HAS THE VALUE OF RESIDENTIAL ENERGY LABELS INCREASED?

A QUANTITATIVE APPROACH TO THE CHANGE IN VALUE OF ENERGY LABELS BETWEEN 2008-2018 FOR DWELLINGS IN THE PROVINCE OF NOORD-HOLLAND, THE NETHERLANDS

Maurits Cassee
January 2019

university of groningen / faculty of spatial sciences

COLOPHON

DOCUMENT:	*Master thesis – Real Estate Studies*
TITLE:	*Has the value of residential energy labels increased?* A quantitative approach to the change in value of energy labels for dwellings in the province of Noord-Holland, The Netherlands
VERSION:	*Final*
AUTHOR:	*Maurits Cassee*
STUDENT NUMBER:	*S2541130*
E-MAIL:	*mauritscassee@gmail.com*
DATE:	*31 January 2019*
WORD COUNT:	*19021 (including appendices and references)*
ILLUSTRATION FRONT PAGE	*European Commission*
SUPERVISORS:	*Prof. Dr. E.F. Nozeman* *Dr. M. van Duijn*

ABSTRACT

This thesis studies the Dutch energy labels for dwellings. An overview is given about the origin of the energy labels for dwellings. The main research question studies the development of energy label value over the past decade. This development is studied for the province of Noord-Holland and its underlying COROP-regions.

This study is based on transaction data provided by the Dutch Association of Realtors and the energy label registrations provided by the Netherlands Enterprise Agency. The research method that is used is the hedonic model that is rooted in econometric modelling.

The results of the analysis suggest that the different types of energy label classes (A-G) attribute different value to dwellings. When looking at the developments over the years the results show that the different energy label classes develop differently. The overall valuation of the energy labels approximately follows the real estate cycle, it valuation fluctuates over time and increased over the period 2008-2018.

The valuation of the energy labels differs greatly among the different COROP-regions. In regions with lower average transaction prices the valuation of energy labels is stronger than in regions where the average transaction price is higher.

Since the energy label policy enforcement of 2015, the valuation of energy labels increased significantly although this might be caused by the increased housing prices in general as there is a structural break between the two subsamples.

Keywords: Hedonic modelling, energy label dwelling, Dutch housing market, 2008-2018

PREFACE

My name is Maurits Cassee and this is my master thesis “Has the value of residential energy labels increased?”. This thesis represents the pièce the résistance of my academic career that started with a bachelor degree in Business Administration, complemented with a propaedeutic degree in Dutch law at the University of Groningen. After completing these degrees, I started the master Real Estate Studies at the University of Groningen in early 2017. The completion of this thesis marks the end of a fantastic time as a student in Groningen and the start of my professional career in real estate. I am looking forward to the exciting times that lie ahead.

I would like to thank my supervisors prof. dr. Ed Nozeman and dr. Mark van Duijn for their constructive feedback during the process of writing this thesis. I also would like to thank the Dutch Association of Realtors (NVM) and the Netherlands Enterprise Agency (RVO) for their willingness to share their precious data for the purpose of this study.

Last but not least, I would like to thank my fellow students, friends, parents, family and others who were willing to share their time and knowledge for supporting me during this writing process.

Maurits Cassee
January 2019

TABLE OF CONTENTS

1. INTRODUCTION

Motivation

Global interest increased in mitigation of climate change. This led to the investigation of the sources of energy consumption. The built environment accounts for a substantial part of the global energy consumption. This creates opportunities for substantial reduction of greenhouse emissions (Rooijers, 2015). The European energy labels are established for appliances, cars, dwellings and offices as an objective instrument to indicate the energy performance. So far, the energy label for dwellings is used as an indicator of energy performance only.

However, offices need to have an energy label that is 'green' by 2023 in The Netherlands. The energy labels range from A to G whereby A to C are considered green. Offices that do not have a green label by 2023 are not allowed to be rented out until it is retrofitted to have a green label (Netherlands Enterprise Agency, 2016). For dwellings there is a policy that all new dwellings are all-electric, meaning that there is no natural gas used in the building. Therefore, the transition to a renewable future seems to happen increasingly rapid and only time will tell when existing dwellings are required to obtain a green label.

To be eligible for a green label, dwellings need to be well insulated. Measures that help to obtain a green label include double (or triple) paned windows, floor-, wall- and roof insulation. Furthermore, the central heating system is required to be hybrid (natural gas combined with electricity) or full-electric meaning that it requires less or no natural gas. These upgrades require a large investment up front. Home owners will eventually be able to re-earn this in a lower energy bill. However, when selling the dwelling prior to the break-even point, home-owners are uncertain about their return on investment.

Therefore, it is important for home-owners to know the value of energy labels. When the current economic value is known, home-owners can make an informed decision whether to invest in the energy performance of their dwelling. Currently, no literature is available on the development of energy label value for dwellings in the Netherlands. Now that there is increased attention for energy labels in the recent Climate Agreement and the adoption rate of energy labels has increased, the value of energy labels might have changed as buyers only attribute value to something that they know holds value.

As house prices differ greatly among different regions, it is interesting to see whether the same is true for the valuation of energy labels. There is a gap in the literature about this topic. This study will tap into this gap to examine the regional differences in energy label valuation.

This study taps into the gap by examining the development of energy label value over the period of 2008-2018.

Literature review

Earlier research on the value of energy labels focused either on offices or dwellings. In the Netherlands, offices have been studied. The results of the study by Kok and Jennen (2012) show that significant value is attributed to the energy label. More recent research on energy label value for dwellings took place in the United Kingdom where the same Energy Performance Coefficient (EPC) system is in place. In that study significant added value of energy labels was found (Fuerst et al., 2015). Another dwelling based study in Spain yields the same results although a different measurement of energy efficiency was used (De Ayala et al., 2016). In the United States, where yet another definition of energy efficiency is used a higher energy efficiency score adds value as well. Even in times of crises a premium is paid for energy efficient dwellings although the premium is smaller than in times of economic growth (Kahn and Kok, 2014). Based on previous published literature energy labels hold a certain value. How the value of the energy labels developed over the recent years has not yet been studied publicly before.

Wong (2002) describes a model of the household housing decision-making process. In the model multiple factors determine the ultimate choice of the household. This means that housing characteristics do not entirely determine the household housing decision. Energy labels are an indicator of certain housing characteristics: e.g. thermal comfort because of insulation and a lower energy bill because of solar panels. For the Dutch housing market the estimation is that half of the price of dwellings is determined by the dwelling attributes and characteristics of its surroundings, including location (Boelhouwer, 2000). The other half of the price of dwellings is determined by more economic factors such as financing-, supply- and socioeconomic factors, including policies and regulations (Galati et al., 2011). It is important to consider the housing transaction in a holistic perspective, taking the dwellings characteristics and transactional conditions into account. Previous studies focused mainly on either these elements, housing characteristics or economic conditions. This study combines these two perspectives into one holistic approach.

Brounen (2018) is the first to take both location and energy labels into account. The study ends with a preliminary conclusion that in the largest cities of The Netherlands, the energy labels seem to hold less value than in the surrounding areas. However, the press statement released by Brounen is based on data of his previous study of 2011 and is not a full research article or study. No further details of these findings have been published by Brounen until now.

Literature on the development of value of energy labels is scarce, especially for the Netherlands. Brounen and Kok (2011) studied the development of the adoption of the energy label, but not the development of its value. The study shows that the energy labels hold value. However, the study does not take into account the recent developments of the Dutch housing market. During the studied period the housing market was in a decline, possibly resulting in a different valuation of the energy labels. Furthermore, at the time of the study by Brounen and Kok (2011) the adoption rate of energy labels was lower than it is now and more than half of the buyers was unaware of the energy label of the dwelling that they bought (Braanker, 2015; Brounen and Kok, 2011). From 2015 on, sellers of dwellings are required to provide an energy label during the transaction, this makes it possible to study the development of the energy label value.

Aim

The motivation shows that no public studies are known showing the development of energy label value in the Netherlands. The research aim of this study is to determine whether the value of energy labels changes over time. After the introduction of energy labels in the Netherlands the housing market recovered from a crisis to a new all-time high (source, Figure 2). The impact of this development on the valuation of energy labels is not yet known. Furthermore, policies changed in in the energy sector. It is interesting to see if this is reflected in the value of energy labels.

The data that will be used for this research are the transaction data provided by the NVM, a branch organization of real estate agents and appraisers in The Netherlands. To establish the energy labels of the dwellings that are taken into account data provided by RVO will be used. RVO is the Netherlands Enterprise Agency and responsible for the registration of dwellings' energy labels.
Because the supply of data is limited to a single province of choice, a selection of a suitable province is needed. For a province to be suitable it preferably has regions with different housing market trends. Figure 2 shows that the province of Noord-Holland meets this criterion.

The central research question of this study is:

How did the value of residential energy labels develop between 2008-2018 in the province of Noord-Holland, the Netherlands?

To answer this main research question, three sub-questions are answered.

1. *What is the impact of an energy label on the value of residential properties?*

Based on a literature study the added value of energy labels for dwellings is determined. The studies laid out in the literature review will be discussed thoroughly to answer this sub-question. Subsequently, empirically tested whether the findings of the literature study apply to the data examined.

2. *What are the developments in economic value of energy labels in the period 2008-2018 in the province of Noord-Holland?*

This question is empirically answered by using a hedonic price model where the energy label classes are valued. By means of interactions the development in the period 2008-2018 is established. In this analysis different geographical scales are taken into account.

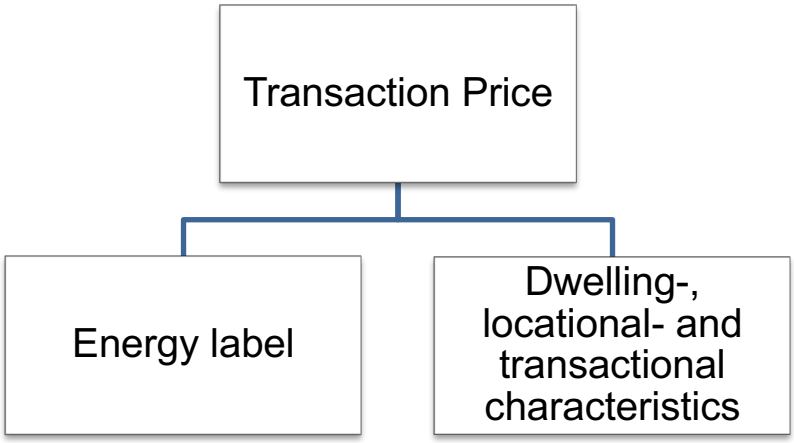
3. *What impact do relevant policies have on the value of energy labels?*

This question is answered by first establishing the relevant policies that affect the energy label policy. Subsequently the moments in time where the policies are made public and become effective are marked. The periods before and after the respective changes are tested for significant changes.

Approach

To determine the current value of energy labels a literature study will be conducted to determine the relevant influencing factors. These factors are included in a hedonic price model that is used to estimate the value of the energy label categories.
The empirical part of the study will be conducted using the transaction data of NVM in conjunction with the energy label information of RVO. These two datasets will be combined to estimate the value of energy labels using a hedonic model. The hedonic model will be estimated over multiple years so see if and when changes in valuation of energy labels occurs.

Figure 1: Conceptual Model



Theoretical Framework

Figure 1 shows the conceptual model which indicates the three types of variables that are included in the hedonic models. The energy label is the (independent) X-variable, the transaction price is the (dependent) Y-variable. The dwelling characteristics, locational- and transactional characteristics are included to control for differences across locations and time of transactions, these are the (control) Z-variables.

The remainder of this paper is organized as follows. Section 2 describes the background of energy label policies in the Netherlands and section 3 the conceptual model including hypotheses. Section 4 describes the empirical approach. Section 5 describes the data and the exploratory analysis. Section 6 presents the results, and section 7 concludes.

2. ENERGY LABEL POLICIES & HOUSING MARKET DEVELOPMENTS

In Europe the Energy Performance of Buildings Directive (EPBD) of 2002 marked the start of the development of Energy Performance Certificates (EPC) by the different countries. In the EPC each dwelling is rated according to a predefined list of performance indicators resulting in an energy label ranging from A to G.

Energy labels in The Netherlands, a brief history

In the Netherlands the EPC became compulsory in 2008 for home-owners wanting to sell or rent-out their dwelling. Although compulsory, commitment to providing the EPC when selling a dwelling was low as there was no enforcement. In 2010 the EPBD was revised resulting in simplified energy labels, ranging from A to G. In 2012 the Dutch Parliament refused to pass a law that enabled enforcement of the registration of energy labels during transactions although the obligation to provide an energy label remained. After the Energy Agreement of 2013 the Dutch Parliament agreed to a law that ensured the application of preliminary labels for all dwellings by the government as well as enforcement of supplying a final energy label during a transaction by the seller, starting from 2015 (Dutch Senate, 2008).

The preliminary and final labels are a simplified label version of the energy-index (EI) which includes over 150 dwelling characteristics. To obtain an EI for a dwelling an expert is needed that conducts a full calculation of the dwelling. This method is required for housing cooperatives that provide public housing whereas the energy-label is aimed at private home-owners. The energy label is based on a set of 10 characteristics of the dwelling that home-owners can supply after which an expert checks the provided evidence. Table 1 summarizes the type different types of obtaining a EPC. Home-owners are free to opt for the more extensive EI system yet are not obliged to do so (Netherlands Enterprise Agency, 2014).

Table 1 Type of Energy Performance Certificates in The Netherlands (source: Netherlands Enterprise Agency)

Type of energy performance certificate	Energy index (EI)	Energy label
Established by	Expert only	Home-owner, supervision by expert
Included characteristics	150	10
Obligatory for	Housing corporations	Private home-owners

Current situation of energy labels

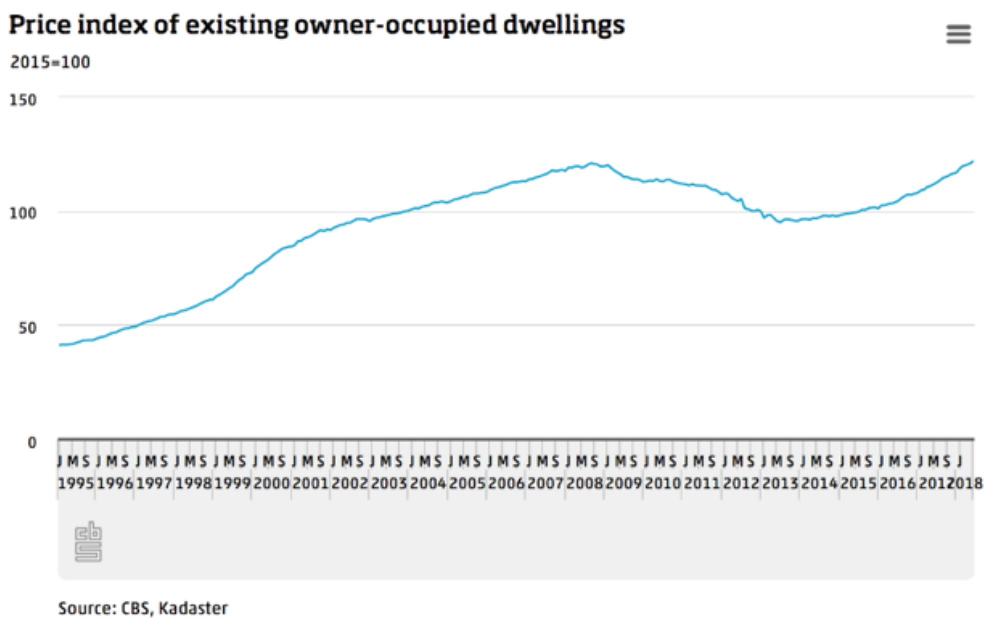
The introduction of preliminary energy labels and the enforcement of providing final labels increased the transparency of dwellings' energy efficiency, enabling home-buyers to support their purchase decision with this information. However, until 2015 more than half of the home-buyers was not aware of the energy label that they purchased and for 87% of the buyers the EPC had not played a role in their negotiations (Braanker, 2015).

Since 2015 the Netherlands has seen a transition in perspective on energy reduction. The first government coalition that was formed after the introduction of the preliminary EPC for all dwellings, in 2017, aimed at a 49% reduction in energy consumption of the built environment before 2030 (Government of The Netherlands, 2017a). The year 2018 started with an announcement of the Minister of Economic Affairs and Climate that the use of natural gas in Groningen should be reduced 8 years sooner than agreed in the coalition agreement (NOS.nl, 2018). A mere two months later, in March 2018 the Dutch government announced that it would stop the production of natural gas in Groningen completely (Government of The Netherlands, 2018). This increased attention for energy performance and sustainability by the Dutch Government has led to increased awareness of the energy label (Government of The Netherlands, 2017b). Whether this accounts to a higher valuation of energy labels is what this study aims to answer.

Recent developments in the Dutch housing market

During the financial crisis starting from 2008 the Dutch housing market experienced the first downward trend over 20 years. Prices declined to a level of a few years before resulting in home-owners with mortgages higher than the value of their homes. In the decade after the financial crisis of 2008 the Dutch housing market has recovered, reaching an all-time high in prices of owner-occupied dwellings in May 2018 as Figure 2 shows (Statistics Netherlands and Land Registry Office, 2018).

Figure 2 Price index of owner-occupied dwellings in The Netherlands



The increase in the price index of dwellings is caused by a growing economy combined with low interest rate and limited supply on the market (Bokeloh, 2018). Potential buyers have difficulty finding a suitable dwelling because of this shortage. However, the impact of the shortage differs across regions in the Netherlands (Aalders et al., 2018). Therefore, the price levels across regions differ greatly.

