X2 represents housing type, X3 represents whether it has a balcony, and X4 represents whether the building has a garage or carport. Additionally, X5 represents the square footage of the house in meters and X6 represents the number of rooms in it. Lastly, $\sum_{c=1}^{c} \varphi_c X_i$ represents dummy variables used for the age of the building, while $\sum_{z=1}^{40} \gamma_1 Z_1$ represents location fixed effects through the form of Dutch COROP-regions. Lastly, ε_i represents the error term.

$$lnWOZ_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \beta_3 X3_i + \beta_4 X4_i + \beta_5 X5_i + \beta_6 {X5_i}^2 + \beta_7 X6_i + \beta 8(\mathrm{X1} * \mathrm{X2})_i + \sum_{c=1}^{c} \varphi_c X_i + \sum_{z=1}^{40} \gamma_1 Z_1 + \varepsilon_i \quad (2)$$

Equation 2 adds to the regression formula of model 4 by including an interaction variable that interacts energy label and housing types. The interaction is represented by $(\mathrm{X1} * \mathrm{X2})$, with $\beta 8$ representing the coefficient associated with the interaction of the two variables.

3.2 DATA COLLECTION

The dataset used for the analysis is the 2015 version of CBS's WoonOnderzoek (CBS, 2015). The dataset contains information regarding the Dutch population and their wants and needs regarding houses, containing details about the current residence, the environment of the residence, and more information that is less relevant for this analysis. The 2015 version of the dataset was chosen as opposed to the more recent 2018 dataset, as the 2015 dataset includes more detailed information regarding the location of the residences in the dataset, which is considered to be an important determinant of house price (e.g. Kahn & Kok, 2014, Fuerst et al., 2015).

3.3 DESCRIPTIVE STATISTICS

Figure 1 shows the geographical distribution of the 2,534 cases that ended up being taken into account for the analysis. This is significantly less than the 62,668 cases included in the standard dataset. Table 1 shows all modifications that were made to the data which resulted in cases being dropped from the analysis. As can be seen, the majority of the cases that had to be dropped were dropped as a result of a missing energy label for the residence in the dataset, or because the residences were rented, instead of occupied by a home-owner. The associated Stata-code used to perform these modifications can be found in Appendix 2.

Table 1: Sample Selection Bookkeeping Process

Data Modification Made	**Remaining Cases**
Starting number of residences	62,668
Dropping residences that are missing 'Energy Label'	21,198
Dropping residences that are missing 'Housing Type'	19,233
Dropping residences that are 'Rented'	4,000
Dropping residences that were not yet constructed when purchased	3,306
Dropping residences built before 1950	2,583
Dropping residences with a WOZ-value over €1,000,000	2,577
Dropping residences whose 'Year Purchased' was before 'Year Built'	2,554
Dropping residences which are in the 'Other' housing type	2,534

COROP-regions are used for both the analysis and Figure 1, as they are the most accurate and most readily available measure way of implementing spatial fixed effects in the analysis. The 2014 version of the COROP regions was used, as the dataset used originates from 2015, and thus did not yet use the 2018 version, which added slightly more sub-regions to the COROP-regions. As can be seen, cases are generally rather evenly distributed, with a lack of cases in the North and a higher concentration of cases near the centre of the country, as well as specifically towards the West, where the more densely populated regions of the Netherlands are. The geographical distribution of the cases is also influenced by the fact that only houses built after 1950 were included in the analysis, as more houses have been built in these densely populated areas since then, as they have experienced more population growth. Because locational fixed effects are included in the analysis, the cases do not need to be geographically distributed in a way that is representative of the distribution of all residences in the Netherlands.

Table 2 shows the descriptive statistics of the main ratio and interval variables that are used in the analysis. The mean WOZ-value of the houses included in the analysis is €192,504, which is adequately close to the mean WOZ-value across the Netherlands in 2015, which was €206,000 (CBS, 2020). The mean surface area is 115, compared to the 2013 national average for houses of 120 (CBS, 2013). The fact that the average year houses in the analysis were built is 1975 seems logical, as only houses that were built between 1950 and 2015 were included in the analysis. In general, none of the values for these descriptive statistics deviate far enough from national averages to generate any concern regarding the representativeness of the dataset.

Table 2: Descriptive statistics of ratio/interval variables included in the analysis. (N=2534)

Ratio Variables	**Mean**	**Std. Deviation**	**Min**	**Max**
WOZ-Value	*€192,504*	*€93,592*	*€10000*	*€897000*
Log of WOZ-Value	12.076	.415	11.156	13.349
Costs of Energy and Water Use	*€142.03*	*€54.82*	*€22.23*	*€493.03*
Costs of Energy and Water Use per Meter	*€1.30*	*€0.45*	*€0.13*	*€4.06*
Number of Rooms	4.486	1.346	1	16
Surface Area	115.206	49.191	25	600
Year Built	1975	14.43	1950	2013

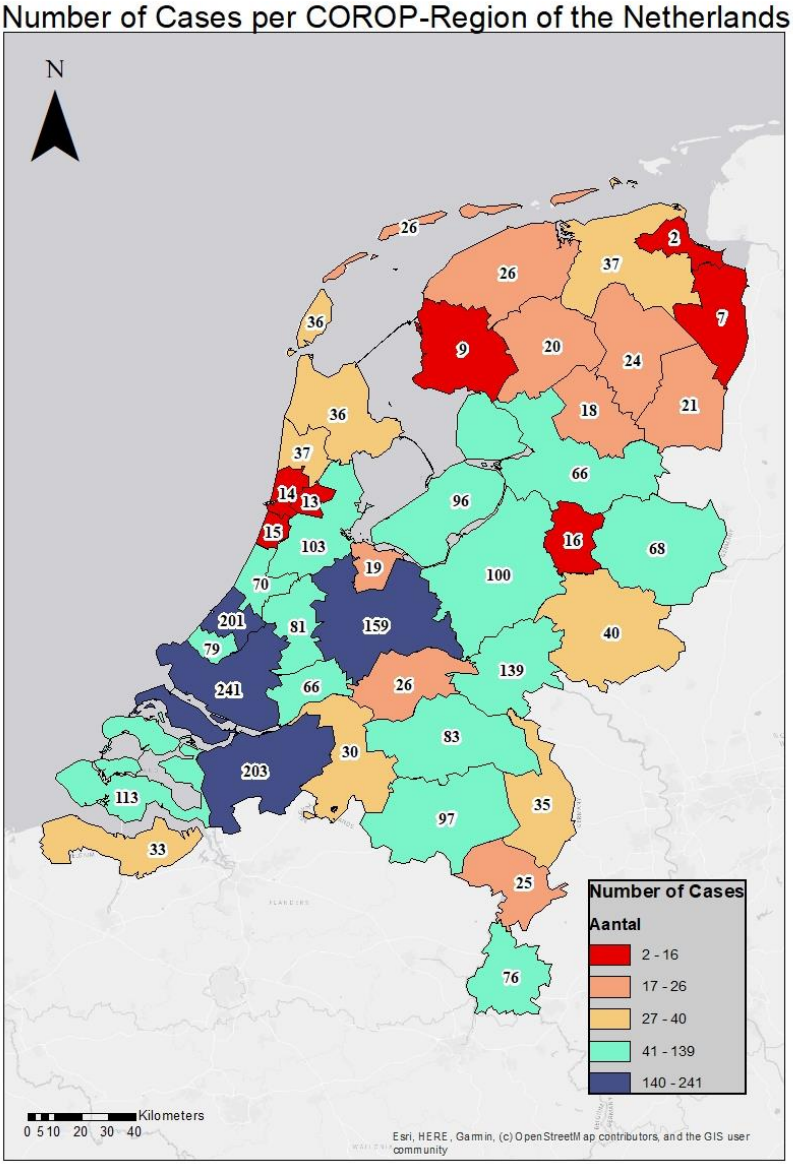


Figure 1: Distribution of the cases included in the analysis. (Esri, 2022)

Table 3 shows the distribution of nominal and ordinal variables included in the analysis. Regarding energy labels, more than half of the observations have an energy label of C or D. The distribution is similar to the distribution of energy labels across all residences in the Netherlands (RVO, 2015). For the analysis, the categories E, F, and G were combined to increase the sample size and to provide more consistent results. For housing type, most residences are either Terraced or Flats/Apartments, with the rest consisting of (semi-)detached properties. The ‘Other’ category was removed here, as it was inconsistent, unclear, and had a very low sample size. Most residences do not have a garage or carport but about 35% do, which is similar to balconies; around 38% of properties have access to a balcony while the rest does not.

Table 3: Descriptive statistics of nominal independent variables included in the analysis.

Other Variables	Value	Frequency	Percentage
Energy Label	A	197	7.77
	B	361	14.25
	C	853	33.66
	D	496	23.52
	E	323	12.75
	F	139	5.49
	G	65	2.57
Housing Type	Flat or Apartment	705	27.82
	Terraced	1260	49.72
	Semi-detached	310	12.23
	Detached	259	10.22
Residence has Garage/ Carport	Garage	801	31.61
	Carport	98	3.87
	Neither	1635	64.52
Residence has a balcony	Private Balcony	937	36.98
	Shared Balcony	21	0.83
	Neither	1576	62.19

Appendix 3 shows the correlation between the dependent variable and the independent variables used in the analysis. A correlation of 0.7 or higher is generally found to be too high (Abachnick & Fidell, 1996), leading to multicollinearity that will induce bias in the results of the regression. However, no variables share a bivariate correlation factor higher than 0.7. Additionally, after the main regression in Model 4 was performed, the Tolerance and VIF of all the variables used were calculated. The results of this can be found in Appendix 4. Some critical values with a VIF of higher than 10 were found, but these were found for certain COROP regions with a very low number of observations. Their low number of observations should not cause bias.

3.4 LIMITATIONS

Due to modifications having to be made to the research, mainly in the form of data management, removing outliers, and removing cases that do not apply to the specifics of this research (for example rented properties), the total sample size fell to 2,534. The exact modifications that were made can be found in the Stata do-file in Appendix 2. Higher sample size could have made it more likely for dependent variables to turn out to be significant, and it could have led to a higher degree of reliability in the results.

Apart from that, the lack of a more reliable dependent variable might hurt the results. WOZ-value was chosen as it was deemed more appropriate than self-judged current value or the most recent transaction price. Transaction price would, generally speaking, be the best option for the dependent variable in research structured such as this one, but the observations in the data set are too spread out, as only a small portion of properties were sold recently, and recent sales are required in this case as other variables, such as energy label, might not be representative of the state of the property at the time of the sale. The results found in this research are limited by the scope of the research and the lack of variables related to depreciation and maintenance in the dataset. Especially these missing or inaccurate variables might impact the results.

3.5 ETHICAL CONSIDERATIONS

All observations in this research are anonymous. The location and owners of properties cannot be determined using the given data, especially considering the limited precision of the locational variable, which is COROP-regions. The research is performed by and with the viewpoint of a student writing their master’s thesis.

4. RESULTS & DISCUSSION

4.1 RESULTS

Table 4: OLS regression results.

Variable	(1)	(2)	(3)	(4)	(5)
Energy Label					
A	-	-	-	-	-
B	-.152***	-.019	-.024	-.023	-.092*
C	-.245***	-.005	-.026	-.018	-.105**
D	-.371***	-.047	-.049*	-.044*	-.176***
E or lower	-.399***	-.042	-.057**	-.05*	-.206***
Housing Type					
Flat or Apartment	-	-	-	-	-
Terraced	-	.154***	.217***	.173***	.031
Semi-Detached	-	.242***	.372***	.302***	.187***
Detached	-	.509***	.643***	.535***	.382***
Housing Characteristics	No	Yes	Yes	Yes	Yes
Construction Year Categories[1]	No	Yes	Yes	Yes	Yes
Location Fixed Effects	No	No	Yes	Yes	Yes
Polynomial of Square Footage	No	No	No	Yes	Yes
Energy Label * Housing Type Interaction [2]	No	No	No	No	Yes
R-Squared	8%	59.5%	73.6%	76.4%	76.9%

Dependent variable is the natural logarithm of residential WOZ-value.

(= Coefficient is significant at 90% confidence level. ** = Coefficient is significant at 95% confidence level. *** = Coefficient is significant at 99% confidence level.)*

[1] *Construction year is categorized into groups of a 10 year, interval, starting with the 1950-1960 category.*
[2] *Precise coefficients can be found in Appendix 1, while calculated price effects can be found in Table 5*

The full results of the regression, including interaction variable values can be found in Appendix 1. The full Stata-code used for these results can be found in Appendix 2. In the first model specified in Table 4, energy labels are the independent variable being regressed on the natural logarithm of the residence's WOZ-value. All categories of energy labels have a significant effect on the natural logarithm of WOZ-value at a 99% confidence level. The coefficients can be transformed into a percentage-based price effect using this Equation 3:

$P_{\%}WOZ = (EXP(\beta_x X_i - 1) * 100$	(3)

Where β is the coefficient of variable X for residence I, and P is the percentage change in the assessed WOZ-value of the property. These findings suggest that in this model, properties with an energy label of E or lower are assessed at a 32.9% lower WOZ-value than properties with an A energy label. The model is estimated to explain approximately 8% of the variation in the natural logarithm of WOZ-value.

In the second model, housing types and characteristics are added to the model, improving the explanatory power to 59.5%. Flats or apartments transact at the lowest price, followed by terraced housed and semi-detached housing, and detached properties transact at the highest price. Though coefficients change when adding or removing variables, this order of price effects regarding the different housing types remains consistent in the following models. The housing characteristics variables added, which are whether the building has a balcony, whether it has a garage, the number of rooms, the square footage of the property, and the year the residence was built divided into 10-year categories, all have a significant effect on the natural logarithm of WOZ-value. As would be expected, buildings with a balcony or garage transact at a significantly higher price, and so do residences with more rooms, higher square footage, and buildings that were built more recently.