Figure 3 Supply of dwellings per COROP-region in the province of Noord-Holland (Source: CBS)

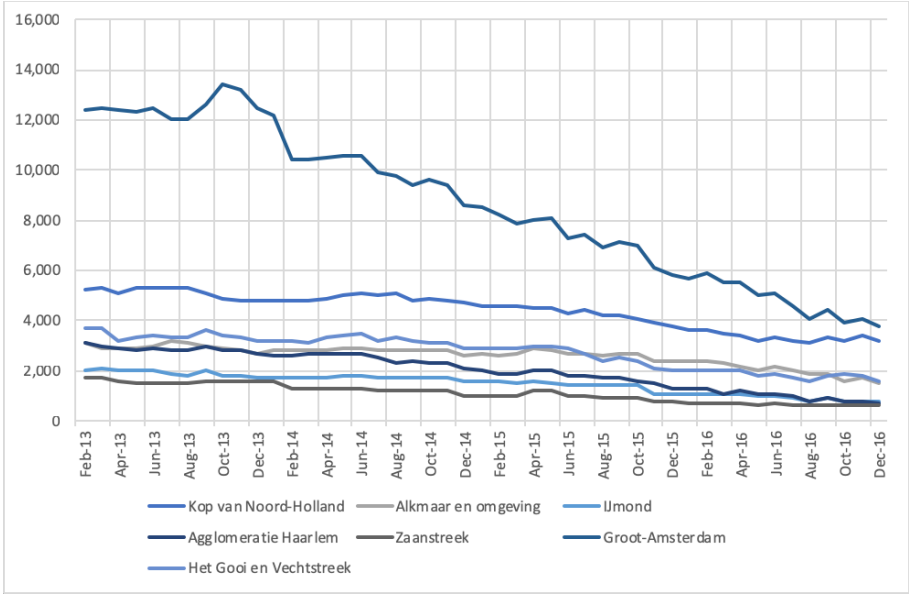


Figure 3 shows the supply of dwellings in the 7 COROP regions of the Noord-Holland province. In Greater Amsterdam the trend of decline in supply is fierce compared to the other regions indicating differences between the regions in the province and large differences over time.

Since 2011 the Dutch housing market and energy label regulation have seen major developments that might impact the valuation of energy labels. As the figure above shows, 2011 was in the midst of the Dutch housing crisis. At that time the energy label was not obligatory resulting in fewer observations. The economic tide and the different supply/demand levels within Noord-Holland make is worthwhile to pay scientific attention to the impact of energy label policy in relation to dwelling price dynamics.

3. THEORY

Establishment of a transaction

In an efficient market with sufficient supply and demand, where there is perfect information and goods are homogenous, suppliers can only compete on the price of their goods. In such a market, buyers will always pick the good with the lowest price since the goods are homogenous. Eventually the price of the good will equal the production cost of the good as it is the lowest price a supplier can offer the good (Evans, 2004).

The housing market functions differently as dwellings are not homogenous. Dwellings are a composite of varying characteristics such as size, location, type and age and are therefore considered heterogenous. The supply and demand of dwellings can vary greatly over time and varies across location which can sometimes make it difficult for potential-buyers to value dwellings based on previous transactions (Evans, 2004).

DiPasquale and Wheaton (1994) describe the process of the establishment of a market equilibrium where supply meets demand. An increase in the number of households leads to a higher demand for dwellings. When the costs for buying decrease, for example because of a lower interest rate or lower transfer tax, buying a dwelling becomes more attractive for households which leads to a higher demand. The supply is formed by all owner-occupied dwellings. When more dwellings are constructed than demolished, the stock of dwelling increases.
The real estate market does not respond directly to a change in supply or demand as the construction or demolishment of dwellings is time-consuming and restricted by legal procedures. This increases the time needed to adjust to the new market equilibrium.

In the neo-classical consumption theory, the effect of utility maximization is described. Buyers are seeking for a good that best suits their needs at the lowest possible cost. Sellers aim to maximize their profit and therefore want a reasonable price for the good they sell. Both parties are willing to agree to a price that satisfies both utilities (Rosen, 1974).

Regional housing markets

Previous literature indicates that housing markets function on different regional scales. Some factors such as nationwide policy changes or macro-economic developments affect all regions within a country. In these cases, the same trend applies to all regions. However, some local differences affect the regional housing markets differently which creates different trends. By approaching housing markets on a regional scale it is possible to get a more detailed insight in the value of dwelling characteristics (Renes et al., 2006). Eskinasi (2011) refined this idea by describing a model that supplements the housing market model by DiPasquale and Wheaton (1994) by adding the focus for regional housing markets. This study will take the regional housing markets into account by using the widely used Dutch COROP-regions which divide the province of Noord-Holland in seven regions.

Determinants of dwelling value

Dwelling value is determined by the characteristics of the dwelling itself and external factors.

Characteristics of the dwelling

As dwellings are differentiated products the housing market is heterogenous. In this heterogenous market competition takes place on characteristics in combination with price whereas homogenous markets compete on price alone. The value of a dwelling is determined by the price that buyers are willing to pay to obtain a specific dwelling. As dwellings are heterogenous goods composed of different characteristics, a transaction price is considered to be the sum of the value of the different characteristics combined. A dwelling value is thus considered to be the same as its transaction price. Hedonic modelling can attribute economic value to the individual characteristics (Rosen, 1974).
In a meta-study that studied 125 hedonic models, the variables in real estate hedonic models that have the most significance in determining the dwelling value are age, floor size and the number of rooms. For the floor size and the number of rooms the log-linear function is oftentimes also included (Sirmans et al., 2005).

Locational characteristics

Apart from the dwelling characteristics, the characteristics of its surroundings are also a determinant of its value. The amenity-based theory is grounded on the notion that amenities in the surrounding of the property are part of the utility maximization equation. The characteristics of the surroundings of the dwelling and the amenities it provides affect dwelling value (Brueckner et al., 1999). The meta-study by Boyle and Kiel (2001) focuses on common included externalities in hedonic real estate modelling. Their results show that air quality, water quality, undesirable land use and neighborhood variables are among the most included externalities that affect dwelling value. Lazrak (2014) describes population density and the share of non-Western immigrants as characteristics that affect dwelling value.

For the Netherlands, many studies have focused on the effect of these types of externalities.
Van Dam and Visser (2006) summarize which externalities affect dwelling value. Furthermore, Van Dam and Visser categorize these externalities in three different types of environmental characteristics that are relevant for the value of dwellings:

Functional characteristics

Functional characteristics include the proximity to different types of amenities. The accessibility of public transportation has a positive impact on the value of dwellings, the closer the public transportation station the higher the value, for tram and metro transportation this effect is even stronger than for trains (van Dam and Visser, 2006). The proximity to the center of a neighborhood or city as a whole has a positive effect on the value of dwellings too (van Dam and Visser, 2006).

Financial economic characteristics

There are multiple financial economic characteristics regarding dwellings that affect its value. Van Dam and Visser (2006) describe the effect of social status of a neighborhood. In their study the share of owner-occupied dwellings in a

neighborhood seems to affect dwelling value. A closer look at the results shows that it is not the share of owner-occupied dwellings that counts but the socio-economic status of the neighbors. The education level and income of neighbors has a larger effect on dwelling value.

Physical characteristics

Physical characteristics indicate the physical presence of certain attributes. These include the proximity to parks or water. The proximity to natural space has a significant positive impact on the value of dwellings (Daams et al., 2016).
The effect of redevelopment of historical industrial sites can be considered a positive neighborhood characteristic upon completion. Before completion the use of land as industrial site close to residential areas is considered an undesired land use. Before completion the redevelopment shows a negative impact on the value of dwellings, after completion this is converted to a positive impact (Van Duijn et al., 2016). The investment in historic amenities in the surroundings of residential areas positively affects dwelling value (Koster and Rouwendal, 2017; Lazrak et al., 2014).

Energy label as dwelling characteristic

The establishment of energy labels adds a characteristic to dwellings. Each increment in energy label represents an investment of some sort in the energy performance of the dwelling. Although energy labels for dwellings are a relatively new research field, some relevant literature is available.

Shortly after the introduction of the compulsory energy label for dwellings the Dutch housing market was studied. In the limited available sample, the energy label shows significant attributed value to the dwelling (Brounen and Kok, 2011). When looking at housing markets abroad both California (Kahn and Kok, 2014) and Spain (De Ayala et al., 2016) show significant impact of better energy labels on dwelling value.

In the Dutch office market, the same results are presented. Offices with better energy labels yield higher rents than equivalent offices with lower energy performance (Kok and Jennen, 2012).

Most studies show that energy labels account for a certain value of the dwelling. However, current research does not account for possible trends in the development of energy label value. With the recent increase in indexed dwelling prices in the Netherlands it is interesting to see whether the energy label value develops in a similar way.

Trend effect

Even if all characteristics and externalities of the dwelling are known and incorporated in the hedonic model, it is not possible to entirely determine the transaction price of the dwelling. The housing market does not function fully rational which creates different developments in times of crises or booms. In a sellers' market, prices increase quickly and a bubble may arise as transactions are established based on future expectations (Case and Shiller, 1988).
Even without booms or busts, dwelling prices can fluctuate over time, the moment of transaction is therefore relevant. Figure 2 shows the development of dwelling value over recent years. The time of transaction is often included as a control variable in

the hedonic model (Brooks and Tsolacos, 2010). Previous literature on energy label value focused on the value on a specific moment in time. This study takes the development of value over time into account.

In the Dutch housing market, the factors that influence the value development of dwellings are described to be demographic developments, regulatory changes, dynamics of the market as well as the factors included in the model by DiPasquale and Wheaton such as developments in the supply of dwellings (Boelhouwer et al., 1996). Another effect that time can have on capital is depreciation. Although rarely studied in the field of dwellings, some research points to capital depreciation of the dwelling offsetting accrued capital gain (Harding et al., 2007). When considering an energy label class as representation of a previous investment in the dwelling the same effect might be true for energy labels. Therefore, the development of energy label value is examined in this study.

Summary of relevant characteristics

In Table 2 Relevant types of characteristics for real estate in hedonic models are listed. The characteristics are categorized as independent variable, dependent variable or control variable. The independent variable in this study is the energy label of the respective dwelling. The dependent variable is the transaction price of the respective dwelling. The control variables include the time of the transaction and numerous characteristics of the dwelling and its location. The literature that used these variables as such is reported in Table 2.

Table 2 Relevant types of characteristics for real estate in hedonic models

Type of variable	Characteristic	Reported by
Independent (X)	Energy label	(Brounen and Kok, 2011; De Ayala et al., 2016; Fuerst et al., 2015; Kahn and Kok, 2014; Kok and Jennen, 2012)
Dependent (Y)	Transaction price	(Boyle and Kiel, 2001; Brounen and Kok, 2011; Daams et al., 2016; Kahn and Kok, 2014; Koster and Rouwendal, 2017; Lazrak et al., 2014; Rosen, 1974; Sirmans et al., 2005; Van Duijn et al., 2016)
Control (Z)	Time	(Brooks and Tsolacos, 2010; Daams et al., 2016; Van Duijn et al., 2016)
	Dwelling characteristics	(Brounen and Kok, 2011; Daams et al., 2016; van Dam and Visser, 2006; Van Duijn et al., 2016)
	Locational characteristics	(Boyle and Kiel, 2001; Daams et al., 2016; van Dam and Visser, 2006; Eskinasi, 2011; Sirmans et al., 2005; Van Duijn et al., 2016)
	Financial economic characteristics	(Boelhouwer, 2000; van Dam and Visser, 2006; Galati et al., 2011; Lazrak et al., 2014)

Based on the theory and conceptual model three hypotheses are established that will be tested through a quantitative study with the use of statistical methods.

Hypotheses

1. *Energy labels have a significant positive impact on the value of dwellings.*

The study by Brounen and Kok (2011) shows that Dutch households are willing to pay a premium for dwellings with a green energy label. By testing the validity of this hypothesis, the results comprise a larger period of time.

2. *Between 2008-2018, the economic value of energy labels has increased.*

As Figure 2 shows, the house prices between 2008 and 2018 have first shown a decrease following the global financial crisis. From 2013 on, the house prices have increased to a new all-time high. Assuming that the valuation of the characteristics of the dwelling follow the same developments, the value of energy labels has shown a similar development. The economic value used in this hypothesis is referred to as the relative value of the energy label.

3. *Between different geographical regions in Noord-Holland, the economic value of energy labels differs. In regions with lower transaction prices the value of energy labels is relatively higher.*

Regional differences can greatly influence the housing market. By establishing the economic energy label value per COROP-region it is possible to test whether significant differences between COROP-regions exist. The hypothesis is that in regions where the transaction prices are lower, the energy label value is higher. It is assumed that the economic value depends on the investment that is needed. In cheaper dwellings these investments are a relatively larger part of the transaction price. The investment that is needed might fluctuate between the different regions because of varying construction costs for example. However, with the large differences in transaction prices it is assumed that the high transaction prices are established because of something other than the energy labels.

4. *The enforcement of relevant energy label policies has a positive significant impact on the economic value of energy labels.*

In 2015 the energy label was simplified with the aim of a higher adoption rate. At the same time the adoption was stimulated by enforcement of registration during a transaction. The preferred method to study this question would be a difference-in-difference approach. However, as the policy affects the whole country no control group is available. Therefore, the effect of this new policy will be tested using an indicator-variable that indicates whether the transaction occurred during the policy being in effect. This indicator-variable will then be interacted with the different energy label classes.

4. METHODOLOGY

Hedonic model

In real estate studies, hedonic models are a common research instrument as many studies are based on this method (Daams et al., 2016; van Dam and Visser, 2006; Lazrak et al., 2014; Van Duijn et al., 2016). In hedonic modelling, the value of a good is determined by the sum of its individual characteristics. As dwellings are heterogenous goods composed of different characteristics, a transaction price is considered to be the sum of the value of the different characteristics combined. Dwelling value is thus considered to be the same as its transaction price. Hedonic modelling can attribute economic value to the individual characteristics of dwellings. The underlying assumption of hedonic models is that there is sufficient supply and demand in the market to establish a market equilibrium (Rosen, 1974).

Multiple linear regression

The hedonic model is based on the principle of multiple linear regression. With a hedonic model it is possible to determine the effect of an independent variable (x) and control variables (z) on a dependent variable (y). The value of the dwelling is considered to be composed of all characteristics. These characteristics consist of the independent variable and control variables.

Requirements for multiple linear regression

For a multiple linear regression result to be valid the residuals of an analysis need to meet five requirements (Brooks and Tsolacos, 2010):

1. The error term needs to have an average of 0.
   This is tested with the help of a P-P plot.
2. The variance of the residuals needs to be constant at all values of x.
   This is tested with a scatterplot on homoscedasticy.
3. The residuals need to be statistical independent, meaning that there is no autocorrelation.
   This is tested with a Durbin-Watson test.
4. There is no relation between the x- or z-variables.
   This is tested by a test for endogeneity.
5. The residuals need to follow an approximate normal distribution.
   This is tested by creating histograms of the residuals.

These five requirements of the multiple linear regression all apply to the residuals of the model. Between the x- and z-variables there might be multicollinearity. This means that there is high correlation between multiple variables. With a correlation matrix it is possible to determine whether this is the case. The correlation matrix for the preferred model is shown in Appendix 3. The answers to the other assumptions are displayed in Appendix 2.

Empirical model

To test the hypotheses of chapter 3, multiple hedonic models are established. In the hedonic model the independent variable (*x)* is given, in this study that is the energy label. The dependent variable is the linear log of the transaction price of the dwelling (*y*). The control variables (*z*) are formed by dwelling characteristics and locational characteristics as well as the time of the transaction. Table 10 specifies the incorporated variables of each of the different empirical models that are discussed in this chapter. The first three models are refinements of the same concept of determining energy label value for the respective years and regions. Model 4 focusses on the possible change in energy label valuation before and after the energy label registration obligation as from January 1st 2015.

In the first model, the attributed economic value of the different energy label categories is estimated. This model is aimed at being as concise as possible, only the dwelling characteristics are taken into account as is the energy label.

Model 1:

$$\ln Y_{it} = \beta + \alpha E + \gamma Z_D + \upsilon Z_T + \omega Z_L + \varepsilon$$

Whereby $\ln Y_{it}$ is the linear log of the transaction price of the dwelling i at time t. The constant is formed by β. The category of the energy label is represented by E, the core dwelling characteristics by Z_D. The year of the transaction is represented by Z_T. The locational characteristics such as the municipality and the urbanity index are represented by Z_L. The error term is represented by ε whereas the coefficient is γ. It is now possible to estimate whether there is a significant difference in energy label valuation across the different COROP-regions. This model is constructed to answer model 1.

Model 2:

$$\ln Y_{it} = \beta + \alpha E + \gamma Z_D + \upsilon Z_T + \omega Z_L + \varphi Z_A + \varepsilon$$

In the second model the first model is enhanced with all available locational data. The distances to several amenities for the respective neighborhood of the transaction are added as well as the urbanity index. These amenities characteristics are represented by Z_A. This model is constructed to answer hypothesis 1 in more detail.

Model 3:

$$\ln Yi_t = \beta + \alpha E + \gamma Z_D + \upsilon Z_T + \omega Z_L + \delta T_T + \phi T_C + \varepsilon$$

The third model builds upon the first two models but adds an interaction to test for the effect of energy label valuation over time and across regions. This model is constructed to answer hypotheses 2 and 3. The interaction-variables for the interaction between energy label and year is represented by T_T. The interaction-variables for the interaction between energy label and COROP-region is represented by ϕT_C. The interactions are composed by the product of the energy label-variable with the respective interaction-variable.
The locational characteristics are omitted in this model as the analyses indicated that the number of transactions dropped significantly because of the inclusion of locational characteristics. This is more discussed in more detail in section 5.

In the fourth model the focus is on the moments of changes in energy label policy by the Dutch Government. This model is constructed to answer hypothesis 4. The

obligation to provide energy labels during a sales transaction has become effective on January 1st 2015. Therefore, a dummy-variable will be added that represents a pre-obligation or post-obligation time-period. This dummy variable is then interacted with the energy label classes. The different subsamples of the pre-obligation and post-obligation will be tested on structural differences with the help of a Chow-test. The previous interactions as mentioned in Model 3 are left out for clarity.

Model 4:

$$\ln \mathrm{Yi}_t = \beta + \alpha E + \gamma Z_D + \upsilon Z_T + \omega Z_L + \sigma D_O + \varepsilon$$

The interaction of the dummy-variable for the pre- or post-obligation timeframe is represented by D_O.

5. DATA

This chapter first describes the selection and collection of the data. Subsequently the operationalization of the data is described. This study uses data provided by Central Bureau of Statistics (CBS), Dutch Association of Realtors (NVM) and the Netherlands Enterprise Agency (RVO).

Selection of timeframe and location

As the Dutch Association of Realtors restricts the use of their data by students the area of interest is limited to a period of 10 years and one province of choice. As the introduction of energy labels for dwellings in The Netherlands took place in 2008 the consecutive 10 years are taken into consideration to study the development of value of energy labels.

The selection of province is based on the assumption that Noord-Holland has a large variety of housing markets; from a booming Amsterdam housing market to the northern part of the province that suffers from a decline in population. This creates the opportunity to study the effect of these trends on the valuation of energy labels.

In order to estimate regional differences within the province of Noord-Holland the regional scale of COROP-region is used. COROP-regions are designed by the government for analytical purposes and aim to cover the same geographical region over consecutive years (Statistics Netherlands, 2018a). The province of Noord-Holland is divided into 7 COROP-regions which are shown in Table 5. Using the fine-grained scale of a municipality would result in more locational variables becoming unavailable for all municipalities and years as a result of missing data.

Included characteristics

The decision to include certain characteristics in this study is made based on the significance of effects of these characteristics in previous literature and the significance in this study. The list of included characteristics in this study is reported in Table 3.

Unfortunately, the supply data of dwellings split out over the different COROP-regions is only available for a specific period as displayed in Figure 3. When including this variable all the transactions that took place outside the timeframe of February 2013 till December 2016 would have to be dropped. As Table 6 indicates, most transactions took place in the most recent years, resulting in a significant loss in number of transactions included in the model. Therefore, the decision is made to keep this variable out of the models.

The average mortgage rate is by choice not included in the model. The uncertainty of the impact of these factors is the reason for this. It would have been possible to include the average mortgage interest rate for newly established mortgages. But it possible to transfer a mortgage to a new home, thereby taking the old terms and conditions for the outstanding balance to a new property. This creates a mixed mortgage, and therefore the data for new mortgages would not be representative. No other public source of data was available that would make it possible to correct for this (e.g. know the type of mortgage for each transaction and its terms and

conditions). Knowing this limitation of possible reliable and still include it in the model would decrease the validity. Therefore, the mortgage interest is left out. The impact of the economic growth (and thus the increasing housing prices) is corrected for by taking the individual years and their average prices into account in the interpretation of the models.

Table 3 Included characteristics, grouped per type

Type of characteristic	Variable	Description
Dwelling	Energy label	Categorical variable for the energy label class (A to G)
	Floor space	Livable area of floor space in m^2
	Rooms	Number of rooms
	Private parking	Dummy for private parking (1 = yes)
	Garden	Dummy for garden (1 = yes)
	Balcony	Dummy for balcony (1 = yes)
	Roof terrace	Dummy for roof terrace (1 = yes)
	Type of dwelling	Categorical variable of different types of dwelling
	Construction period	Categorical variable for the different specified construction periods.
Locational	COROP-region	Categorical variable of the COROP-region that the dwelling is situated in
	Urbanity index	Categorical variable for the different urban density classes per neighborhood defined by the Central Statistics Bureau.
	Distance to hospital	Average distance in kilometer from neighborhood.
	Distance to freeway entrance	Average distance in kilometer from neighborhood.
	Distance to large grocery store	Average distance in kilometer from neighborhood.
	Distance to train station	Average distance in kilometer from neighborhood.
	Distance to restaurant	Average distance in kilometer from neighborhood.
Transactional	Time of the transaction	Date of the sale
	Transaction price	Nominal transaction price in €
	Days listed	Number of days the dwellings was listed for sale

Energy labels

The energy label data is provided by the Netherlands Enterprise Agency that is accountable for the registration of energy labels during dwelling transactions. In the acquired dataset a total of 3.721.779 valid energy labels are collected. This dataset covers all the energy labels for dwellings that are registered in the Netherlands, from as early as September 2008 till September 2018.
Table 4 shows the distribution of the different energy labels across all the transactions included in the dataset provided by the Netherlands Enterprise Agency.

Table 4 Distribution of energy labels

Energy label	Number of transactions	Relative share
A	599.843	16,1 %
B	584.183	15,7%
C	1.121.272	30,1%
D	617.219	18,0%
E	370.049	9,9%
F	215.845	5,8%
G	159.368	4,3%

Dwelling and transactional data

The data used in this study are secondary data acquired by the Dutch Association of Realtors for the transaction data. The acquired dataset covers the province of Noord-Holland in the period between 2008 and Q2 2018. This selection includes 245.383 registered transactions. The included data provide the address of the dwelling, its transaction price, the transaction date and dwelling characteristics such as floor space and number of rooms. Table 5 shows the included transactions per COROP-region.

Table 5 Total transactions per COROP-region

COROP-region	Number of transactions	Relative share
Kop van Noord-Holland	21.155	8,6%
Alkmaar en omgeving	21.388	8,7%
IJmond	16.497	6,7%
Agglomeratie Haarlem	32.111	13,1%
Zaanstreek	13.237	5,4%
Groot-Amsterdam	114.424	46,5%
Gooi en Vechtstreek	27.168	11,0%

Locational data

To enrich the hedonic model characteristics locational data are included. These data are secondary data provided by CBS Statistics Netherlands, the Dutch national statistics bureau.

This dataset includes information on the urbanity-index per COROP-region as well as average distances to several types of amenities per COROP-region. Appendix 1 lists all included variables. The list of included variables is based on the literature as listed in Table 2 as well as the availability of data. Not all relevant locational characteristics are available for all regions and years and can therefore not be used reliably. An important missing variable is the distance to (several different types of) public green space.

Operationalization

During a transaction the Dutch Association of Realtors registers a variety of characteristics of the dwelling as well as the transaction price and date. The Dutch Association of Realtors does not register the energy label during a transaction.
The provided energy label data is updated on a daily basis by the Netherlands Enterprise Agency. The data therefore is always up-to-date. However, old data is overwritten which makes it impossible to track the energy label history of a specific address. If an address has been sold multiple times it is not possible to determine whether the energy label has changed between the transactions. Therefore, only the most recent transaction, that matches the registration of the energy label, is taken into account.

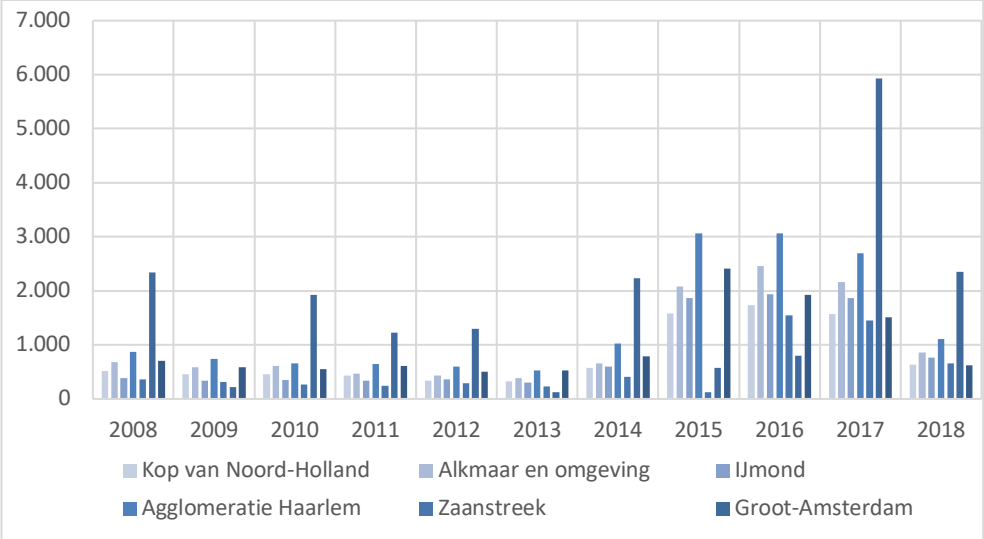
To be able to estimate the energy label value both datasets need to be joined. The transactions that do not have a registered energy label after this merger are dropped. This results in a reduction of the number of included transactions from 245.383 to 96.376.

Afterwards the CBS dataset is merged to the transactions. The COROP-region of the transaction as well as the moment (year) of the transaction need to be matched. Not all CBS data is available for all years or regions as shown in Table 7.

The first step was to adjust the datasets in such a way that the address notation matches and the datasets can be merged. After the merger of datasets, only the transactions that match with a registered energy label are kept. As only the most recent energy label is registered, the transactions that took place more than 90 days prior to the registration of the energy label are dropped. This is to ensure that the transaction matches with the most recent label that is registered.
If observations miss data on one or more of the included variables the observations are dropped during the regression analysis.

Table 6 shows the distribution of the matched transactions per COROP-region in the period 2008-2018.

Table 6 Transactions per COROP-region and year.



Energy label

For dwellings that have been sold multiple times only the transactions that took place during or after the registration of the most recent energy label are included. By removing the transactions that took place before the registration of the energy label only the most recent available energy label is used for all transactions. This is a result of the unavailability of historic energy label data.

Outliers

As all transactions of dwellings that took place in the province of Noord-Holland are still included some special dwellings exist in the dataset. This might affect the representativeness of the final conclusions and therefore these outliers are removed. Some criteria for outliers to be removed are a high transaction price (>€2,500,000), floor space (>700 m^2) or number of rooms (>10). The exact steps to remove outliers and all other steps that are taken to prepare the data for analysis and the steps taken to analyze the data are documented in

Appendix 6.

Descriptive statistics

The descriptive statistics of the variables that are included in the regression models are displayed in this section. Table 7 displays the descriptive statistics of the variables that are continuous or categorical variables with two different categories. In Table 9 the frequencies are given for the categorical variables that offer more than two possible categories.

Table 7 Descriptive statistics

Variable	N	Mean	SD	Min	Max
Transaction price €	96.376	283.459	187.466	20.000	2.500.000
Days listed	96.376	136,17	225,38	0	1.824
Roof terrace	96.376	0,08	0,27	0	1
Garden	96.376	0,59	0,49	0	1
Balcony	96.376	0,34	0,48	0	1
Parking space	96.376	0,27	0,44	0	1
After 2015 Energy Label policy change	96.376	0,65	0,48	0	1
Floor space square meters	96.376	107,46	41,79	26	536
Address density	92.719	2.549,83	2.313,34	2	12.259
Distance to hospital	23.271	1,17	1,03	0	10
Distance to large supermarket	85.954	1,61	3,12	0	71
Distance to public swimming pool	23.271	1,05	3,09	0	10
Distance to freeway entrance	19.357	5,57	7,81	0	159
Distance to train station	15.614	4,86	6,62	0	10

The descriptive statistics in Table 7 show that most data is available for 96.376 observations. The average time on the market, or days listed is 4,5 months. The dwelling that has been for sale for the longest period in the dataset is almost 5 years. The mean transactions price for all regions and years combined is €283.459.
In Table 8 the mean transaction prices are split out per COROP-region. In the CBS-column the average prices per COROP-region are displayed, based on the CBS Statistics Netherlands data (2018b).

Table 8 Transaction prices per COROP-region

Variable	Mean	Std. Err.	CBS
Kop van Noord-Holland	€215.587	€1.129	€208.545
Alkmaar en omgeving	€252.188	€1.393	€242.182
IJmond	€252.713	€1.448	€233.091
Agglomeratie Haarlem	€333.825	€1.842	€301.545
Zaanstreek	€221.769	€1.231	€193.091
Groot-Amsterdam	€297.934	€1.013	€259.545
Gooi en Vechtstreek	€320.920	€2.394	€326.182

The transaction data and the data provided by CBS Statistics Netherlands both show comparable differences across the different COROP-Regions. The differences

between the two data sources can be explained by the different establishment methods. The transaction data is based on the actual transaction price whereas the CBS data is based on an estimate for all the dwellings and not only the ones that are sold.

The data for some variables, including the distance to the different amenities, is not available for all transactions, the number of observations for these variables is substantially lower. This is a result of the absence of data for certain periods of time and certain amenities. In other words, not all selected amenities are recorded for the full period of this study.

Table 9 Descriptive statistics, frequencies per energy label class

Variable	Energy label class							
Transaction year	A	B	C	D	E	F	G	Total
2008	473	849	1.776	1.037	856	559	298	5.848
2009	445	727	1.609	919	762	511	270	5.243
2010	411	619	1.535	857	705	450	255	4.832
2011	396	537	1.162	722	567	373	212	3.969
2012	347	492	1.166	717	554	351	195	3.822
2013	322	523	1.033	672	549	294	196	3.589
2014	597	897	1.901	1.061	786	594	453	6.289
2015	2.432	2.488	5.075	2.505	2.146	1.678	1.643	17.967
2016	3.179	299	5.802	2.879	2.390	1.811	1.581	20.632
2017	2.761	2.493	4.711	2.373	1.955	155	1.335	17.188
2018	1.115	984	1.966	943	787	642	560	6.997
COROP-region								
Kop van Noord-Holland	13	1.558	3.138	1.089	550	475	516	8.626
Alkmaar en omgeving	1.675	1.639	3.907	1.721	1.024	774	647	11.387
IJmond	1.007	1.122	2.712	1.397	1.161	922	777	9.098
Agglomeratie Haarlem	1.489	1.536	3.506	2.099	2.475	2.337	1.556	14.998
Zaanstreek	104	842	1.735	1.074	924	732	713	706
Groot-Amsterdam	5.483	6.104	10.072	5.181	4.081	2.031	1.506	34.458
Gooi en Vechtstreek	484	798	2.666	2.124	1.852	1.542	1.283	10.749
Rooms								
1	122	140	278	139	122	52	65	918
2	1.119	1.477	3.211	173	1.463	627	372	9.999
3	2.952	3.905	6.199	352	274	1.789	1.287	22.392
4	2.572	2.867	6.095	3.586	3.333	2.682	2.001	23.136
5	3.557	3.557	8.771	38	2.748	1.929	1.595	25.957
6	157	1.228	2.402	1.298	1.093	1	846	9.437
7	404	276	547	411	367	443	423	2.871
8	120	92	160	127	130	182	210	1.021
9	43	41	54	62	46	84	131	461
10	19	16	19	12	25	25	68	184
Construction period								
1500-1905	151	195	653	796	801	1.063	103	4.689
1906-1930	183	358	1.227	2.079	2.205	3.017	2.961	1.203
1931-1944	40	73	509	1.098	152	1.783	1.667	669
1945-1959	81	103	1.078	1.813	1.988	848	542	6.453
1960-1970	110	466	3.993	4.399	3.983	1.584	634	15.169
1971-1980	156	1.211	713	3.146	1.407	467	106	13.623
1981-1990	215	2.791	10.219	1.175	108	26	21	14.555
1991-2000	2.374	5.728	2.588	95	22	9	22	10.838
>2001	9.168	2.674	339	84	33	16	15	12.329
Dwelling type								
Terraced house	449	3.759	10.495	4.247	4.273	3.482	1.406	32.152
Corner house	1.158	1.342	3.652	203	1.348	121	1.547	12.287
Semi-detached house	999	1.331	1.658	1.017	820	1.036	147	8.331

Detached house	939	908	1.058	1.048	447	101	1.462	6.872
Ground floor apartment	654	873	1.404	700	391	138	91	4.251
Regular apartment	1.641	2.353	3.521	188	1.453	662	454	11.964
Multi-floor apartment	539	581	1.041	361	277	98	77	2.974
Portico hall flat	1.253	1.529	3.121	1.936	149	641	304	10.274
Gallery hall apartment	805	923	1.786	1.466	1.568	536	187	7.271
Urbanity index								
Very highly urban	3.713	4.158	7.483	4.848	4.672	3.405	2.353	30.632
Highly urban	4.222	4.392	1.048	4.965	3.968	2.782	2.074	32.883
Moderately urban	1.718	1.902	3.688	1.834	1.417	1.048	966	12.573
Mildly urban	1.563	1.613	3.186	1.565	107	784	742	10.523
Not urban	1.258	1.526	2.866	1.458	936	791	860	9.695
Total	12.478	13.599	27.736	14.685	12.067	8.813	6.998	96.376

The descriptive statistics in Table 9 are in line with what is to be expected, most transactions took place in the Amsterdam area and in the years following the financial crisis.
Most of the houses sold are terraced houses and most energy label ratings are in the mid-range, C and D.
The share of transactions sharply increases from 2015 on. It is important to note that the data for 2018 are only available for the months January till September.

The urbanity index is composed of 5 different classes ranging from “Not urban” for places where the address-density is lower than 500 per square kilometer to “Very highly urban” for places where the address-density is above 2500 per square kilometer.

6. RESULTS

Based on the hedonic models from chapter 4, multiple regressions are performed. The results of these quantitative analyses will be presented in this chapter. Based on the presentation of the results the answers to the hypotheses of chapter 3 are discussed here.