The third model adds locational fixed effects in the form of COROP-regions. This causes a substantial increase in explanatory power when compared to model 2, from 59.5% to 73.6%. Nearly all of the 40 COROP-regions have a significant effect on the dependent variable, with the few exceptions generally being regions with lower sample size. For specifics regarding the sample size and coefficients for certain COROP-regions, refer to the map that is Figure 1 and the full coefficient of the regression in Appendix 1.

Model 4 adds a polynomial, which is the surface area of the residence but squared, allowing for a better model fit in the case of non-linear relationships between the square footage variable and the WOZ-value of the residence. The square root variable is significant at the 99% level and increases the explanatory power of the model by 2.8% when compared to the previous model. The variable's negative signs indicate diminishing returns for price effects of the square footage of the building. The estimated effects of square footage on WOZ-value are visualised in Appendix 6.

In model 5, the interactions between energy labels and housing types are added to the model. Doing so increases the explanatory power by 0.5%. Specific coefficients for the interactions can once again be found in Appendix 1 .

Table 5 shows the percentual price effects of energy label on housing price for the various housing types. The data was computed by adding up the coefficients of Model 5 for the energy label and the interaction between energy label and housing types (for the relevant category)– exact coefficients can be found in Appendix 1 – and then transforming them into a marginal price effect using Equation 3. The effect of housing type is accounted for in Model 5, but not included in the effects in Table 5 as they would overcomplicate the table and make the effect less comparable.

Table 5: Model 5 Marginal Price Effect of Energy Label for various housing types. (= Coefficient is significant at 90% confidence level. ** = Coefficient is significant at 95% confidence level. *** = Coefficient is significant at 99% confidence level.)*

Percentage	Flat or Apartment	Terraced	Semi-Detached	Detached
A	0%	0%	0%	0%
B	-8.76%*	0.54%*	-7.49%	2.37%
C	-9.96%**	-0.72%**	-1.02%	5.13%**
D	-16.15%***	-0.03%***	-2.55%**	-4.52%
E or lower	-18.59%***	0.08%***	-0.33%***	1.95%***

The values that can be found in Table 5 are visualised in a graph in Figure 2. Looking at this table, most coefficients for the flats and apartments, and terraced housing types are significant, with a B energy label being the only label significant at a level of less than 5%. Apart from that, only D and E or lower are significant for semi-detached residences, and only C and E or lower are significant for detached properties. Keep in mind that the A label would have likely been significant, but it was chosen as the reference category due to its large sample size and the fact that it was expected to have the most beneficial effect out of the available energy labels, and it thus is a normative standard. The lack of significance of about half of the interactions means that the estimated coefficients for them are not all accurate, which reduces the interpretability of these results, as they may include inconsistencies.

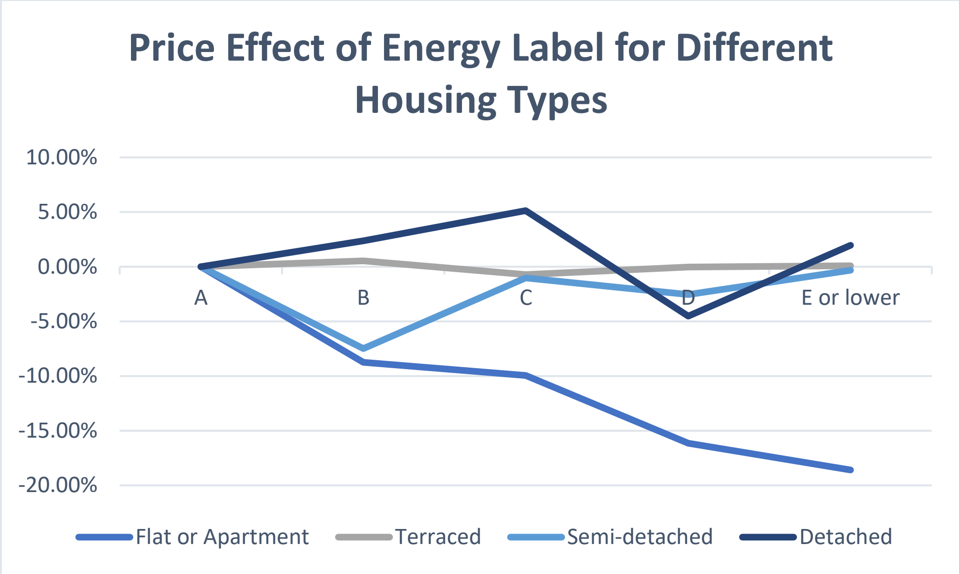


Figure 2: Price effect of energy label on residential WOZ-value for various different housing types.

Lastly, Table 6 shows an alternative method of attempting to record the effects of energy labels on housing prices for the different housing types. In this table, separate regressions were performed for each housing type instead of adding interaction effects. The regressions were performed in the same way and with the same variables as for Model 4, but only for specific housing types. This drastically reduces the sample size for all housing types, and therefore none of the coefficients are significant as a result.

Table 6: Regression price effects for separate regressions performed separately for each housing type

	Flat or Apartment	**Terraced**	**Semi-detached**	**Detached**
A	0	0	0	0
B	-6.69%	-0.08%	-1.76%	-3.54%
C	-3.45%	-3.03%	-2.86%	-1.35%
D	-7.82%	-2.01%	-3.88%	-10.24%
E or lower	-7.60%	-3.44%	0.01%	-3.84%

A table showing the sample sizes for all different combinations of housing type and energy label can be found in Appendix 7, which can help with the interpretation of both these results and the results illustrated in Table 5.

4.2 DISCUSSION

Though not always significant, the coefficient for energy label E or lower when compared to residences with an energy label of A indicates a negative price effect of approximately -4.1% (Model 2) to -5.5% (Model 4). In models 2 to 4, the price effect of the energy label goes down as the energy label of the residence gets worse, as would be expected. The price premium for detached residences when compared to flats or apartments seems to vary from a 66% price premium (Model 2) to a price premium of approximately 90% (Model 3).

The results of other research on this topic seem to find somewhat comparable results regarding the price effects of energy labels and housing types. Brounen & Kok (2011) estimate the price effect difference between A-level properties and properties rated E or lower to be between 11.29% and 16.17% in their Heckman two-step estimation. Fuerst et al. (2015) find that dwellings rated at an A or B sell at a premium of 5% compared to dwellings rated at a D. They find a smaller difference between detached dwellings and flats/apartments, a difference of around 20.4%. Other research is less comparable as other measures of energy efficiency are used in a large number of cases, and the context often differs.

In this research, after using a pooled sample of all housing types, two different ways of analysing the extent to which energy labels are capitalized for different housing types were performed. First, interaction effects between housing type and energy labels were added in Model 5. Additionally, separate regressions using the form of Model 4 were performed for each housing type. The price effects associated with energy labels for each different housing type can be found in Table 5.