Table 10 Results regression models, summary

Variable	Model 1	se	Model 2	se	Model 3	se	Model 4	se
Energy label class								
A	0.1204***	(0.005)	0.045***	(0.011)	0.082***	(0.020)	0.118***	(0.007)
B	0.1203***	(0.004)	0.057***	(0.009)	0.013	(0.018)	0.088***	(0.006)
C	0.1037**	(0.004)	0.035***	(0.008)	-0.021	(0.015)	0.064***	(0.006)
D	0.0841***	(0.003)	0.018**	(0.008)	-0.030*	(0.016)	0.035***	(0.006)
E	0.0748***	(0.004)	0.015*	(0.008)	-0.039**	(0.016)	0.023***	(0.006)
F	0.0594***	(0.004)	0.005	(0.008)	-0.010	(0.017)	0.015**	(0.006)
G	-		-		-		-	
Floor space (log)	0.765***	(0.004)	0.776***	(0.008)	0.759***	(0.004)	0.764***	(0.004)
Construction period								
1500-1905	0.130***	(0.005)	0.034***	(0.011)	0.124***	(0.005)	0.126***	(0.005)
1906-1930	0.033***	(0.005)	-0.042***	(0.010)	0.032***	(0.005)	0.029***	(0.005)
1931-1944	0.063***	(0.005)	-0.012	(0.011)	0.065***	(0.005)	0.060***	(0.005)
1945-1959	-0.073***	(0.005)	-0.126***	(0.010)	-0.075***	(0.005)	-0.077***	(0.005)
1960-1970	-0.178***	(0.004)	-0.218***	(0.009)	-0.182***	(0.004)	-0.181***	(0.004)
1971-1980	-0.171***	(0.004)	-0.186***	(0.009)	-0.173***	(0.004)	-0.173***	(0.004)
1981-1990	-0.110***	(0.004)	-0.140***	(0.008)	-0.112***	(0.004)	-0.113***	(0.004)
1991-2000	-0.021***	(0.003)	-0.041***	(0.008)	-0.023***	(0.003)	-0.023***	(0.003)
>2001	-		-		-		-	
Dwelling type								
Terraced house	0.088***	(0.004)	0.034***	(0.009)	0.090***	(0.004)	0.088***	(0.004)
Corner house	0.114***	(0.005)	0.063***	(0.010)	0.116***	(0.005)	0.113***	(0.005)
Semi-detached house	0.209***	(0.005)	0.169***	(0.011)	0.215***	(0.005)	0.208***	(0.005)
Detached house	0.360***	(0.005)	0.348***	(0.012)	0.373***	(0.005)	0.360***	(0.005)
Ground floor apartment	0.118***	(0.005)	0.061***	(0.010)	0.116***	(0.005)	0.117***	(0.005)
Regular apartment	0.119***	(0.004)	0.046***	(0.007)	0.118***	(0.004)	0.118***	(0.004)
Multi-floor apartment	-0.013**	(0.005)	-0.011	(0.010)	-0.012**	(0.005)	-0.013***	(0.005)
Portico hall flat	0.040***	(0.003)	-0.004	(0.007)	0.043***	(0.003)	0.040***	(0.003)
Gallery hall apartment	-		-		-		-	
Roof terrace = 1	0.033***	(0.003)	0.024***	(0.006)	0.033***	(0.003)	0.033***	(0.003)
Garden = 1	-0.005**	(0.002)	0.009*	(0.005)	-0.004	(0.002)	-0.005**	(0.002)
Balcony = 1	0.031***	(0.002)	0.029***	(0.004)	0.031***	(0.002)	0.031***	(0.002)
Parking space = 1	0.108***	(0.002)	0.124***	(0.005)	0.107***	(0.002)	0.108***	(0.002)
Transaction year								
2009	-0.041***	(0.004)	-0.062***	(0.011)	-0.048***	(0.018)	-0.042***	(0.004)
2010	-0.034***	(0.004)	-0.052***	(0.011)	-0.052***	(0.019)	-0.035***	(0.004)
2011	-0.038***	(0.005)			0.020	(0.020)	-0.039***	(0.005)
2012	-0.108***	(0.005)			-0.144***	(0.020)	-0.109***	(0.005)
2013	-0.155***	(0.005)			-0.228***	(0.020)	-0.155***	(0.005)
2014	-0.103***	(0.004)			-0.117***	(0.016)	-0.105***	(0.004)
2015	-0.040***	(0.003)			-0.103***	(0.014)	-0.091***	(0.007)
2016	0.033***	(0.003)			-0.009	(0.014)	-0.017***	(0.007)
2017	0.146***	(0.003)			0.088***	(0.014)	0.096***	(0.007)
2018	0.226***	(0.004)			0.157***	(0.016)	0.176***	(0.007)

Urbanity index								
Very highly urban	0.110***	(0.004)			0.113***	(0.003)	0.110***	(0.004)
Highly urban	0.023***	(0.003)	-0.352***	(0.014)	0.024***	(0.003)	0.023***	(0.003)
Moderately urban	0.038***	(0.004)	-0.260***	(0.012)	0.037***	(0.004)	0.038***	(0.004)
Mildly urban	0.037***	(0.004)	-0.196***	(0.012)	0.036***	(0.004)	0.037***	(0.004)
Not urban	-		-		-		-	
Address density (log)			0.142***	(0.005)				
Distance to hospital (log)			-0.001**	(0.000)				
Distance to large supermarket (log)			0.003	(0.003)				
Distance to public swimming pool (log)			-0.001	(0.001)				
Distance to freeway entrance (log)			-0.006*	(0.003)				
Distance to train station (log)			-0.004***	(0.001)				
Constant	8.861***	(0.029)	8.536***	(0.075)	8.957***	(0.031)	8.904***	(0.029)
Observations	96,308		15,614		96,308		96,306	
R-squared	0.809		0.837		0.813		0.810	
Address density	No		Yes		No		No	
Distance to amenities	No		Yes		No		No	
Number of Rooms	Yes		Yes		Yes		Yes	
Municipality	Yes		Yes		Yes		Yes	
Year * Energy Label interactions	No		No		Yes		No	
COROP-Region * Energy Label interactions	No		No		Yes		No	
Energy Label policy enforcement interaction	No		No		No		Yes	
Adjusted R-squared	0.809		0.835		0.813		0.809	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

No base levels have been manually set in the regression, as a result some levels of categorical variables are omitted because of collinearity, this is indicated with "–" in the coefficient column.

Regression results

As Table 10 shows, most results (model 1) are significant. More floor space adds significant positive value to a dwelling, as do a roof terrace, balcony or parking space. Compared to the base construction period of newer than 2001, dwellings that are constructed between 1945-2000 sell at different discounts. The dwellings that are constructed before 1945 sell at a premium compared to the newer dwellings. This is in line with previous literature.

The transactions years show developments that are in line with the price index as displayed in Figure 2. The years of the housing crisis show negative value compared to the base year of 2008 whereas the recent years show a premium paid for dwellings.

Model 2 shows that increased distance to amenities lowers the value of dwellings, for some amenities such as the distance to a train station or freeway entrance these results are significant. The results for the distance to a supermarket or public swimming pool are not significant.

Notes

It is important to note that model 2 has less observations included in the model compared to models 1 and 3. This is due not all amenity-data being available for all regions in all years. Therefore, the amenities are left out in the third and fourth model which include the interaction between energy labels and the transaction years and regions (model 3) and the label enforcement policy and energy labels (model 4). Based on the number of included observations and the adjusted R-squared score, model 3 is the preferred model to answer hypotheses 2 and 3. For hypothesis 4 model 4 is preferred as it includes the interaction of energy label enforcement. Table 10 displays the main regression results of the regression models. The full regression results are displayed in Appendix 5. Unfortunately, the value of energy label G could not be estimated as it is omitted because of collinearity. This happens in all the models and interactions where the G-label is included.

Hypotheses

The following section will discuss and answer the hypotheses.

Hypothesis 1

"Energy labels have a significant positive impact on the value of dwellings."

To answer hypothesis 1 the most concise model with the no interactions is used (model 1). In this model all the results for the energy labels are significant. Table 11 shows the relevant results based on Table 10.

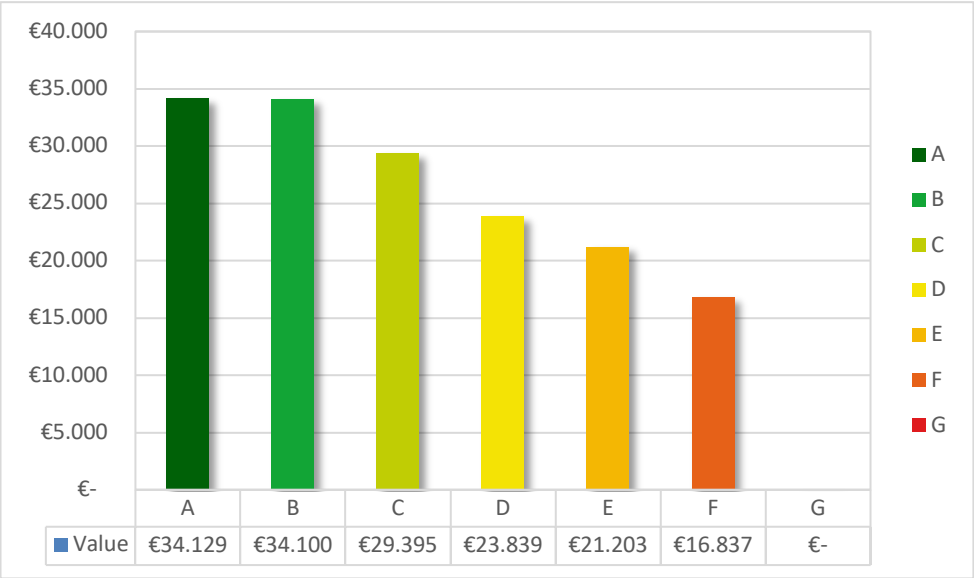
Table 11 Results summary model 1

Energy label class	Model 1	se
A	0.1204***	(0.005)
B	0.1203***	(0.004)
C	0.1037**	(0.004)
D	0.0841***	(0.003)
E	0.0748***	(0.004)
F	0.0594***	(0.004)
G	-	

An A-label represents 12,04% of the transaction price of a dwelling. For the labels B (12,03%), C (10,37%), D (8,41%), E (7,48%) and F (5,94%) the attributed value is lower.

Figure 4 is a graphical representation of the results of model 1 where the results are monetized. In this figure the value of the energy label is monetized by taking the coefficients and multiplying them by average transaction price of the dataset. Note that the G-class is omitted and therefore has no value.

Figure 4 Energy label valuation



The results suggest that for an A-label energy rating the highest premium is paid. The premium of a B-label is slightly lower compared to an A-label. The largest increase in premium is paid between the base G-label and F-label. From the F-label the increase per label category is significantly smaller.

The results suggest that Hypothesis 1 cannot be rejected.

A possible explanation for the differences in increased valuation for each label class upgrade is the steps between the different energy label classes have different requirements. For example, to be eligible for an F-label the dwelling at least needs to have a central heating system, which is can be seen as a technical upgrade from independent gas heaters per room as well as a comfort improvement as the air quality is increased (no open fire in the room). The increase from a B-rating to an A-rating requires the heat to be provided electrically (by means of a heat pump), this is mainly a technical improvement but not necessarily a comfort improvement (Milieu Centraal, 2019).

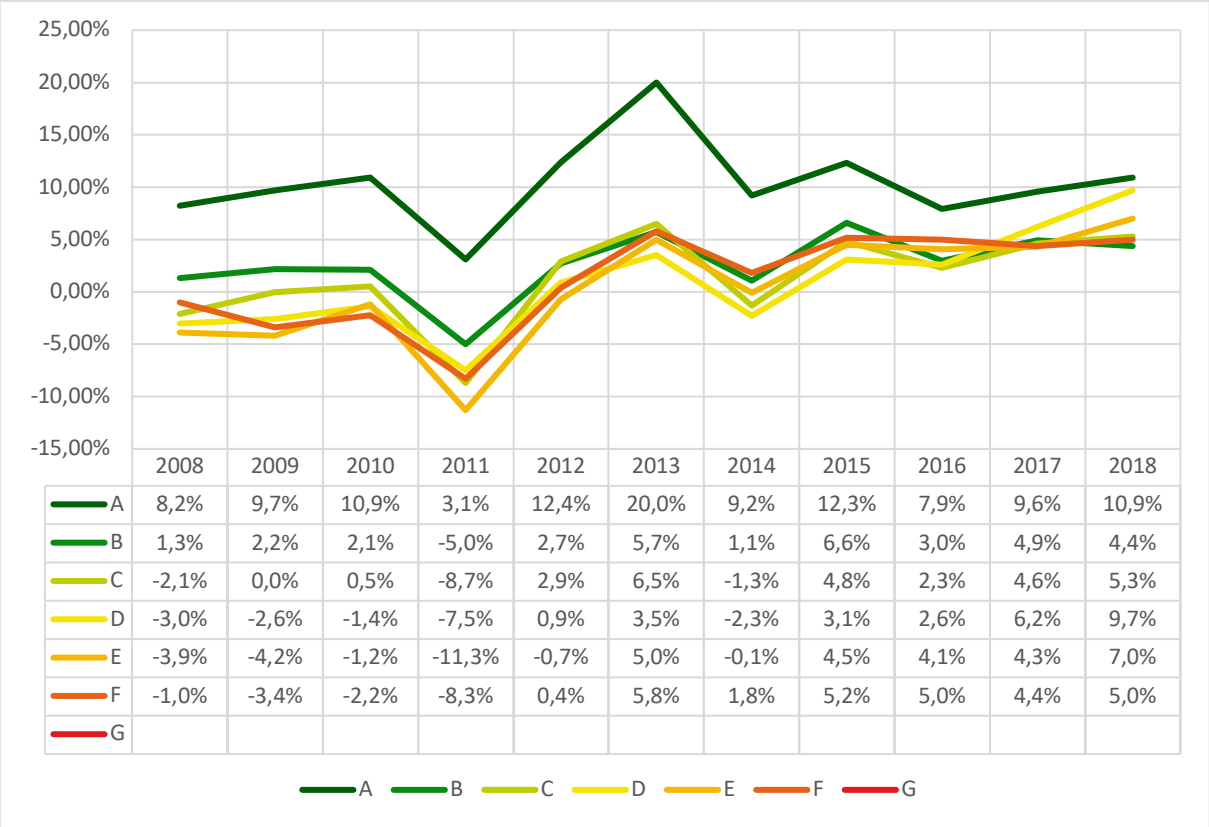
This possible difference in ways to value the improvement of the energy label rating might explain the difference in paid premium per label category.

Hypothesis 2

"Between 2008-2018, the economic value of energy labels has increased."

The regression results of the interaction between the years and energy label classes are displayed in Appendix 5. A graphical representation of the results is made in Figure 5, the percentages that are shown are the sum of the energy label coefficients and the coefficients of the interaction of the respective energy labels with the different years. It is important to note however that the interaction does not yield any significant results in the years 2009 and 2010, in the consecutive years the significance varies across the different energy label classes.

Figure 5 Development energy label as share of transaction price

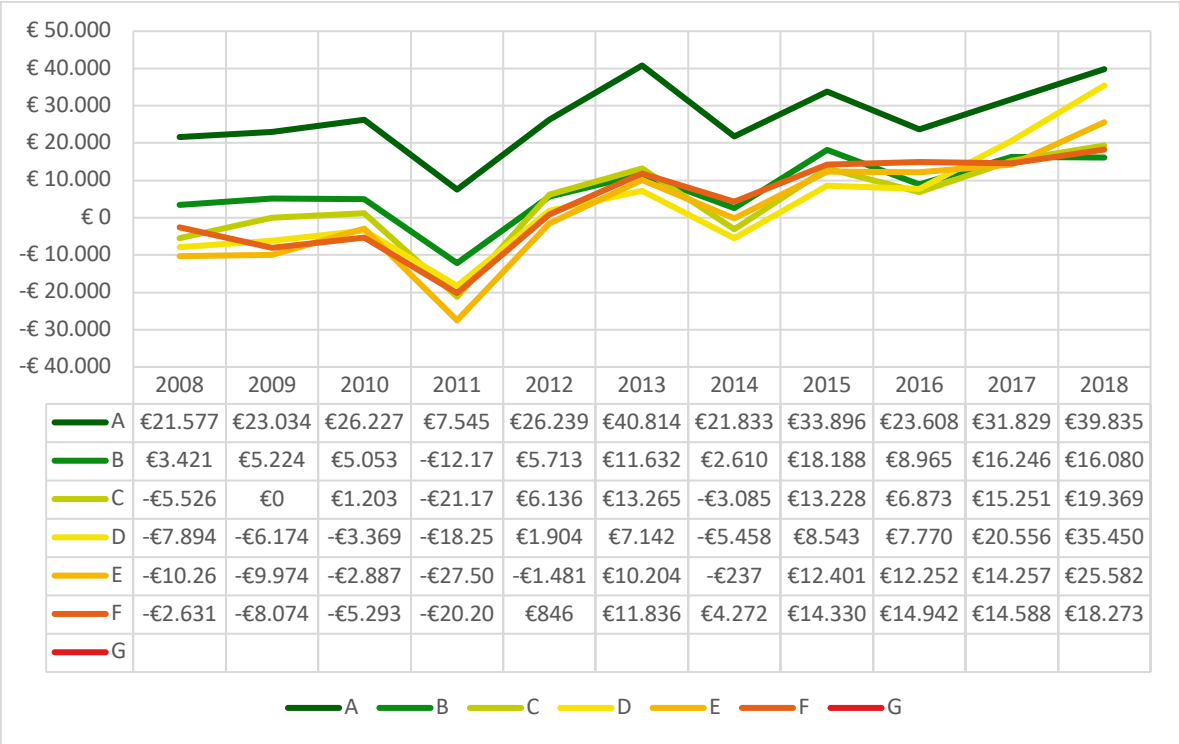


	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
A	8,2%	9,7%	10,9%	3,1%	12,4%	20,0%	9,2%	12,3%	7,9%	9,6%	10,9%
B	1,3%	2,2%	2,1%	-5,0%	2,7%	5,7%	1,1%	6,6%	3,0%	4,9%	4,4%
C	-2,1%	0,0%	0,5%	-8,7%	2,9%	6,5%	-1,3%	4,8%	2,3%	4,6%	5,3%
D	-3,0%	-2,6%	-1,4%	-7,5%	0,9%	3,5%	-2,3%	3,1%	2,6%	6,2%	9,7%
E	-3,9%	-4,2%	-1,2%	-11,3%	-0,7%	5,0%	-0,1%	4,5%	4,1%	4,3%	7,0%
F	-1,0%	-3,4%	-2,2%	-8,3%	0,4%	5,8%	1,8%	5,2%	5,0%	4,4%	5,0%
G											

The results clearly indicate that the A-label attributes the highest value to a dwelling. As the figure shows, the overall value of the energy label approximately follows the real estate cycle when compared to the price index of dwellings displayed in Figure 2. The valuation of energy labels peaked during the housing market crisis of 2013. However, in 2011 the opposite was the case with a discount for all label classes. The reason for this is uncertain, there are no significant differences in the types of dwellings sold in the years 2011 and 2013 as well as the region where the transactions took place. There must be another external factor that explains the decrease in valuation in 2011.

In Figure 6 the value of the energy label is monetized by adding the coefficients of the energy labels and the interaction of energy labels with the different years. This sum of coefficients is then multiplied by the average transaction price of the respective year. The energy label class G is omitted and therefore has no value.

Figure 6 Development energy label value monetized



	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
A	€21.577	€23.034	€26.227	€7.545	€26.239	€40.814	€21.833	€33.896	€23.608	€31.829	€39.835
B	€3.421	€5.224	€5.053	-€12.17	€5.713	€11.632	€2.610	€18.188	€8.965	€16.246	€16.080
C	-€5.526	€0	€1.203	-€21.17	€6.136	€13.265	-€3.085	€13.228	€6.873	€15.251	€19.369
D	-€7.894	-€6.174	-€3.369	-€18.25	€1.904	€7.142	-€5.458	€8.543	€7.770	€20.556	€35.450
E	-€10.26	-€9.974	-€2.887	-€27.50	-€1.481	€10.204	-€237	€12.401	€12.252	€14.257	€25.582
F	-€2.631	-€8.074	-€5.293	-€20.20	€846	€11.836	€4.272	€14.330	€14.942	€14.588	€18.273
G											

When looking at the development over the period 2008-2018 the valuation of energy labels increased in 2018 compared to the begin of the period.

The results suggest that Hypothesis 2 cannot be rejected.

The results as displayed in Figure 6 show very different results per year than the valuation for all years combined as in Figure 4, in the most recent years the premium for an A-label is smaller than the premium for any other label (except the base G-label). This not in line with previous literature that had a similar valuation pattern as discussed with hypothesis 1. A possible explanation for these differences might be the housing market cycle. In times of low supply of dwellings other factors might have larger impact on the buyers-decision. The peak of energy label valuation is in line with the moment of highest supply of dwellings as can be seen in Figure 3.

As no supply-data is available for the period before 2013 so it is not possible to test this possible explanation. Table 6 shows that the number of transactions in 2011 was comparable to the years 2010 and 2012 so a possible explanation must be an external factor.

Hypothesis 3

"Between different geographical regions in Noord-Holland, the economic value of energy labels differs. In regions with lower transaction prices the value of energy labels is relatively higher."

The regression results of the interaction between the different COROP-regions and the energy label classes are displayed in Appendix 5. A graphical representation of the results is made in Figure 7. It is important to note however that the interaction does not yield any significant results in the COROP-region Groot-Amsterdam, in the other COROP-regions the significance varies across the different energy label classes as Appendix 5 shows.

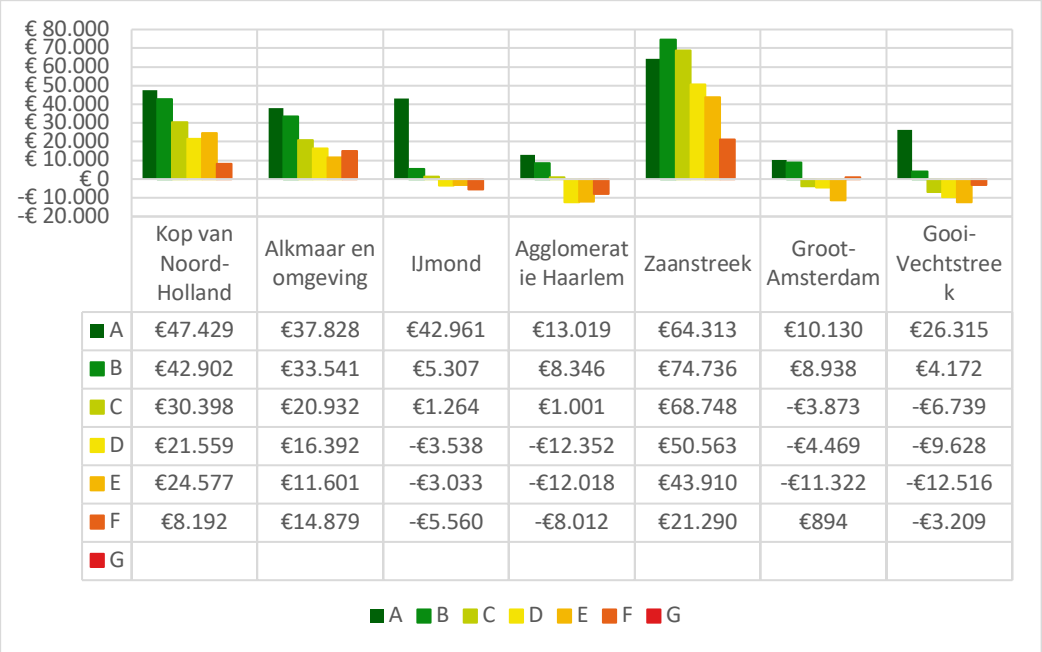
Figure 7 Energy label value per COROP-region as share of transaction price

	Kop van Noord-Holland	Alkmaar en omgeving	IJmond	Agglomeratie Haarlem	Zaanstreek	Groot-Amsterdam	Gooi-Vechtstreek
A	22,00%	15,00%	17,00%	3,90%	29,00%	3,40%	8,20%
B	19,90%	13,30%	16,80%	2,50%	33,70%	3,00%	1,30%
C	14,10%	8,30%	14,30%	0,30%	31,00%	-1,30%	-2,10%
D	10,00%	6,50%	7,90%	-3,70%	22,80%	-1,50%	-3,00%
E	11,40%	4,60%	3,60%	-3,60%	19,80%	-3,80%	-3,90%
F	3,80%	5,90%	1,90%	-2,40%	9,60%	0,30%	-1,00%
G							

The results indicate that the impact of energy labels differ greatly between the different COROP-regions. The different coefficients might be due to the absolute monetary value of the energy label being comparable between the different regions but the mean transaction prices are not. Therefore, the energy label is monetized to see the economic value of the energy labels in the different regions.

In Figure 8 the value of the energy label is monetized by taking the coefficients of the energy label and the coefficient of the interaction between energy label and COROP-region. This sum of coefficients is then multiplied by the average transaction price of the respective COROP-region.

Figure 8 Energy label valuation per COROP-region monetized



	Kop van Noord-Holland	Alkmaar en omgeving	IJmond	Agglomeratie Haarlem	Zaanstreek	Groot-Amsterdam	Gooi-Vechtstreek
A	€47.429	€37.828	€42.961	€13.019	€64.313	€10.130	€26.315
B	€42.902	€33.541	€5.307	€8.346	€74.736	€8.938	€4.172
C	€30.398	€20.932	€1.264	€1.001	€68.748	-€3.873	-€6.739
D	€21.559	€16.392	-€3.538	-€12.352	€50.563	-€4.469	-€9.628
E	€24.577	€11.601	-€3.033	-€12.018	€43.910	-€11.322	-€12.516
F	€8.192	€14.879	-€5.560	-€8.012	€21.290	€894	-€3.209
G							

What the results clearly indicate is that even when the value is monetized there are significant differences in valuation of energy labels across the different COROP-regions.

The results suggest that Hypothesis 3 cannot be rejected.

A possible explanation for the different valuation of energy labels might be the popularity of a certain region. The Agglomeratie Haarlem, Groot-Amsterdam and Gooi en Vechtstreek-region show comparable results. These three regions form the core of the Randstad within the province of Noord-Holland and are highly popular. Table 5 shows that in recent years the most transactions took place in those three regions.

Furthermore, the average transaction price is the highest in these three regions, as

Table 12 indicates.

Table 12 Mean transaction prices (per square meter) per COROP-region

COROP-region	Mean	se
Kop van Noord-Holland	€1.821	€6
Alkmaar en Omgeving	€2.212	€7
IJmond	€2.279	€8
Agglomeratie Haarlem	€2.877	€7
Zaanstreek	€2.131	€6
Groot-Amsterdam	€3.052	€7

Gooi en Vechtstreek €2.593 €9

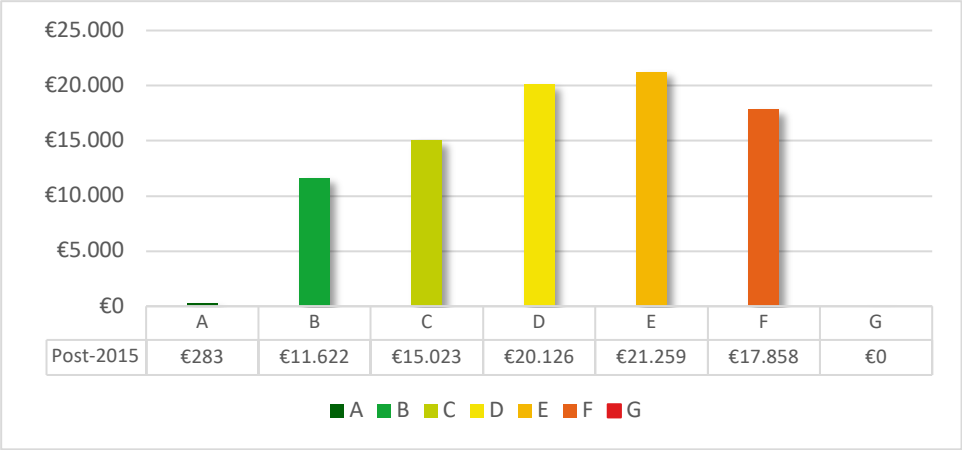
This suggests that the more expensive a dwelling, the less important its energy label. When looking at the regions Kop van Noord-Holland, Alkmaar en Omgeving and IJmond the number of transactions is comparable, so is the mean transaction price per square meter. Figure 8 shows that for these three regions the valuation is also comparable, supporting this possible explanation.

Hypothesis 4

"The enforcement of relevant energy label policies has a positive significant impact on the economic value of energy labels."

The results of the analysis of model 4 are graphically displayed in Figure 9.

Figure 9 Effect of energy label policy change on value



The transactions that took place after the label enforcement policy of 2015 and onwards yields a significant positive valuation for the energy label classes B, C, D, E and F. For the A-label class the results are positive as well although the results are not significant. As the correlation matrix in Appendix 3 indicates the correlation between the transaction year and the label enforcement dummy is high (0.861). This might have to do with the upward trend in the housing market cycle that is visible from 2015 onwards, the energy label policy enforcement was introduced at a moment where the housing market was booming and has been since. The dummy represents the timeframe in which the enforcement is in place, as the time since than only showed increasing house prices this might affect the results of this analysis.

To determine whether there is a structural break in the dataset on the moment that the policy change took place a Chow test is performed (Chow, 1960). The dataset is split in two subsets with the help of a dummy variable, one set before the policy change and one after the policy change.

Equation 1 Chow test

$$F = \frac{(RSS_P - (RSS_{<2015} + RSS_{2015>}))/k}{\frac{(RSS_{<2015} + RSS_{2015>})}{(N_{<2015} + N_{2015>} - 2k)}}$$

With the Chow-test the following hypothesis is tested:

H_0: The parameters are constant over time
H_1: The parameters are not constant over time

Appendix 7 displays the results of the Chow test. On a 99% confidence interval the H_0-hypothesis is rejected. This indicates that the coefficients of the model are not equal in both subsamples. Concretely this impacts the results as it gives reason to believe that in the previous analysis the effects of the label energy enforcement are underestimated or overestimated.

The results indicate that Hypothesis 4 cannot be accepted or rejected.

The results imply that the valuation of energy labels increased after the enforcement came in place. It is not possible to conclude that it is because of the enforcement or that the change in valuation is caused by externalities as the Chow-test indicates that the coefficients in both subsamples differ significantly. Furthermore, not all interactions are significant and the results highly correlate with the transaction year.

This might indicate that the real estate cycle, which is indicated by the year, creates such a high demand for dwellings in general that the energy label is not a priority in the buying decision. The availability of a suitable dwelling of the right size might have a higher priority in the buying decision than the energy label of the dwelling.

7. Discussion & Conclusion

Conclusion

This study explored the valuation of energy labels for dwellings in the Netherlands. The data that was used is based on the province of Noord-Holland for the timeframe 2008-2018. The results indicate that significant value is attributed to the different energy label classes. The first research question, "*What is the impact of an energy label on the value of residential properties?"*, can empirically be answered as the discussion of the results of hypothesis 1 show. The energy labels attribute significant value to dwellings. These results are mainly in line with the previous literature although the difference between the green-label premiums is smaller than the previous literature indicates.

The second research question, "*What are the developments in economic value of energy labels in the period 2008-2018 in the province of Noord-Holland?"*, is answered in the discussion of hypothesis 2 and 3. The value of energy labels has changed significantly over the years, resulting in a slight increased value over the period 2008-2018. The valuation differs significantly over the different COROP-regions, in regions with lower transaction prices the valuation is higher than in regions with higher transactions prices.

The third research question, "*What impact do the relevant policies have on the value of energy labels?"*, is answered in the discussion of hypothesis 4. Since the energy label policy enforcement is in place the value of energy labels increased significantly. However, it is not completely certain if this is because of the enforcement and increased attention for energy labels or a general trend. The structural break in coefficients as indicated by the Chow-test supports this view. Further research with an enhanced methodology is needed to answer this question with more certainty.

Limitations & recommendations

As only the most recent energy label is registered no historic data is available, still a lot of dwellings do not get a registered energy label during a transaction. To account for this a future study could start early with the collection of the energy label data to generate historic data. This would enrich the number of included transactions.

Not all results of the interaction-analyses in this study are significant, both the interaction of the energy label and the different COROP-regions as well as the different years (model 3). As the energy-label results are significant when all the years and regions are combined (model 1) this might be because of the relatively small sample size that occur when splitting one province in multiple years and COROP-regions. Further research could be based on a larger dataset which might overcome this limitation.

Based on this study the exact effect of the policy change that occurred in 2015 is difficult to determine (model 4). As the policy changed for the whole dataset no control group is available, therefore the preferred difference-in-difference approach was therefore not possible. Further research could focus on a solution for this question.

APPENDICES:

Appendix 1 Included variables with categories of categorical variables

Type of characteristic	Variable	Description
Dwelling	Energy label	Categorical variable for the energy label class (A to G)
	Floor space	Livable area of floor space in m^2
	Number of rooms	Categorical variable for the number of rooms
	Garage	Dummy for garage (1 = yes)
	Garden	Dummy for garden (1 = yes)
	Balcony / Terrace	Dummy for balcony or terrace (1 = yes)
	Type of dwelling	Categorical variable of different types of dwelling
	Construction period	Categorical variable for the different specified construction periods.
Locational	COROP-region	Categorical variable of the COROP-region that the dwelling is situated in
	Address density	The density of addresses within a x km range from the location.
	Urbanity index	Categorical variable for the different urban density classes defined by the Central Statistics Bureau.
	Distance to closest hospital	Average distance in kilometer within municipality.
	Distance to closest large grocery store	Average distance in kilometer within municipality.
	Distance to closest restaurant	Average distance in kilometer within municipality.
	Distance to closest cinema	Average distance in kilometer within municipality.
	Distance to closest public swimming pool	Average distance in kilometer within municipality.
	Distance to closest freeway entrance	Average distance in kilometer within municipality.
	Distance to closest train station	Average distance in kilometer within municipality.
Transactional	Time of the transaction	Date of the sale
	Days listed	Days that the dwelling was listed on the market
	Transaction price	Nominal transaction price in €
	Relative transaction price discount	Relative difference between the original list price and the actual transaction price.

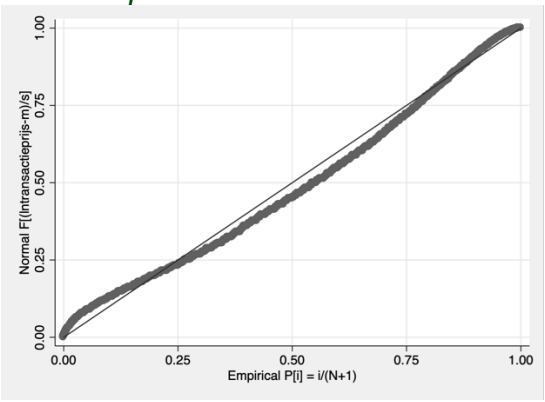
Categorical variable	Possible categories
Energy label	A (=1)
	B (=2)
	C (=3)
	D (=4) (base level)
	E (=5)
	F (=6)
	G (=8)
Type of dwelling	Terraced house (base level)
	Detached house
	Semi-detached house
	Corner house
	Staggered terraced house
Type of apartment	Ground floor apartment
	Regular apartment
	Multi-floor apartment
	Portico hall flat

	Gallery hall entrance flat
Construction period	< 1500-1905 1906-1930 1931-1944 1945-1959 1960-1970 1971-1980 1981-1990 1991-2000 2001> (base level)
Period of energy label enforcement	No (=0) Yes (=1)

Appendix 2 Model assumptions

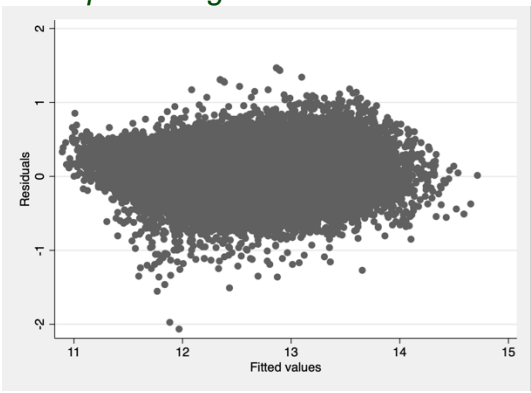
This appendix shows the results of the model assumption test, based on the preferred model 3.

Error term of the dependent variable Transaction price (log)



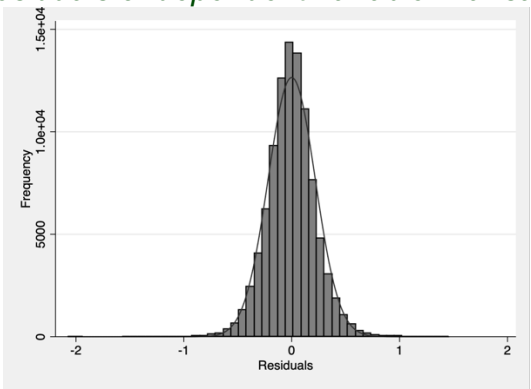
The error term approaches the average of 0 in the P-P plot.

Variance of residuals is equal along the values of x for Transaction price (log)



The variance of the residuals is approximately constant at all values of *x*.