Table 5 shows the fact that it only the 'flat and apartment housing' type displays consistent significance and results that, as would be expected in general, display continually decreasing price effects as the energy label gets lower. The other housing types displayed in the research do not share this same level of significance as the flat and apartment housing type, and the coefficients associated with these other housing types also display more non-linear, seemingly more random patterns. Fuerst et al. (2015), who also researched potential differences in energy label capitalization for different housing types found that capitalization is higher for the housing types 'flat' and 'terraced'. In this case, significantly negative results are found for the 'flat and apartments' housing type, but the 'terraced' and the other housing types do not display significant and consistent results.

Similarly, Table 5 shows the partial results of the separate regressions performed for each housing type. Once again, flats and apartments show the most relevant price effect patterns, yet none of the coefficients were significant, likely due to the lower sample size as a result of splitting up the sample. The differences in coefficients for the pooled model with interactions compared to the separate models is a result of the housing type being removed from the analysis, as all observations for the separate models have the same housing type.

Using both the results from Table 5 and Table 6, it seems that for the potential existence of an interaction between housing type and energy label on WOZ value, the main housing category to observe is the flat and apartment housing type. The increased significance of this interaction is, however, probable to be correlated to the increased depreciation of flats compared to other housing types. Flats and apartments are often older than other housing types, showing increased depreciation, and the exclusion of variables that directly impact the independent variable can lead to the price effect variation of such exogenous variables being partially absorbed by other variables which are in the model, and correlated with the exogenous variable. In this case, the increased depreciation of flats and apartments might thus lead us to overestimate the negative price effect that a poor energy label has, as flats and apartments with a poor energy label seem likely to be severely depreciated properties. This issue could be fixed by including an accurate and up-to-date variable measuring residence's depreciation, but the dataset used for the analysis does not include such a variable.

The other housing types, which are not assumed to suffer as much from the aforementioned issue with depreciation, do not display significant enough effects to be able to assume that an interaction effect between housing type and energy label exists for these categories. This could indicate that there is a lack of interaction effects, or it could be a result of missing variables and other inaccuracies in data, such as the lack of a more accurate dependent variable.

The high negative price effect associated with flats and apartments with a poor energy label could be a result of such apartments being correlated with a low building quality and a high degree of depreciation in general. The effects of such poor building quality/depreciation cannot be perfectly distinguished from the effect of a low energy label, but the results still bring attention to the severity of the issue that is energy poverty, and the fact that it is mainly prevalent in low-budget flats and apartments. The value of such flats and apartments is low as a result of low building quality and low energy efficiency, while utility costs are high due to low energy efficiency, and residents likely do not have the spare capital to invest into energy efficiency, which may not even be worth it in the first place in these low building quality residences.

Relevant missing variables that could improve the analysis could be depreciation, as mentioned before, and more accurate locational characteristics – the COROP-regions used in this research are not at an accurate enough spatial scale, and they allow for too much differentiation regarding spatial characteristics within these regions. Especially more detailed neighbourhood characteristics would be useful to more accurately capture the variation in housing price as a result of these spatial differences.

5. CONCLUSION

This research attempts to quantify the effect of energy labels on housing as well as to attempt to identify any differences in this energy label premium for different housing types. In the first model, all energy label coefficients are significant at a high significance level; The negative price effects seem to increase at a steady rate as the energy label goes down, which is as would be expected. The coefficients overestimate the actual price effect of energy labels, as a result of a large number of missing control variables. Adding housing type, construction year in 10-year intervals and other control variables related to housing characteristics significantly increases the R-square. Adding locational fixed effects in Model 3 further increases R-square as a result of capturing the geographical variation in housing prices. Results are found to be similar in model 4, which included a polynomial for the square footage of the residence, accounting for potential non-linear relationships between housing price and square footage. Values of the coefficients for energy labels remain similar to those in the results of Model 3. It can be seen in these first models that there does seem to be a relevant relation between housing characteristics (such as housing type) and the price effect of energy label, as the estimated price effects change as more housing characteristics are added to the models.

The final model adds interactions. All energy labels are significant in this model except B energy labels, being the only one not significant at a significance level of 5% or lower. Adding interaction effects reduces the interpretability of the base variants of the variables included if the interactions themselves are not taken into account. In Figure 2, the effect of energy label on WOZ-value for the different housing types that were included in the research is shown.

Energy efficiency, proxied by energy labels, affects the WOZ-value of residences in the Netherlands. This was determined using hedonic regression analysis, and the assessed WOZ-value of residences with an A energy label is estimated to be between 4.1% and 5.5% higher than that of residences with an energy label of E or worse.

When splitting the sample up or looking for interaction effects, flats and apartments are the only housing type showing indication of a potential interaction between housing type and energy label. This could be a result of the increased depreciation that flats and apartments tend to experience correlating with the poor energy labels of the residence. Other housing types are not assumed to have a significant interaction effect with energy labels.

These results indicate maintaining an acceptable energy efficiency level is more important to prevent prices from falling for flats and apartments, as the negative price effect that is associated with a decreasing energy label is more severe for flats and apartments than it is for other housing types.

Future research would benefit from a dataset on a larger scope, with a large sample of transaction prices of properties that were sold recently, and more detailed information regarding neighbourhood characteristics and maintenance level/depreciation of the relevant residences. Additionally, future research could look at the effect that price shocks in energy price have on the extent of capitalization and on the extent to which energy poverty is relevant in a society and who experience it.

6. REFERENCES

Alcott, H. & Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of economic perspectives*, 26(1), 3-28.

Aydin, E., Brounen, D. & Kok, N. (2020). The capitalization of energy efficiency: Evidence from the housing market. *Journal of Urban Economics*, 117, 103243.

Brounen, D. & Kok, N. (2011). On the economics of energy labels in the housing market. *Journal of Environmental Economics and Management, 62(2)*, 166-179.

Centraal Burea Statistiek (2013). *Twee derde van alle woningen eengezingswoning*. Retrieved on May 12, 2022 from https://www.cbs.nl/nl-nl/achtergrond/2013/04/twee-derde-van-alle-woningen-eengezinswoning#:~:text=De%20gemiddelde%20vloeroppervlakte%20van%20woningen,keer%20zo%20groot%20als%20appartementen.

Centraal Burea Statistiek (2015). *Woononderzoek Nederland.* Retrieved on December 1, 2021 from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/woononderzoek-nederland--woon--

Centraal Bureau Statistiek (2020). *WOZ-waarde gemiddeld 8.9 procent hoger*. Retrieved on May 12, 2022 from https://www.cbs.nl/nl-nl/nieuws/2020/37/woz-waarde-gemiddeld-8-9-procent-hoger.

Cerin, P., Hassel, L.G. & Semenova, N. (2014). Energy performance and housing prices. *Sustainable Development*, 22(6), 404-419.

Cozza, S., Chambers, J. & Patel, M.K. (2020). Measuring the thermal energy performance gap of labelled residential buildings in Switzerland. *Energy Policy*, 137, 111085.