Distribution of residuals of dependent variable Transaction price (log)



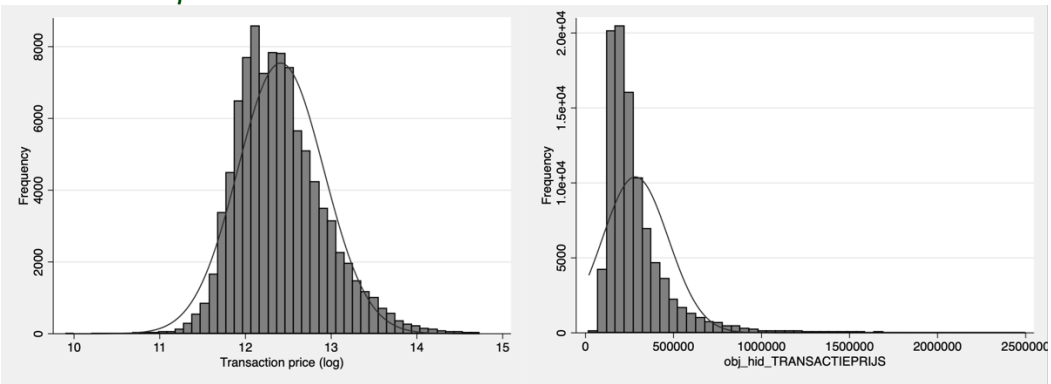
The residuals follow the normal distribution.

Appendix 3 Correlation Matrix

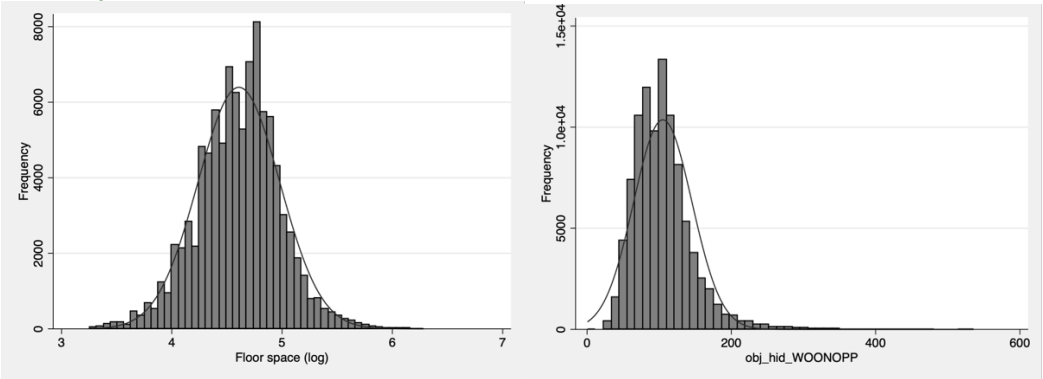
	Transaction price (log)	Energy label	Floor space (log)	Number of rooms	Construction period	Dwelling type	Roof terrace	Garden	Balcony	Parking space	Transaction year	Residence	Urbanity index	COROP-region	Label enforcement
Transaction price (log)	1														
Energy label	-0.067	1													
Floor space (log)	0.697	-0.089	1												
Number of rooms	0.546	0.041	0.782	1											
Construction period	0.047	-0.734	0.069	-0.076	1										
Dwelling type	-0.251	-0.043	-0.473	-0.562	0.142	1									
Roof terrace	0.136	-0.044	0.106	0.056	-0.005	-0.021	1								
Garden	0.170	0.044	0.394	0.465	-0.107	-0.745	0.003	1							
Balcony	-0.070	0.071	-0.218	-0.225	0.004	0.540	-0.098	-0.485	1						
Parking space	0.399	-0.167	0.408	0.254	0.197	-0.044	0.050	0.081	-0.051	1					
Transaction year	0.228	-0.054	0.137	0.123	0.044	-0.086	0.033	0.065	-0.052	0.102	1				
Residence	-0.088	0.005	0.144	0.153	0.033	-0.198	-0.015	0.209	-0.159	0.056	0.054	1			
Urbanity index	0.033	-0.043	0.213	0.186	0.097	-0.155	-0.061	0.165	-0.153	0.182	0.003	0.251	1		
COROP-region	0.144	0.047	-0.133	-0.125	-0.082	0.174	0.007	-0.184	0.185	-0.077	-0.036	-0.080	-0.167	1	
Label enforcement	0.243	-0.054	0.152	0.121	0.050	-0.074	0.036	0.056	-0.046	0.114	0.861	0.043	0.008	-0.031	1

Appendix 4 Distribution transformed variables

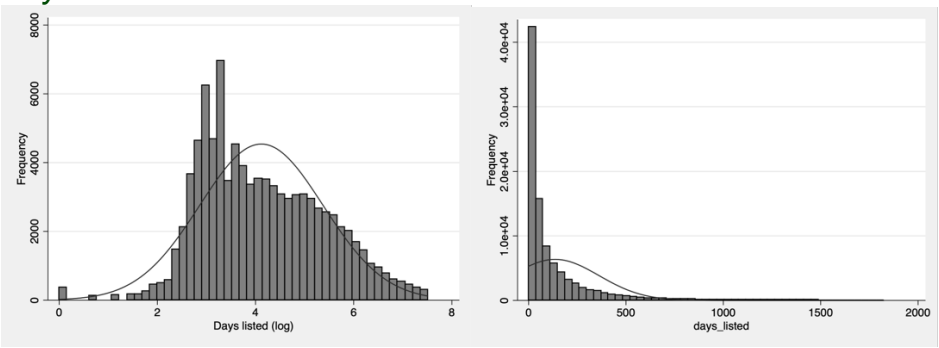
Transaction price



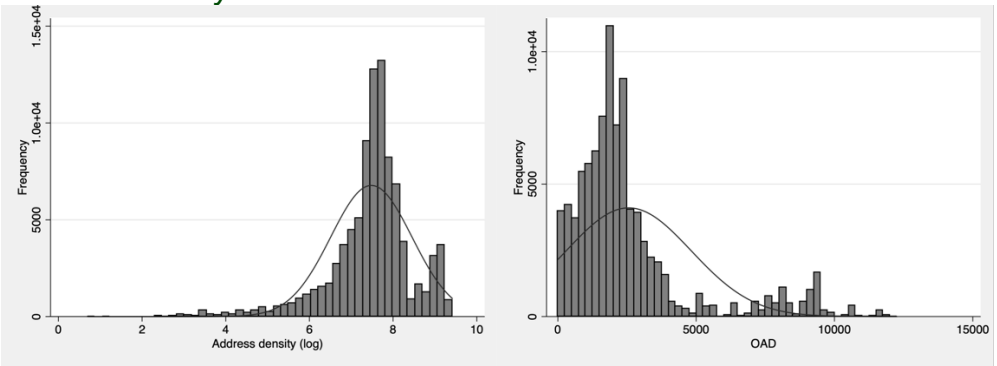
Floor space



Days listed



Address density



Appendix 5 Regression results

Variable	Model 1	se	Model 2	se	Model 3	se	Model 4	se
Energy label class								
A	0.1204***	(0.005)	0.045***	(0.011)	0.082***	(0.020)	0.118***	(0.007)
B	0.1203***	(0.004)	0.057***	(0.009)	0.013	(0.018)	0.088***	(0.006)
C	0.1037**	(0.004)	0.035***	(0.008)	-0.021	(0.015)	0.064***	(0.006)
D	0.0841***	(0.003)	0.018**	(0.008)	-0.030*	(0.016)	0.035***	(0.006)
E	0.0748***	(0.004)	0.015*	(0.008)	-0.039**	(0.016)	0.023***	(0.006)
F	0.0594***	(0.004)	0.005	(0.008)	-0.010	(0.017)	0.015**	(0.006)
G	-		-		-		-	
Floor space (log)	0.765***	(0.004)	0.776***	(0.008)	0.759***	(0.004)	0.764***	(0.004)
Rooms								
1	-0.222***	(0.019)	-0.365***	(0.042)	-0.210***	(0.019)	-0.222***	(0.019)
2	-0.262***	(0.017)	-0.389***	(0.039)	-0.247***	(0.017)	-0.262***	(0.017)
3	-0.279***	(0.017)	-0.441***	(0.038)	-0.262***	(0.017)	-0.278***	(0.017)
4	-0.284***	(0.017)	-0.467***	(0.038)	-0.266***	(0.017)	-0.282***	(0.017)
5	-0.251***	(0.017)	-0.446***	(0.037)	-0.235***	(0.016)	-0.249***	(0.017)
6	-0.208***	(0.017)	-0.374***	(0.038)	-0.193***	(0.016)	-0.207***	(0.017)
7	-0.147***	(0.017)	-0.299***	(0.038)	-0.134***	(0.017)	-0.145***	(0.017)
8	-0.096***	(0.018)	-0.265***	(0.041)	-0.085***	(0.018)	-0.094***	(0.018)
9	-0.082***	(0.019)	-0.192***	(0.044)	-0.075***	(0.019)	-0.080***	(0.019)
10	-		-		-			
Construction period								
1500-1905	0.130***	(0.005)	0.034***	(0.011)	0.124***	(0.005)	0.126***	(0.005)
1906-1930	0.033***	(0.005)	-0.042***	(0.010)	0.032***	(0.005)	0.029***	(0.005)
1931-1944	0.063***	(0.005)	-0.012	(0.011)	0.065***	(0.005)	0.060***	(0.005)
1945-1959	-0.073***	(0.005)	-0.126***	(0.010)	-0.075***	(0.005)	-0.077***	(0.005)
1960-1970	-0.178***	(0.004)	-0.218***	(0.009)	-0.182***	(0.004)	-0.181***	(0.004)
1971-1980	-0.171***	(0.004)	-0.186***	(0.009)	-0.173***	(0.004)	-0.173***	(0.004)
1981-1990	-0.110***	(0.004)	-0.140***	(0.008)	-0.112***	(0.004)	-0.113***	(0.004)
1991-2000	-0.021***	(0.003)	-0.041***	(0.008)	-0.023***	(0.003)	-0.023***	(0.003)
>2001	-		-		-		-	
Dwelling type								
Terraced house	0.088***	(0.004)	0.034***	(0.009)	0.090***	(0.004)	0.088***	(0.004)
Corner house	0.114***	(0.005)	0.063***	(0.010)	0.116***	(0.005)	0.113***	(0.005)
Semi-detached house	0.209***	(0.005)	0.169***	(0.011)	0.215***	(0.005)	0.208***	(0.005)
Detached house	0.360***	(0.005)	0.348***	(0.012)	0.373***	(0.005)	0.360***	(0.005)
Ground floor apartment	0.118***	(0.005)	0.061***	(0.010)	0.116***	(0.005)	0.117***	(0.005)
Regular apartment	0.119***	(0.004)	0.046***	(0.007)	0.118***	(0.004)	0.118***	(0.004)
Multi-floor apartment	-0.013**	(0.005)	-0.011	(0.010)	-0.012**	(0.005)	-0.013***	(0.005)
Portico hall flat	0.040***	(0.003)	-0.004	(0.007)	0.043***	(0.003)	0.040***	(0.003)
Gallery hall apartment	-		-		-		-	
Roof terrace = 1	0.033***	(0.003)	0.024***	(0.006)	0.033***	(0.003)	0.033***	(0.003)
Garden = 1	-0.005**	(0.002)	0.009*	(0.005)	-0.004	(0.002)	-0.005**	(0.002)
Balcony = 1	0.031***	(0.002)	0.029***	(0.004)	0.031***	(0.002)	0.031***	(0.002)
Parking space = 1	0.108***	(0.002)	0.124***	(0.005)	0.107***	(0.002)	0.108***	(0.002)
Transaction year								
2009	-0.041***	(0.004)	-0.062***	(0.011)	-0.048***	(0.018)	-0.042***	(0.004)
2010	-0.034***	(0.004)	-0.052***	(0.011)	-0.052***	(0.019)	-0.035***	(0.004)
2011	-0.038***	(0.005)			0.020	(0.020)	-0.039***	(0.005)
2012	-0.108***	(0.005)			-0.144***	(0.020)	-0.109***	(0.005)
2013	-0.155***	(0.005)			-0.228***	(0.020)	-0.155***	(0.005)
2014	-0.103***	(0.004)			-0.117***	(0.016)	-0.105***	(0.004)
2015	-0.040***	(0.003)			-0.103***	(0.014)	-0.091***	(0.007)
2016	0.033***	(0.003)			-0.009	(0.014)	-0.017***	(0.007)
2017	0.146***	(0.003)			0.088***	(0.014)	0.096***	(0.007)
2018	0.226***	(0.004)			0.157***	(0.016)	0.176***	(0.007)
Urbanity index								
Very highly urban	0.110***	(0.004)			0.113***	(0.003)	0.110***	(0.004)
Highly urban	0.023***	(0.003)	-0.352***	(0.014)	0.024***	(0.003)	0.023***	(0.003)

Moderately urban	0.038***	(0.004)	-0.260***	(0.012)	0.037***	(0.004)	0.038***	(0.004)
Mildly urban	0.037***	(0.004)	-0.196***	(0.012)	0.036***	(0.004)	0.037***	(0.004)
Not urban	-		-		-		-	
Interaction year#energy label class								
2008#A					0.000	(0.000)		
2008#B					0.000	(0.000)		
2008#C					0.000	(0.000)		
2008#D					0.000	(0.000)		
2008#E					0.000	(0.000)		
2008#F					0.000	(0.000)		
2008#G					0.000	(0.000)		
2009#A					0.015	(0.023)		
2009#B					0.009	(0.021)		
2009#C					0.021	(0.020)		
2009#D					0.004	(0.021)		
2009#E					-0.003	(0.021)		
2009#F					-0.024	(0.023)		
2009#G					0.000	(0.000)		
2010#A					0.027	(0.024)		
2010#B					0.008	(0.022)		
2010#C					0.026	(0.020)		
2010#D					0.016	(0.021)		
2010#E					0.027	(0.022)		
2010#F					-0.012	(0.023)		
2010#G					0.000	(0.000)		
2011#A					-0.051**	(0.025)		
2011#B					-0.063***	(0.023)		
2011#C					-0.066***	(0.021)		
2011#D					-0.045**	(0.022)		
2011#E					-0.074***	(0.023)		
2011#F					-0.073***	(0.025)		
2011#G					0.000	(0.000)		
2012#A					0.042	(0.025)		
2012#B					0.014	(0.024)		
2012#C					0.050**	(0.022)		
2012#D					0.039*	(0.023)		
2012#E					0.032	(0.023)		
2012#F					0.014	(0.025)		
2012#G					0.000	(0.000)		
2013#A					0.118***	(0.026)		
2013#B					0.044*	(0.024)		
2013#C					0.086***	(0.022)		
2013#D					0.065***	(0.023)		
2013#E					0.089***	(0.023)		
2013#F					0.068***	(0.026)		
2013#G					0.000	(0.000)		
2014#A					0.010	(0.021)		
2014#B					-0.002	(0.019)		
2014#C					0.008	(0.018)		
2014#D					0.007	(0.019)		
2014#E					0.038**	(0.020)		
2014#F					0.028	(0.021)		
2014#G					0.000	(0.000)		
2015#A					0.041**	(0.018)		
2015#B					0.053***	(0.016)		
2015#C					0.069***	(0.015)		
2015#D					0.061***	(0.016)		
2015#E					0.084***	(0.016)		
2015#F					0.062***	(0.017)		
2015#G					0.000	(0.000)		
2016#A					-0.003	(0.018)		
2016#B					0.017	(0.016)		
2016#C					0.044***	(0.015)		
2016#D					0.056***	(0.016)		
2016#E					0.080***	(0.016)		

2016#F	0.060***	(0.017)
2016#G	0.000	(0.000)
2017#A	0.014	(0.018)
2017#B	0.036**	(0.017)
2017#C	0.067***	(0.015)
2017#D	0.092***	(0.016)
2017#E	0.082***	(0.017)
2017#F	0.054***	(0.018)
2017#G	0.000	(0.000)
2018#A	0.027	(0.020)
2018#B	0.031*	(0.019)
2018#C	0.074***	(0.017)
2018#D	0.127***	(0.019)
2018#E	0.109***	(0.019)
2018#F	0.060***	(0.020)
2018#G	0.000	(0.000)
Interaction COROP-region#energy label class		
Kop van Noord-Holland#A	0.138***	(0.017)
Kop van Noord-Holland#B	0.186***	(0.015)
Kop van Noord-Holland#C	0.162***	(0.013)
Kop van Noord-Holland#D	0.130***	(0.014)
Kop van Noord-Holland#E	0.153***	(0.016)
Kop van Noord-Holland#F	0.048***	(0.016)
Kop van Noord-Holland#G	0.000	(0.000)
Alkmaar en omgeving#A	0.068***	(0.016)
Alkmaar en omgeving#B	0.120***	(0.014)
Alkmaar en omgeving#C	0.104***	(0.012)
Alkmaar en omgeving#D	0.095***	(0.013)
Alkmaar en omgeving#E	0.085***	(0.014)
Alkmaar en omgeving#F	0.069***	(0.014)
Alkmaar en omgeving#G	0.000	(0.000)
IJmond#A	0.088***	(0.016)
IJmond#B	0.155***	(0.015)
IJmond#C	0.164***	(0.012)
IJmond#D	0.109***	(0.013)
IJmond#E	0.075***	(0.013)
IJmond#F	0.029**	(0.013)
IJmond#G	0.000	(0.000)
Agglomeratie Haarlem#A	-0.043***	(0.014)
Agglomeratie Haarlem#B	0.012	(0.013)
Agglomeratie Haarlem#C	0.024**	(0.010)
Agglomeratie Haarlem#D	-0.007	(0.011)
Agglomeratie Haarlem#E	0.003	(0.011)
Agglomeratie Haarlem#F	-0.014	(0.011)
Agglomeratie Haarlem#G	0.000	(0.000)
Zaanstreek#A	0.208***	(0.016)
Zaanstreek#B	0.324***	(0.015)
Zaanstreek#C	0.331***	(0.012)
Zaanstreek#D	0.258***	(0.013)
Zaanstreek#E	0.237***	(0.014)
Zaanstreek#F	0.106***	(0.014)
Zaanstreek#G	0.000	(0.000)
Groot-Amsterdam#A	-0.048***	(0.013)
Groot-Amsterdam#B	0.017	(0.012)
Groot-Amsterdam#C	0.008	(0.010)
Groot-Amsterdam#D	0.015	(0.010)
Groot-Amsterdam#E	0.001	(0.010)
Groot-Amsterdam#F	0.013	(0.011)
Groot-Amsterdam#G	0.000	(0.000)
Gooi en Vechtstreek#A	0.000	(0.000)
Gooi en Vechtstreek#B	0.000	(0.000)
Gooi en Vechtstreek#C	0.000	(0.000)
Gooi en Vechtstreek#D	0.000	(0.000)
Gooi en Vechtstreek#E	0.000	(0.000)
Gooi en Vechtstreek#F	0.000	(0.000)

Gooi en Vechtstreek#G					0.000	(0.000)		
Interaction label enforcement#energy label class								
0#A							-	
0#B							-	
0#C							-	
0#D							-	
0#E							-	
0#F							-	
0#G							-	
1#A							0.001	(0.008)
1#B							0.041***	(0.007)
1#C							0.053***	(0.007)
1#D							0.071***	(0.007)
1#E							0.075***	(0.007)
1#F							0.063***	(0.008)
1#G							0.000	(0.000)
Address density (log)			0.142***	(0.005)				
Distance to hospital (log)			-0.001**	(0.000)				
Distance to large supermarket (log)			0.003	(0.003)				
Distance to public swimming pool (log)			-0.001	(0.001)				
Distance to freeway entrance (log)			-0.006*	(0.003)				
Distance to train station (log)			-0.004***	(0.001)				
Constant	8.861***	(0.029)	8.536***	(0.075)	8.957***	(0.031)	8.904***	(0.029)
Observations	96,308		15,614		96,308		96,306	
R-squared	0.809		0.837		0.813		0.810	
Address density	No		Yes		No		No	
Distance to amenities	No		Yes		No		No	
Number of Rooms	Yes		Yes		Yes		Yes	
Municipality	Yes		Yes		Yes		Yes	
Year * Energy Label interactions	No		No		Yes		No	
COROP-Region * Energy Label interactions	No		No		Yes		No	
Energy Label policy enforcement interaction	No		No		No		Yes	
Adjusted R-squared	0.809		0.835		0.813		0.809	

Standard errors in parentheses

No base levels have been manually set in the regression, as a result some levels of categorical variables are omitted because of collinearity, this is indicated with "–" in the coefficient column.