Dinan, T.M. & Miranowski, J.A. (1989). Estimating the implicit price of energy efficiency improvements in the residential housing market - A hedonic approach. *Journal of Urban Economics*, 25(1), 52–67.

Eichholtz, P., Kok, N. & Quigley, J.M. (2010). Doing well by doing good? Green office buildings. *American Economic Review*, 100(5), 2492–2509.

European Commission, 2020. *About the energy label and ecodesign*. Retrieved on May 12, 2022 from https://ec.europa.eu/info/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/energy-label-and-ecodesign/about_en.

Fuerst, F., McAllister, P., Nanda, A & Wyatt, P (2015). Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. *Energy Economics*, 48, 145-156.

Gillingham, K. & Palmer, K. (2020). Bridging the Energy Efficiency Gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy*, 18-38.

Gilmer, R.W. (1989). Energy labels and economic search: an example from the residential real estate market. *Energy Economics*, 11(3), 213-218.

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

Kholodilin, K.A., Mense, A. & Michelsen, C. (2017). The market value of energy efficiency in buildings and the mode of tenure. *Urban Studies*, 54(14), 3218–3238.

Lyons, R.C., Lyons, S. & Stanley, S. (2016). The price effect of building energy ratings in the Dublin residential market. *Energy Efficiency*, 9(4), 875-885.

Nu.nl (2021). *Gasprijzen op recordhoogte: huishoudens gaan nog meer betalen*. Retrieved on December 1, 2021 from https://www.nu.nl/economie/6160742/gasprijzen-op-recordhoogte-huishoudens-gaan-nog-meer-betalen.html.

RFF (2020). *Energy Efficiency 101*. Retrieved on December 1, 2021 from https://www.rff.org/publications/explainers/energy-efficiency-101/#:~:text=Energy%20efficiency%20refers%20to%20using,to%2080%20percent%20less%20electricity.

Rijksdienst voor Ondernemend Nederland (2015). *Energielabel woningen*. Retrieved on May 12, 2022 from https://www.rvo.nl/onderwerpen/wetten-regels/energielabel-woningen.

Rijksoverheid (2021). *150 miljoen euro voor aanpak energiearmoede kwetsbare huishoudens*. Retrieved on December 1, 2021 from https://www.rijksoverheid.nl/actueel/nieuws/2021/10/15/150-miljoen-euro-voor-aanpak-energiearmoede-kwetsbare-huishoudens.

Rijksoverheid (2022). *Energielabel woning.* Retrieved on May 12, 2022 from https://www.rijksoverheid.nl/onderwerpen/energielabel-woningen-en-gebouwen/energielabel-woning.

Rooyakkers, E. (2021). *Trends: Energiearmoede*. Retrieved on December 1, 2021 from https://www.aedesmagazine.nl/edities/3-2021/artikelen/energiearmoede. Aedes Magazine.

Semple, S. & Jenkins, D. (2020). Variation of energy performance certificate assessments in the European Union. *Energy Policy*, 137, 111127.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics*. Northridge, Cal: Harper Collins.

Tardy, F. & Lee, B. (2019). Building related energy poverty in developed countries – Past, present, and future from a Canadian perspective. *Energy and Buildings*, 194, 46,61.

Zhang, L., Stangenberg, L. & van Wickeren, S. (2020). The information value of energy labels: Evidence from the Dutch residential housing market? *CPB discussion Paper*.

7. APPENDIXES

7.1 FULL REGRESSION RESULTS

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5
Energylabel Definitive = 2, B	-0.152***	-0.0186	-0.0244	-0.0227	-0.0917*
	(0.0357)	(0.0292)	(0.0246)	(0.0229)	(0.0507)
Energylabel Definitive = 3, C	-0.245***	-0.00544	-0.0263	-0.0183	-0.105**
	(0.0315)	(0.0303)	(0.0257)	(0.0239)	(0.0493)
Energylabel Definitive = 4, D	-0.371***	-0.0472	-0.0489*	-0.0441*	-0.176***
	(0.0330)	(0.0320)	(0.0272)	(0.0256)	(0.0503)
Energylabel Definitive = 5, E or lower	-0.399***	-0.0418	-0.0566**	-0.0492*	-0.206***
	(0.0337)	(0.0331)	(0.0280)	(0.0262)	(0.0505)
Housing Type= 2, Terraced	12.35***	0.154***	0.217***	0.173***	0.0310
	(0.0286)	(0.0251)	(0.0224)	(0.0183)	(0.0455)
Housing Type= 3, Semi-detached		0.242***	0.372***	0.302***	0.187***
		(0.0362)	(0.0329)	(0.0243)	(0.0582)
Housing Type= 4, Detached		0.509***	0.643***	0.535***	0.382***
		(0.0534)	(0.0515)	(0.0333)	(0.0677)
1.EnergyLabel#1.HousingType					0
					(0)
1.EnergyLabel#2.HousingType					0
					(0)
1.EnergyLabel#3.HousingType					0
					(0)
1.EnergyLabel#4.HousingType					0
					(0)
2.EnergyLabel#1.HousingType					0
					(0)
2.EnergyLabel#2.HousingType					0.0970*
					(0.0526)
2.EnergyLabel#3.HousingType					0.0138
					(0.0648)
2.EnergyLabel#4.HousingType					0.115
					(0.0769)
3.EnergyLabel#1.HousingType					0
					(0)
3.EnergyLabel#2.HousingType					0.0977**
					(0.0470)
3.EnergyLabel#3.HousingType					0.0947
					(0.0590)

3.EnergyLabel#4.HousingType					0.155**
					(0.0694)
4.EnergyLabel#1.HousingType					0
					(0)
4.EnergyLabel#2.HousingType					0.176***
					(0.0481)
4.EnergyLabel#3.HousingType					0.150**
					(0.0622)
4.EnergyLabel#4.HousingType					0.130
					(0.0847)
5.EnergyLabel#1.HousingType					0
					(0)
5.EnergyLabel#2.HousingType					0.206***
					(0.0486)
5.EnergyLabel#3.HousingType					0.202***
					(0.0641)
5.EnergyLabel#4.HousingType					0.225***
					(0.0799)
(4.1) Balcony= 1, Private Balcony		0.0547***	0.0310**	0.0155	0.0174
		(0.0166)	(0.0140)	(0.0126)	(0.0125)
(4.1) Balcony = 2, Shared Balcony		0.112*	0.112	0.0760	0.0766
		(0.0619)	(0.0683)	(0.0612)	(0.0637)
(Yes) Garage or carport = 1, Yes, garage		0.0984***	0.125***	0.107***	0.102***
		(0.0164)	(0.0144)	(0.0136)	(0.0135)
(4.7) Garage or carport = 2, Yes, carport		0.0645**	0.0646**	0.0459*	0.0470*
		(0.0311)	(0.0283)	(0.0277)	(0.0277)
(4.9) Number of rooms = 2		0.209***	0.203*	0.141	0.140
		(0.0522)	(0.113)	(0.107)	(0.109)
(4.9) Number of rooms = 3		0.336***	0.333***	0.228**	0.225**
		(0.0525)	(0.113)	(0.106)	(0.109)
(4.9) Number of rooms = 4		0.375***	0.364***	0.243**	0.243**
		(0.0537)	(0.113)	(0.107)	(0.109)
(4.9) Number of rooms = 5		0.438***	0.400***	0.255**	0.258**
		(0.0586)	(0.116)	(0.107)	(0.110)
(4.9) Number of rooms = 6		0.483***	0.450***	0.291***	0.294***
		(0.0643)	(0.118)	(0.109)	(0.111)
(4.9) Number of rooms = 7		0.513***	0.482***	0.305***	0.308***
		(0.0719)	(0.122)	(0.110)	(0.113)
(4.9) Number of rooms = 8		0.493***	0.479***	0.307***	0.308***
		(0.0775)	(0.123)	(0.113)	(0.115)

(4.9) Number of rooms = 9		0.432***	0.436**	0.238	0.246
		(0.161)	(0.171)	(0.152)	(0.151)
(4.9) Number of rooms = 10		0.621***	0.579***	0.441***	0.446***
		(0.166)	(0.170)	(0.169)	(0.164)
(4.9) Number of rooms = 12,		0.863***	0.887***	0.738***	0.718***
		(0.0981)	(0.208)	(0.193)	(0.189)
(4.9) Number of rooms = 16,		0.699***	0.769***	0.482***	0.473***
		(0.0921)	(0.141)	(0.115)	(0.118)
Surface Area (source: bag 2015)		0.00251***	0.00237***	0.00569***	0.00576***
		(0.000555)	(0.000556)	(0.000494)	(0.000481)
Surface Area Squared				-5.73e-06***	-5.86e-06***
				(1.04e-06)	(1.01e-06)
Built1971_1980		0.0545***	0.0686***	0.0386***	0.0400***
		(0.0160)	(0.0137)	(0.0114)	(0.0115)
Built1981_1990		0.0777***	0.0763***	0.0633***	0.0638***
		(0.0161)	(0.0134)	(0.0126)	(0.0127)
Built1991_2000		0.198***	0.183***	0.164***	0.166***
		(0.0227)	(0.0198)	(0.0183)	(0.0180)
Built2001_2010		0.283***	0.279***	0.257***	0.231***
		(0.0346)	(0.0309)	(0.0291)	(0.0303)
Built2011_2014		0.398***	0.344***	0.312***	0.273***
		(0.0785)	(0.0830)	(0.0763)	(0.0742)
COROP-region (40) = 2, Delfzijl en omgeving			-0.268***	-0.254***	-0.226**
			(0.0902)	(0.0776)	(0.107)
COROP-region (40) = 3, Overig Groningen			0.239***	0.227***	0.241***
			(0.0720)	(0.0610)	(0.0640)
COROP-region (40) = 4, Noord-Friesland			0.126*	0.131**	0.139**
			(0.0713)	(0.0626)	(0.0644)
COROP-region (40) = 5, Zuidwest-Friesland			0.238***	0.228***	0.232***
			(0.0695)	(0.0600)	(0.0640)
COROP-region (40) = 6, Zuidoost-Friesland			0.0714	0.0654	0.0919
			(0.0669)	(0.0584)	(0.0600)
COROP-region (40) = 7, Noord-Drenthe			0.243***	0.234***	0.243***
			(0.0664)	(0.0577)	(0.0598)
COROP-region (40) = 8, Zuidoost-Drenthe			0.0597	0.0710	0.0868
			(0.0711)	(0.0622)	(0.0641)
COROP-region (40) = 9, Zuidwest-Drenthe			0.275***	0.277***	0.301***
			(0.0654)	(0.0592)	(0.0617)
COROP-region (40) = 10, Noord-Overijssel			0.389***	0.386***	0.401***
			(0.0632)	(0.0538)	(0.0565)
COROP-region (40) = 11, Zuidwest-Overijssel			0.412***	0.399***	0.406***
			(0.0746)	(0.0631)	(0.0629)

COROP-region (40) = 12, Twente			0.237***	0.233***	0.247***
			(0.0635)	(0.0537)	(0.0558)
COROP-region (40) = 13, Veluwe			0.537***	0.529***	0.544***
			(0.0616)	(0.0518)	(0.0542)
COROP-region (40) = 14, Achterhoek			0.299***	0.292***	0.303***
			(0.0665)	(0.0564)	(0.0592)
COROP-region (40) = 15, Arnhem/Nijmegen			0.394***	0.378***	0.390***
			(0.0625)	(0.0516)	(0.0542)
COROP-region (40) = 16, Zuidwest-Gelderland			0.507***	0.499***	0.514***
			(0.0709)	(0.0611)	(0.0648)
COROP-region (40) = 17, Utrecht			0.624***	0.608***	0.622***
			(0.0639)	(0.0534)	(0.0558)
COROP-region (40) = 18, Kop van Noord-Holland			0.313***	0.311***	0.327***
			(0.0670)	(0.0575)	(0.0596)
COROP-region (40) = 19, Alkmaar en omgeving			0.556***	0.554***	0.562***
			(0.0699)	(0.0623)	(0.0647)
COROP-region (40) = 20, IJmond			0.590***	0.586***	0.604***
			(0.0686)	(0.0581)	(0.0597)
COROP-region (40) = 21, Agglomeratie Haarlem			0.808***	0.814***	0.835***
			(0.0797)	(0.0716)	(0.0744)
COROP-region (40) = 22, Zaanstreek			0.508***	0.499***	0.523***
			(0.0743)	(0.0660)	(0.0657)
COROP-region (40) = 23, Groot-Amsterdam			0.720***	0.727***	0.735***
			(0.0643)	(0.0543)	(0.0567)
COROP-region (40) = 24, Het Gooi en Vechtstreek			0.823***	0.800***	0.809***
			(0.0764)	(0.0649)	(0.0653)
COROP-region (40) = 25, Agglomeratie Leiden en Bollenstreek			0.751***	0.742***	0.757***
			(0.0642)	(0.0542)	(0.0567)
COROP-region (40) = 26, Agglomeratie 's-Gravenhage			0.587***	0.592***	0.612***
			(0.0629)	(0.0530)	(0.0556)
COROP-region (40) = 27, Delft en Westland			0.663***	0.657***	0.670***
			(0.0617)	(0.0513)	(0.0540)
COROP-region (40) = 28, Oost-Zuid-Holland			0.548***	0.550***	0.562***
			(0.0647)	(0.0545)	(0.0567)
COROP-region (40) = 29, Groot-Rijnmond			0.483***	0.477***	0.494***
			(0.0609)	(0.0506)	(0.0532)