*** p<0.01, ** p<0.05, * p<0.1

Appendix 6 STATA Do-file syntax

```
**NVM data preparation*

import delimited "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/NVM.csv",
delimiter(";") encoding(ISO-8859-1)clear
label variable ïobj_hid_straatnaam "obj_hid_STRAATNAAM"
rename ïobj_hid_straatnaam obj_hid_straatnaam
rename obj_hid_postcode obj_hid_postcode_spatie
rename obj_hid_postcodespatieloos obj_hid_postcode
label variable obj_hid_postcode "obj_hid_POSTCODE"
label variable obj_hid_postcode_spatie "obj_hid_POSTCODE_SPATIE"
drop obj_hid_postcode_spatie


**Prepare house number to be merged with RVO house number*

tostring obj_hid_huisnummer, generate(obj_hid_huisnummer_string)
rename obj_hid_huisnummer obj_hid_huisnummer_int
rename obj_hid_huisnummer_string obj_hid_huisnummer
save "/Users/Cassee/Documents/Master Thesis/STATA/NVM.dta", replace


*NVM & RVO data merge*

use "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/RVO.dta"
merge m:m obj_hid_huisnummer obj_hid_postcode obj_hid_huisnummertoevoeging using
"/Users/Cassee/Documents/Master Thesis/STATA/NVM.dta"
save "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/NVM RVO.dta", replace
use "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/NVM RVO.dta"
rename _merge _mergervo
gen jaar = substr(obj_hid_datum_afmelding,6,9)
recast str4 jaar, force
destring jaar, replace
destring obj_buurt_id, replace
destring obj_wijk_id, replace
keep if _mergervo==3
save "/Users/Cassee/Documents/Master Thesis/STATA/NVM RVO.dta", replace

*CBS Kerncijfers preparation*
*Format and variables are the same between 2008-2012, 2013-2016, 2017 and 2018*
*Keep only the relevant variables*

*2008*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2008.xls",
sheet("KWB2008_versie20111123") firstrow case(lower) clear
keep obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb af_oprith
af_treinst
generate int jaar = 2008
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2008.dta", replace

*2009*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2009.xls",
sheet("kwb2009") firstrow case(lower) clear
keep obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb af_oprith
af_treinst
generate int jaar = 2009
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2009.dta", replace

*2010*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2010.xls",
sheet("KWB_2010_versie_december2013") firstrow case(lower) clear
keep obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb af_oprith
af_treinst
generate int jaar = 2010
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
```

```
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2010.dta", replace

*2011*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2011.xls",
sheet("_KWB2011") firstrow case(lower) clear
keep obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb af_oprith
af_treinst
generate int jaar = 2011
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
tostring af_ondbas, replace
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2011.dta", replace

*2012*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2012.xls",
sheet("KWB2012") firstrow case(lower) clear
keep obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb af_oprith
generate int jaar = 2012
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2012.dta", replace

*2013*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2013.xls",
sheet("KWB2013") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2013
rename ste_oad oad
rename ste_mvs sted
tostring sted, replace

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
destring oad, replace force

save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2013.dta", replace

*2014*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2014.xls",
sheet("KWB2014") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2014
rename ste_oad oad
rename ste_mvs sted

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)

*destring obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb
af_oprith, replace
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2014.dta", replace

*2015*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2015.xls",
sheet("_KWB2015") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2015
rename ste_oad oad
rename ste_mvs sted

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)
```

```
*destring obj_buurt_id oad sted af_ziek_e af_ondbas af_restau af_superm af_zwemb
af_oprith, replace
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2015.dta", replace

*2016*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2016.xls",
sheet("KWB2016") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2016
rename ste_oad oad
rename ste_mvs sted

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)

save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2016.dta", replace

*2017*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2017.xls",
sheet("_KWB2017") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2017
rename ste_oad oad
rename ste_mvs sted

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)

save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2017.dta", replace

*2018*
import excel "/Users/Cassee/Documents/Master Thesis/STATA/Origineel/kwb-2018.xls",
sheet("KWB2018") firstrow case(lower) clear
keep obj_buurt_id ste_oad ste_mvs g_afs_hp g_afs_gs g_afs_sc
generate int jaar = 2018
rename ste_oad oad
rename ste_mvs sted

rename g_afs_hp af_huisa
rename g_afs_gs af_superm
rename g_afs_sc af_ondbas
tostring obj_buurt_id, generate(obj_buurt_id_str)
recast str5 obj_buurt_id_str, force
destring obj_buurt_id_str, generate(obj_wijk_id)

save "/Users/Cassee/Documents/Master Thesis/STATA/CBS2018.dta", replace


*CBS Kerncijfers merge*
use "/Users/Cassee/Documents/Master Thesis/STATA/CBS2008.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2009.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2010.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2011.dta", force
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2012.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2013.dta", force
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2014.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2015.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2016.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2017.dta"
append using "/Users/Cassee/Documents/Master Thesis/STATA/CBS2018.dta"
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS.dta", replace



*CBS and NVM+RVO merge*
use "/Users/Cassee/Documents/Master Thesis/STATA/NVM RVO.dta"
merge m:m jaar obj_wijk_id using "/Users/Cassee/Documents/Master Thesis/STATA/CBS.dta",
```

```
generate(_mergecbs)
keep if _mergecbs==3
keep if _mergervo==3
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS NVM RVO.dta", replace

**COROP-regions*

use "/Users/Cassee/Documents/Master Thesis/STATA/CBS NVM RVO.dta"
gen coropcode =.
replace coropcode =24 if obj_hid_woonplaats=="'S-GRAVELAND"
replace coropcode =18 if obj_hid_woonplaats=="'T VELD"
replace coropcode =18 if obj_hid_woonplaats=="'T ZAND"
replace coropcode =18 if obj_hid_woonplaats=="'T ZAND NH"
replace coropcode =23 if obj_hid_woonplaats=="AALSMEER"
replace coropcode =23 if obj_hid_woonplaats=="AALSMEERDERBRUG"
replace coropcode =18 if obj_hid_woonplaats=="AARTSWOUD"
replace coropcode =18 if obj_hid_woonplaats=="ABBEKERK"
replace coropcode =23 if obj_hid_woonplaats=="ABBENES"
replace coropcode =21 if obj_hid_woonplaats=="AERDENHOUT"
replace coropcode =20 if obj_hid_woonplaats=="AKERSLOOT"
replace coropcode =19 if obj_hid_woonplaats=="ALKMAAR"
replace coropcode =23 if obj_hid_woonplaats=="AMSTELVEEN"
replace coropcode =23 if obj_hid_woonplaats=="AMSTERDAM"
replace coropcode =23 if obj_hid_woonplaats=="AMSTERDAM ZUIDOOST"
replace coropcode =18 if obj_hid_woonplaats=="ANDIJK"
replace coropcode =24 if obj_hid_woonplaats=="ANKEVEEN"
replace coropcode =18 if obj_hid_woonplaats=="ANNA PAULOWNA"
replace coropcode =22 if obj_hid_woonplaats=="ASSENDELFT"
replace coropcode =18 if obj_hid_woonplaats=="AVENHORN"
replace coropcode =23 if obj_hid_woonplaats=="BADHOEVEDORP"
replace coropcode =18 if obj_hid_woonplaats=="BARSINGERHORN"
replace coropcode =23 if obj_hid_woonplaats=="BEETS NH"
replace coropcode =23 if obj_hid_woonplaats=="BEINSDORP"
replace coropcode =21 if obj_hid_woonplaats=="BENNEBROEK"
replace coropcode =18 if obj_hid_woonplaats=="BENNINGBROEK"
replace coropcode =21 if obj_hid_woonplaats=="BENTVELD"
replace coropcode =19 if obj_hid_woonplaats=="BERGEN (NH)"
replace coropcode =19 if obj_hid_woonplaats=="BERGEN AAN ZEE"
replace coropcode =19 if obj_hid_woonplaats=="BERGEN NH"
replace coropcode =18 if obj_hid_woonplaats=="BERKHOUT"
replace coropcode =20 if obj_hid_woonplaats=="BEVERWIJK"
replace coropcode =24 if obj_hid_woonplaats=="BLARICUM"
replace coropcode =21 if obj_hid_woonplaats=="BLOEMENDAAL"
replace coropcode =18 if obj_hid_woonplaats=="BLOKKER"
replace coropcode =23 if obj_hid_woonplaats=="BOESINGHELIEDE"
replace coropcode =18 if obj_hid_woonplaats=="BOVENKARSPEL"
replace coropcode =18 if obj_hid_woonplaats=="BREEZAND"
replace coropcode =24 if obj_hid_woonplaats=="BREUKELEVEEN"
replace coropcode =23 if obj_hid_woonplaats=="BROEK IN WATERLAND"
replace coropcode =19 if obj_hid_woonplaats=="BROEK OP LANGEDIJK"
replace coropcode =23 if obj_hid_woonplaats=="BUITENKAAG"
replace coropcode =18 if obj_hid_woonplaats=="BURGERBRUG"
replace coropcode =23 if obj_hid_woonplaats=="BURGERVEEN"
replace coropcode =24 if obj_hid_woonplaats=="BUSSUM"
replace coropcode =23 if obj_hid_woonplaats=="Beets"
replace coropcode =18 if obj_hid_woonplaats=="CALLANTSOOG"
replace coropcode =20 if obj_hid_woonplaats=="CASTRICUM"
replace coropcode =23 if obj_hid_woonplaats=="CRUQUIUS"
replace coropcode =18 if obj_hid_woonplaats=="DE COCKSDORP"
replace coropcode =18 if obj_hid_woonplaats=="DE GOORN"
replace coropcode =18 if obj_hid_woonplaats=="DE KOOG"
replace coropcode =23 if obj_hid_woonplaats=="DE KWAKEL"
replace coropcode =19 if obj_hid_woonplaats=="DE RIJP"
replace coropcode =18 if obj_hid_woonplaats=="DE WAAL"
replace coropcode =18 if obj_hid_woonplaats=="DE WEERE"
replace coropcode =20 if obj_hid_woonplaats=="DE WOUDE"
replace coropcode =18 if obj_hid_woonplaats=="DEN BURG"
replace coropcode =18 if obj_hid_woonplaats=="DEN HELDER"
replace coropcode =18 if obj_hid_woonplaats=="DEN HOORN"
replace coropcode =18 if obj_hid_woonplaats=="DEN HOORN TEXEL"
replace coropcode =23 if obj_hid_woonplaats=="DEN ILP"
replace coropcode =18 if obj_hid_woonplaats=="DEN OEVER"
replace coropcode =23 if obj_hid_woonplaats=="DIEMEN"
replace coropcode =18 if obj_hid_woonplaats=="DIRKSHORN"
replace coropcode =20 if obj_hid_woonplaats=="DRIEHUIS NH"
replace coropcode =19 if obj_hid_woonplaats=="DRIEHUIZEN"
```

```
replace coropcode =23 if obj_hid_woonplaats=="DUIVENDRECHT"
replace coropcode =23 if obj_hid_woonplaats=="EDAM"
replace coropcode =19 if obj_hid_woonplaats=="EGMOND AAN DEN HOEF"
replace coropcode =19 if obj_hid_woonplaats=="EGMOND AAN ZEE"
replace coropcode =19 if obj_hid_woonplaats=="EGMOND-BINNEN"
replace coropcode =18 if obj_hid_woonplaats=="ENKHUIZEN"
replace coropcode =19 if obj_hid_woonplaats=="GRAFT"
replace coropcode =19 if obj_hid_woonplaats=="GROET"
replace coropcode =18 if obj_hid_woonplaats=="GROOTEBROEK"
replace coropcode =19 if obj_hid_woonplaats=="GROOTSCHERMER"
replace coropcode =21 if obj_hid_woonplaats=="HAARLEM"
replace coropcode =21 if obj_hid_woonplaats=="HAARLEMMERLIEDE"
replace coropcode =21 if obj_hid_woonplaats=="HALFWEG NH"
replace coropcode =18 if obj_hid_woonplaats=="HAUWERT"
replace coropcode =20 if obj_hid_woonplaats=="HEEMSKERK"
replace coropcode =21 if obj_hid_woonplaats=="HEEMSTEDE"
replace coropcode =19 if obj_hid_woonplaats=="HEERHUGOWAARD"
replace coropcode =19 if obj_hid_woonplaats=="HEILOO"
replace coropcode =18 if obj_hid_woonplaats=="HEM"
replace coropcode =18 if obj_hid_woonplaats=="HENSBROEK"
replace coropcode =24 if obj_hid_woonplaats=="HILVERSUM"
replace coropcode =18 if obj_hid_woonplaats=="HIPPOLYTUSHOEF"
replace coropcode =23 if obj_hid_woonplaats=="HOBREDE"
replace coropcode =21 if obj_hid_woonplaats=="HOOFDDORP"
replace coropcode =18 if obj_hid_woonplaats=="HOOGKARSPEL"
replace coropcode =18 if obj_hid_woonplaats=="HOOGWOUD"
replace coropcode =18 if obj_hid_woonplaats=="HOORN"
replace coropcode =18 if obj_hid_woonplaats=="HOORN NH"
replace coropcode =24 if obj_hid_woonplaats=="HUIZEN"
replace coropcode =21 if obj_hid_woonplaats=="Halfweg"
replace coropcode =20 if obj_hid_woonplaats=="IJMUIDEN"
replace coropcode =23 if obj_hid_woonplaats=="ILPENDAM"
replace coropcode =22 if obj_hid_woonplaats=="JISP"
replace coropcode =18 if obj_hid_woonplaats=="JULIANADORP"
replace coropcode =23 if obj_hid_woonplaats=="KATWOUDE"
replace coropcode =19 if obj_hid_woonplaats=="KOEDIJK"
replace coropcode =18 if obj_hid_woonplaats=="KOLHORN"
replace coropcode =22 if obj_hid_woonplaats=="KOOG AAN DE ZAAN"
replace coropcode =24 if obj_hid_woonplaats=="KORTENHOEF"
replace coropcode =18 if obj_hid_woonplaats=="KREILEROORD"
replace coropcode =22 if obj_hid_woonplaats=="KROMMENIE"
replace coropcode =23 if obj_hid_woonplaats=="KUDELSTAART"
replace coropcode =23 if obj_hid_woonplaats=="KWADIJK"
replace coropcode =18 if obj_hid_woonplaats=="LAMBERTSCHAAG"
replace coropcode =23 if obj_hid_woonplaats=="LANDSMEER"
replace coropcode =24 if obj_hid_woonplaats=="LAREN"
replace coropcode =24 if obj_hid_woonplaats=="LAREN NH"
replace coropcode =23 if obj_hid_woonplaats=="LEIMUIDERBRUG"
replace coropcode =23 if obj_hid_woonplaats=="LIJNDEN"
replace coropcode =20 if obj_hid_woonplaats=="LIMMEN"
replace coropcode =23 if obj_hid_woonplaats=="LISSERBROEK"
replace coropcode =24 if obj_hid_woonplaats=="LOOSDRECHT"
replace coropcode =18 if obj_hid_woonplaats=="LUTJEBROEK"
replace coropcode =18 if obj_hid_woonplaats=="LUTJEWINKEL"
replace coropcode =23 if obj_hid_woonplaats=="MARKEN"
replace coropcode =23 if obj_hid_woonplaats=="MARKENBINNEN"
replace coropcode =18 if obj_hid_woonplaats=="MEDEMBLIK"
replace coropcode =23 if obj_hid_woonplaats=="MIDDELIE"
replace coropcode =23 if obj_hid_woonplaats=="MIDDENBEEMSTER"
replace coropcode =18 if obj_hid_woonplaats=="MIDDENMEER"
replace coropcode =18 if obj_hid_woonplaats=="MIDWOUD"
replace coropcode =23 if obj_hid_woonplaats=="MONNICKENDAM"