COROP-region (40) = 30, Zuidoost-Zuid-Holland			0.432***	0.427***	0.446***
			(0.0667)	(0.0571)	(0.0588)
COROP-region (40) = 31, Zeeuwsch-Vlaanderen			0.112*	0.112**	0.132**
			(0.0646)	(0.0541)	(0.0573)
COROP-region (40) = 32, Overig Zeeland			0.336***	0.346***	0.356***
			(0.0613)	(0.0514)	(0.0544)
COROP-region (40) = 33, West-Noord-Brabant			0.401***	0.392***	0.405***
			(0.0604)	(0.0501)	(0.0528)
COROP-region (40) = 34, Midden-Noord-Brabant			0.476***	0.467***	0.482***
			(0.0806)	(0.0682)	(0.0694)
COROP-region (40) = 35, Noordoost-Noord-Brabant			0.466***	0.449***	0.463***
			(0.0619)	(0.0511)	(0.0537)
COROP-region (40) = 36, Zuidoost-Noord-Brabant			0.514***	0.495***	0.504***
			(0.0619)	(0.0514)	(0.0540)
COROP-region (40) = 37, Noord-Limburg			0.238***	0.225***	0.239***
			(0.0640)	(0.0539)	(0.0573)
COROP-region (40) = 38, Midden-Limburg			0.203**	0.202***	0.206***
			(0.0841)	(0.0725)	(0.0755)
COROP-region (40) = 39, Zuid-Limburg			0.236***	0.219***	0.233***
			(0.0688)	(0.0581)	(0.0601)
COROP-region (40) = 40, Flevoland			0.265***	0.248***	0.266***
			(0.0635)	(0.0536)	(0.0562)
Constant	**12.35*** **	**11.13*** **	**10.66*** **	**10.57*** **	**10.66*** **
	(0.0286)	**(0.0598)**	**(0.129)**	**(0.121)**	**(0.131)**
Observations	**2,534**	**2,534**	**2,534**	**2,534**	**2,534**
R-squared	**0.080**	**0.595**	**0.738**	**0.764**	**0.769**

*Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1)*

7.2 STATA DO-FILE

```
//Open Dataset
cd "C:\Users\daans\Documents\University\Master\Master's Thesis\Data\1637254356256-
WoON2015__release_1.0_-_W"
use WoOn2015_e_1.1.dta

//Data Cleaning
drop if missing(WOZwaarde)
drop if missing(Energieklasse)
drop if missing(SrtWon)
drop if missing(gebruiksopp)

drop if EigHuurA==2
drop if BestndWon==2

drop if Kamers>20
drop if bjaarbag<1900
drop if WOZwaarde>1000000
drop if JrGekocht < bjaarbag

winsor WOZwaarde,p(.01) gen(WOZwaardeWin)

gen LogAankPrs = ln(AankPrs)
gen LogVerkwaar = ln(Verkwaar)
gen LogWOZwaarde = ln(WOZwaardeWin)

gen BuildingAge = 2014 - bjaarbag
gen BuildingAgeSquared = BuildingAge * BuildingAge
gen gebruiksoppSquared = gebruiksopp * gebruiksopp
```

```
gen Built1900_1910 = (bjaarbag < 1911)
gen Built1911_1920 = (bjaarbag > 1910 & bjaarbag <1921)
gen Built1921_1930 = (bjaarbag > 1920 & bjaarbag <1931)
gen Built1931_1940 = (bjaarbag > 1930 & bjaarbag <1941)
gen Built1941_1950 = (bjaarbag > 1940 & bjaarbag <1951)
gen Built1951_1960 = (bjaarbag > 1950 & bjaarbag <1961)
gen Built1961_1970 = (bjaarbag > 1960 & bjaarbag <1971)
gen Built1971_1980 = (bjaarbag > 1970 & bjaarbag <1981)
gen Built1981_1990 = (bjaarbag > 1980 & bjaarbag <1991)
gen Built1991_2000 = (bjaarbag > 1990 & bjaarbag <2001)
gen Built2001_2010 = (bjaarbag > 2000 & bjaarbag <2011)
gen Built2011_2014 = (bjaarbag > 2010)

gen AgeWhenSold = JrGekocht - bjaarbag

gen UsePerMeter = totener / gebruiksopp

//Removing dated buildings to reduce bias, and recoding variable with low sample size categories
drop if bjaarbag<1950
recode Energieklasse (5/7 = 5), gen(NewEnergieklasse)
recode SrtWon (5/8 = 5), gen(NewSrtWon)
recode BalkDakt (2/3 = 2) (4 = 3), gen(NewBalkDakt)

//Labelling newly created variables
label variable NewEnergieklasse "Energielabel definitied (bron:rvo) Nieuw"
label variable NewSrtWon "(3.5) Type Woning Functioneel Nieuw"
label variable NewBalkDakt "(4.1) Balkon of dakterras"

label define energylabels 1 "A" 2 "B" 3 "C" 4 "D" 5 "E or lower"
label values NewEnergieklasse energylabels

label define housingtypes 1 "Flat or Apartment" 2 "Terraced" 3 "Semi-detached" 4 "Detached" 5
"Other"
label values NewSrtWon housingtypes

label define balconytypes 1 "Private Balcony" 2 "Shared Balcony" 3 "No Balcony"
label values NewBalkDakt balconytypes
```

```
//Dropping "Other" residential category
drop if NewSrtWon==5

//Regression Models
//Model 1 - Base Model
reg LogWOZwaarde ib1.NewEnergieklasse, vce(robust) allbaselevels

//Model 2 - Adding Building Characteristics
reg LogWOZwaarde ib1.NewEnergieklasse ib1.NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010 Built2011_2014,
vce(robust) allbaselevels

//Model 3 - Adding location fixed effects
reg LogWOZwaarde ib1.NewEnergieklasse ib1.NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010 Built2011_2014
i.corop, vce(robust) allbaselevels

//Model 4 - Adding Polynomial
reg LogWOZwaarde ib1.NewEnergieklasse ib1.NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop, vce(robust) allbaselevels

//Model 5 - adding interactions
reg LogWOZwaarde NewEnergieklasse##NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop, vce(robust) allbaselevels
```

```
//Seperate regression models
reg LogWOZwaarde ib1.NewEnergieklasse ib3.NewBalkDakt ib3.GarCarp i.Kamers gebruiksopp
gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop if NewSrtWon==1, vce(robust) allbaselevels

reg LogWOZwaarde ib1.NewEnergieklasse ib3.NewBalkDakt ib3.GarCarp i.Kamers gebruiksopp
gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop if NewSrtWon==2, vce(robust) allbaselevels

reg LogWOZwaarde ib1.NewEnergieklasse ib3.NewBalkDakt ib3.GarCarp i.Kamers gebruiksopp
gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop if NewSrtWon==3, vce(robust) allbaselevels

reg LogWOZwaarde ib1.NewEnergieklasse ib3.NewBalkDakt ib3.GarCarp i.Kamers gebruiksopp
gebruiksoppSquared Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010
Built2011_2014 i.corop if NewSrtWon==4, vce(robust) allbaselevels

//Checking assumptions
regcheck

//Export Results
outreg2 using RegressionMay.doc, replace ctitle(Model 5)

//Model using Energy Use instead of Energy Labels
reg LogWOZwaarde UsePerMeter ib1.NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp Built1971_1980 Built1981_1990 Built1991_2000 Built2001_2010 Built2011_2014
i.corop

//Model using Year Built Dummies
reg LogWOZwaarde ib1.NewEnergieklasse ib1.NewSrtWon ib3.NewBalkDakt ib3.GarCarp i.Kamers
gebruiksopp i.bjaarbag i.corop

//Descriptive Statistics
//Ratio variables
summarize WOZwaarde LogWOZwaarde totener UsePerMeter Kamers gebruiksopp bjaarbag
```

7.3 CORRELATION MATRIX OF MAIN VARIABLES

	WOZ	Label	Housing	Balcony	Garage	Rooms	Footage	COROP
WOZ Value	1							
Energy Label	-0.275	1						
Housing Type	0.625	-0.129	1					
Balcony	0.172	-0.110	0.409	1				
Garage	-0.471	0.138	-0.508	-0.089	1			
Rooms	0.507	-0.092	0.524	0.246	-0.321	1		
Square Footage	0.643	-0.185	0.563	0.137	-0.429	0.506	1	
COROP Region	0.012	0.023	-0.017	0.003	0.014	-0.009	-0.001	1

7.4 TOLERANCE/VIF

Variable	VIF	1/VIF
Energy Label		
B	3.07	0.326221
C	5.56	0.179995
D	4.99	0.200433
E	4.81	0.207696
Housing Type		
Terraced	3.85	0.260055
Semi-detached	2.84	0.352423
Detached	3.09	0.324099
Balcony		
Private Balcony	1.85	0.539367
Shared Balcony	1.10	0.907789
Garage/Carport		
Garage	1.66	0.600601
Carport	1.10	0.909676
Number of Rooms		
2	26.11	0.038300
3	73.79	0.013552
4	114.30	0.008749
5	108.20	0.009242
6	61.09	0.016370

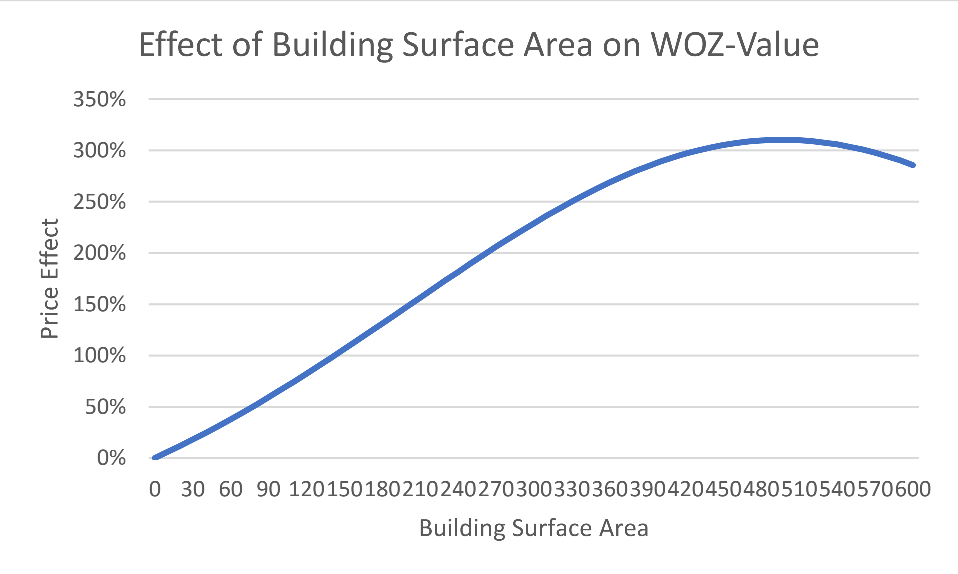
7	24.93	0.040108
8	8.41	0.118967
9	2.53	0.394928
10	1.86	0.537481
12	1.46	0.686284
16	1.24	0.808472
Surface Area	7.64	0.130890
Surface Area Squared	4.84	0.206800
Built19~1980	1.46	0.685314
Built19~1990	1.53	0.653212
Built19~2000	1.63	0.612425
Built20~2010	1.72	0.580189
Built20~2014	1.10	0.909878
COROP-region		
2	1.30	0.771637
3	6.28	0.159153
4	4.70	0.212688
5	2.30	0.435451
6	3.87	0.258574
7	4.28	0.233577
8	3.87	0.258321
9	3.58	0.279550
10	10.30	0.097119
11	3.29	0.303824
12	10.42	0.095986
13	14.77	0.067696
14	6.68	0.149779
15	19.59	0.051035
16	4.75	0.210697
17	22.39	0.044668
18	5.98	0.167298
19	6.26	0.159793
20	3.03	0.330305
21	3.18	0.314593
22	2.88	0.346995
23	15.10	0.066214

24	3.72	0.268752
25	10.86	0.092120
26	27.90	0.035837
27	12.08	0.082804
28	12.36	0.080938
29	32.49	0.030778
30	10.31	0.097008
31	5.55	0.180104
32	16.54	0.060448
33	27.65	0.036166
34	5.15	0.194185
35	13.92	0.071833
36	14.48	0.069039
37	6.01	0.166304
38	4.58	0.218322
39	11.53	0.086748
40	14.12	0.070819
Mean VIF	12.73	

7.5 REGRESSION ASSUMPTION TESTS

Assumption	Test	Model 2	Model 2	Model 3	Model 4	Model 5
Heteroskedasticity Problem	**Breusch-Pagan hettest**	Chi2: 0.26 P: 0.62	Chi2: 165.67 P: 0.00	Chi2: 230.48 P: 0.00	Chi2: 76.49 P: 0.00	Chi2: 73.65 P: 0.00
Multicollinearity Problem	**VIF**	VIF < 5	VIF < 5	VIF < 5	VIF < 5	VIF > 5
Non-normally distributed residuals	**Shapiro-Wilk W**	Z: 7.77 P: 0.00	Z: 8.41 P: 0.00	Z: 11.52 P:0.00	Z: 10.75 P: 0.00	Z: 10.73 P: 0.00
Specification problem	**Linktest**	T: -0.00 P: 1.000	T: -8.58 P:0.00	T: -7.10 P: 0.00	T: -3.49 P: 0.00	T: -2.478 P: 0.013
Functional Form Problem	**Reset test**	X	F:68.99 P: 0.00	F: 73.44 P: 0.00	F: 17.34 P: 0.00	F:14.57 P: 0.00
Influential Observations	**Cook's Distance**	X	Not above cutoff	Not above cutoff	Not above cutoff	Not above cutoff

7.6 NON-LINEAR EFFECT OF BUILDING SURFACE AREA ON WOZ-VALUE



7.7 SAMPLE SIZE OF ALL ENERGY LABEL AND HOUSING TYPE COMBINATIONS

	Flat or Apartment	Terraced	Semi-detached	Detached	**Total**
A	54	80	32	31	**197**
B	84	168	59	50	**361**
C	183	470	115	85	**853**
D	187	312	46	51	**596**
E or lower	197	230	58	42	**527**
Total	**705**	**1260**	**310**	**259**	**2534**