replace coropcode =24 if obj_hid_woonplaats=="MUIDEN"
replace coropcode =24 if obj_hid_woonplaats=="MUIDERBERG"
replace coropcode =24 if obj_hid_woonplaats=="NAARDEN"
replace coropcode =24 if obj_hid_woonplaats=="NEDERHORST DEN BERG"
replace coropcode =18 if obj_hid_woonplaats=="NIBBIXWOUD"
replace coropcode =23 if obj_hid_woonplaats=="NIEUW VENNEP"
replace coropcode =23 if obj_hid_woonplaats=="NIEUW-VENNEP"
replace coropcode =18 if obj_hid_woonplaats=="NIEUWE NIEDORP"
replace coropcode =19 if obj_hid_woonplaats=="NOORD-SCHARWOUDE"
replace coropcode =23 if obj_hid_woonplaats=="NOORDBEEMSTER"
replace coropcode =19 if obj_hid_woonplaats=="NOORDEINDE"
replace coropcode =18 if obj_hid_woonplaats=="OBDAM"
replace coropcode =19 if obj_hid_woonplaats=="OOST-GRAFTDIJK"
replace coropcode =18 if obj_hid_woonplaats=="OOSTERBLOKKER"
```

```
replace coropcode =18 if obj_hid_woonplaats=="OOSTEREND"
replace coropcode =18 if obj_hid_woonplaats=="OOSTEREND NH"
replace coropcode =18 if obj_hid_woonplaats=="OOSTERLEEK"
replace coropcode =23 if obj_hid_woonplaats=="OOSTHUIZEN"
replace coropcode =22 if obj_hid_woonplaats=="OOSTKNOLLENDAM"
replace coropcode =18 if obj_hid_woonplaats=="OOSTWOUD"
replace coropcode =23 if obj_hid_woonplaats=="OOSTZAAN"
replace coropcode =18 if obj_hid_woonplaats=="OPMEER"
replace coropcode =18 if obj_hid_woonplaats=="OPPERDOES"
replace coropcode =19 if obj_hid_woonplaats=="OTERLEEK"
replace coropcode =23 if obj_hid_woonplaats=="OUDE MEER"
replace coropcode =18 if obj_hid_woonplaats=="OUDE NIEDORP"
replace coropcode =18 if obj_hid_woonplaats=="OUDENDIJK"
replace coropcode =18 if obj_hid_woonplaats=="OUDENDIJK NH"
replace coropcode =23 if obj_hid_woonplaats=="OUDERKERK AAN DE AMSTEL"
replace coropcode =18 if obj_hid_woonplaats=="OUDESCHILD"
replace coropcode =18 if obj_hid_woonplaats=="OUDESLUIS"
replace coropcode =19 if obj_hid_woonplaats=="OUDKARSPEL"
replace coropcode =19 if obj_hid_woonplaats=="OUDORP"
replace coropcode =19 if obj_hid_woonplaats=="OUDORP NH"
replace coropcode =21 if obj_hid_woonplaats=="OVERVEEN"
replace coropcode =18 if obj_hid_woonplaats=="PETTEN"
replace coropcode =23 if obj_hid_woonplaats=="PURMER"
replace coropcode =23 if obj_hid_woonplaats=="PURMEREND"
replace coropcode =23 if obj_hid_woonplaats=="PURMERLAND"
replace coropcode =23 if obj_hid_woonplaats=="RIJSENHOUT"
replace coropcode =23 if obj_hid_woonplaats=="ROZENBURG"
replace coropcode =20 if obj_hid_woonplaats=="SANTPOORT-NOORD"
replace coropcode =20 if obj_hid_woonplaats=="SANTPOORT-ZUID"
replace coropcode =18 if obj_hid_woonplaats=="SCHAGEN"
replace coropcode =18 if obj_hid_woonplaats=="SCHAGERBRUG"
replace coropcode =23 if obj_hid_woonplaats=="SCHARDAM"
replace coropcode =18 if obj_hid_woonplaats=="SCHARWOUDE"
replace coropcode =18 if obj_hid_woonplaats=="SCHELLINKHOUT"
replace coropcode =19 if obj_hid_woonplaats=="SCHERMERHORN"
replace coropcode =19 if obj_hid_woonplaats=="SCHOORL"
replace coropcode =18 if obj_hid_woonplaats=="SIJBEKARSPEL"
replace coropcode =18 if obj_hid_woonplaats=="SINT MAARTEN"
replace coropcode =18 if obj_hid_woonplaats=="SINT MAARTENSBRUG"
replace coropcode =18 if obj_hid_woonplaats=="SINT MAARTENSVLOTBRUG"
replace coropcode =18 if obj_hid_woonplaats=="SINT PANCRAS"
replace coropcode =18 if obj_hid_woonplaats=="SLOOTDORP"
replace coropcode =21 if obj_hid_woonplaats=="SPAARNDAM"
replace coropcode =18 if obj_hid_woonplaats=="SPANBROEK"
replace coropcode =18 if obj_hid_woonplaats=="SPIERDIJK"
replace coropcode =22 if obj_hid_woonplaats=="SPIJKERBOOR"
replace coropcode =22 if obj_hid_woonplaats=="SPIJKERBOOR NH"
replace coropcode =19 if obj_hid_woonplaats=="STARNMEER"
replace coropcode =19 if obj_hid_woonplaats=="STOMPETOREN"
replace coropcode =21 if obj_hid_woonplaats=="Spaarndam Gem. Haarlem"
replace coropcode =18 if obj_hid_woonplaats=="TUITJENHORN"
replace coropcode =18 if obj_hid_woonplaats=="TWISK"
replace coropcode =23 if obj_hid_woonplaats=="UITDAM"
replace coropcode =20 if obj_hid_woonplaats=="UITGEEST"
replace coropcode =23 if obj_hid_woonplaats=="UITHOORN"
replace coropcode =18 if obj_hid_woonplaats=="URSEM"
replace coropcode =20 if obj_hid_woonplaats=="VELSEN-NOORD"
replace coropcode =20 if obj_hid_woonplaats=="VELSEN-ZUID"
replace coropcode =20 if obj_hid_woonplaats=="VELSERBROEK"
replace coropcode =18 if obj_hid_woonplaats=="VENHUIZEN"
replace coropcode =23 if obj_hid_woonplaats=="VIJFHUIZEN"
replace coropcode =21 if obj_hid_woonplaats=="VOGELENZANG"
replace coropcode =23 if obj_hid_woonplaats=="VOLENDAM"
replace coropcode =18 if obj_hid_woonplaats=="WAARLAND"
replace coropcode =23 if obj_hid_woonplaats=="WARDER"
replace coropcode =18 if obj_hid_woonplaats=="WARMENHUIZEN"
replace coropcode =23 if obj_hid_woonplaats=="WATERGANG"
replace coropcode =24 if obj_hid_woonplaats=="WEESP"
replace coropcode =18 if obj_hid_woonplaats=="WERVERSHOOF"
replace coropcode =19 if obj_hid_woonplaats=="WEST GRAFTDIJK"
replace coropcode =19 if obj_hid_woonplaats=="WEST-GRAFTDIJK"
replace coropcode =23 if obj_hid_woonplaats=="WESTBEEMSTER"
replace coropcode =18 if obj_hid_woonplaats=="WESTERLAND"
replace coropcode =22 if obj_hid_woonplaats=="WESTKNOLLENDAM"
replace coropcode =18 if obj_hid_woonplaats=="WESTWOUD"
replace coropcode =22 if obj_hid_woonplaats=="WESTZAAN"
```

```
replace coropcode =23 if obj_hid_woonplaats=="WETERINGBRUG"
replace coropcode =18 if obj_hid_woonplaats=="WIERINGERWAARD"
replace coropcode =18 if obj_hid_woonplaats=="WIERINGERWERF"
replace coropcode =18 if obj_hid_woonplaats=="WIJDENES"
replace coropcode =22 if obj_hid_woonplaats=="WIJDEWORMER"
replace coropcode =20 if obj_hid_woonplaats=="WIJK AAN ZEE"
replace coropcode =18 if obj_hid_woonplaats=="WINKEL"
replace coropcode =18 if obj_hid_woonplaats=="WOGNUM"
replace coropcode =22 if obj_hid_woonplaats=="WORMER"
replace coropcode =22 if obj_hid_woonplaats=="WORMERVEER"
replace coropcode =22 if obj_hid_woonplaats=="ZAANDAM"
replace coropcode =22 if obj_hid_woonplaats=="ZAANDIJK"
replace coropcode =21 if obj_hid_woonplaats=="ZANDVOORT"
replace coropcode =18 if obj_hid_woonplaats=="ZIJDEWIND"
replace coropcode =18 if obj_hid_woonplaats=="ZUID-SCHARWOUDE"
replace coropcode =18 if obj_hid_woonplaats=="ZUIDERMEER"
replace coropcode =23 if obj_hid_woonplaats=="ZUIDERWOUDE"
replace coropcode =23 if obj_hid_woonplaats=="ZUIDOOSTBEEMSTER"
replace coropcode =19 if obj_hid_woonplaats=="ZUIDSCHERMER"
replace coropcode =18 if obj_hid_woonplaats=="ZWAAG"
replace coropcode =18 if obj_hid_woonplaats=="ZWAAGDIJK"
replace coropcode =18 if obj_hid_woonplaats=="ZWAAGDIJK-OOST"
replace coropcode =18 if obj_hid_woonplaats=="ZWAAGDIJK-WEST"
replace coropcode =23 if obj_hid_woonplaats=="ZWAANSHOEK"
replace coropcode =23 if obj_hid_woonplaats=="ZWANENBURG"
label define coropregio 18 "Kop van Noord-Holland" 19 "Alkmaar en omgeving" 20 "IJmond" 21
"Agglomeratie Haarlem" 22 "Zaanstreek" 23 "Groot-Amsterdam" 24 "Gooi en Vechtstreek"
label values coropcode coropregio

**Transaction date and energy label registration date**
generate str datum_nvm_afmelding = obj_hid_datum_afmelding
generate str datum_nvm_aanmelding = obj_hid_datum_aanmelding
generate str datum_rvo_registratie = pand_registratiedatum

*Energy label registration date*
generate str4 rvo_registratie_jaar= substr(datum_rvo_registratie,1,4)
generate str2 rvo_registratie_maand= substr(datum_rvo_registratie,5,6)
generate str2 rvo_registratie_dag= substr(datum_rvo_registratie,7,8)
destring rvo_registratie_maand, replace
destring rvo_registratie_dag, replace
destring rvo_registratie_jaar, replace
generate date_rvo_registration = mdy(rvo_registratie_maand, rvo_registratie_dag,
rvo_registratie_jaar)
format date_rvo_registration %td

*Transaction listing date*
generate str4 nvm_aanmelding_jaar= substr(datum_nvm_aanmelding,1,4)
generate str2 nvm_aanmelding_maand= substr(datum_nvm_aanmelding,6,7)
generate str2 nvm_aanmelding_dag= substr(datum_nvm_aanmelding,9,10)
destring nvm_aanmelding_maand, replace
destring nvm_aanmelding_dag, replace
destring nvm_aanmelding_jaar, replace
generate date_nvm_listing = mdy(nvm_aanmelding_maand, nvm_aanmelding_dag,
nvm_aanmelding_jaar)
format date_nvm_listing %td

*Transaction delisting date*
recast str09 datum_nvm_afmelding, force
generate date_nvm_delisting = date(datum_nvm_afmelding, "DMY")
format date_nvm_delisting %td


*Energy label class*
encode pand_energieklasse, generate(energieklasse)
sort obj_hid_postcode obj_hid_huisnummer_int

**Only the most recent energy label is recorded. For dwellings that sold multiple times,
this means that the observations of the transactions that took place prior to the
registration of the energy label are dropped**
*Identify duplicate address transactions*
sort obj_hid_postcode obj_hid_huisnummer obj_hid_huisnummertoevoeging
quietly by obj_hid_postcode obj_hid_huisnummer obj_hid_huisnummertoevoeging: gen
meermaals_verkocht = cond(_N==1,0,_n)
* When meermaalsverkocht = 0 than the address is sold only once*
count if meermaals_verkocht > 0
```

```
**Transaction characteristics**
*Generate integer variable with relative difference of listing price and transaction
price*
destring obj_hid_procverschil, generate(transactieprijs_procverschil) dpcomma

*Transaction year*
destring jaar, replace

*Time on the market*
generate int days_listed = date_nvm_delisting -  date_nvm_listing

**Urbanity index**
destring sted, replace force
destring oad, replace force

label define stedelijkheidsklasse 1 "Very highly urban" 2 "Highly urban" 3 "Moderately
urban" 4 "Mildly urban" 5 "Not urban"
label values sted stedelijkheidsklasse

**Generate dummy variables**
*Balcony*
gen balkon=1
replace balkon=0 if obj_hid_nbalkon==0

*Roof terrace*
gen dakterras=1
replace dakterras=0 if obj_hid_ndakterras==0

*Garden*
gen tuin=0
replace tuin=1 if obj_hid_tuinlig>0

*Outdoor space*
gen buitenruimte=0
replace buitenruimte=1 if tuin==1
replace buitenruimte=1 if dakterras==1
replace buitenruimte=1 if balkon==1

*Green energy label classes*
generate int groenlabel=0
replace groenlabel=1 if energieklasse==1
replace groenlabel=1 if energieklasse==2
replace groenlabel=1 if energieklasse==3

*Parking*
gen parkeergelegenheid=0
replace parkeergelegenheid=1 if obj_hid_parkeer>0

**Distance to amenties**
*Hospital*
destring af_ziek_e, generate(ziekenhuis) ignore(`"x."') force dpcomma
*Restaurant*
destring af_restau, generate(restaurant) ignore(`"x."') force dpcomma
*Supermarkt*
destring af_superm, generate(supermarkt) ignore(`"x."') force dpcomma
*Elementary school*
destring af_ondbas, generate(school) ignore(`"x."') force dpcomma
*Public swimming pool*
destring af_zwemb, generate(zwembad) ignore(`"x."') force dpcomma
*Freeway entrance*
destring af_oprith, generate(oprit) ignore(`"x."') force dpcomma
*Train station*
destring af_treinst, generate(trein) ignore(`"x."') force dpcomma
*General practitioner*
destring af_huisa, generate(huisarts) ignore(`"x."') force dpcomma

**Genereate log-linear variables**
*Amenities*
format %3.0f ziekenhuis
format %3.0f restaurant
format %3.0f supermarkt
format %3.0f zwembad
format %3.0f oprit
format %3.0f trein
format %3.0f huisarts
```

```
gen lnziekenhuis = ln(ziekenhuis)
gen lnrestaurant = ln(restaurant)
gen lnsupermarkt = ln(supermarkt)
gen lnzwembad = ln(zwembad)
gen lnoprit = ln(oprit)
gen lntrein = ln(trein)
gen lnhuisarts = ln(huisarts)

*Transaction price*
gen lntransactieprijs = ln(obj_hid_transactieprijs)
bysort jaar: egen gemiddelde_transactieprijs = mean(obj_hid_transactieprijs)
*Price per square meter*
gen vierkantemeterprijs = (obj_hid_transactieprijs / obj_hid_m2)
*Address density*
gen lnoad = ln(oad)
*Floor space*
gen lnoppervlakte = ln(obj_hid_m2)
*Rooms*
gen lnkamers = ln(obj_hid_nkamers)


**Factoral variables**
*Residency*
encode obj_hid_woonplaats, generate(woonplaats)

*Construction period*
label define bouwperiode 1 "1500-1905" 2 "1906-1930" 3 "1931-1944" 4 "1945-1959" 5 "1960-
1970" 6 "1971-1980" 7 "1981-1990" 8 "1991-2000" 9 ">2001"
label values obj_hid_bwper bouwperiode

*Dwelling type*
gen typewoning=0
replace typewoning=1 if obj_hid_nvmcijfers==2 |obj_hid_nvmcijfers==3
replace typewoning=2 if obj_hid_nvmcijfers==4
replace typewoning=3 if obj_hid_nvmcijfers==5
replace typewoning=4 if obj_hid_nvmcijfers==6
replace typewoning=5 if obj_hid_soortapp==1
replace typewoning=6 if obj_hid_soortapp==2
replace typewoning=7 if obj_hid_soortapp==3
replace typewoning=8 if obj_hid_soortapp==4
replace typewoning=9 if obj_hid_soortapp==5
replace typewoning=5 if obj_hid_soortapp==7
label define dwellingtype 0 "Unknown" 1 "Terraced house" 2 "Corner house" 3 "Semi-detached
house" 4 "Detached house" 5 "Ground floor apartment" 6 "Regular apartment" 7 "Multi-floor
apartment" 8 "Portico hall flat" 9 "Gallery hall apartment", replace
label values typewoning dwellingtype
count if typewoning==0

*Dummy house or apartment*
gen typehouse=0
replace typehouse=1 if typewoning<5
gen typeapartment=0
replace typeapartment=1 if typewoning>4

*Generate transaction year catergorical variable*
gen jaarcat=0
replace jaarcat=1 if jaar==2008
replace jaarcat=2 if jaar==2009
replace jaarcat=3 if jaar==2010
replace jaarcat=4 if jaar==2011
replace jaarcat=5 if jaar==2012
replace jaarcat=6 if jaar==2013
replace jaarcat=7 if jaar==2014
replace jaarcat=8 if jaar==2015
replace jaarcat=9 if jaar==2016
replace jaarcat=10 if jaar==2017
replace jaarcat=11 if jaar==2018
label define jaarcat 1 "2008" 2 "2009" 3 "2010" 4 "2011" 5 "2012" 6 "2013" 7 "2014" 8
"2015" 9 "2016" 10 "2017" 11 "2018"
label values jaarcat jaarcat

*Dummy regulated*
gen labelplicht=0
replace labelplicht=1 if jaar>2014
```

```
*Log of days listed*
gen lndays_listed = ln(days_listed)

**Drop invalid or missing data**
*Missing*
drop if obj_hid_m2 == 0
drop if obj_hid_bwper == 0
*Invalid*
drop if pand_energieklasse == "A+++"
drop if pand_energieklasse == "A++"
drop if pand_energieklasse == "A+"
drop if typewoning==0
drop if sted ==0

*Outlier*
drop if obj_hid_transactieprijs >2500000
drop if obj_hid_m2 >750
drop if obj_hid_nkamers >10

*When the dwelling is sold multiple times and the most recent energy label was established
more than 3 months prior to the transaction than this transaction is dropped*
drop if meermaals_verkocht > 0 & (date_rvo_registration < date_nvm_listing-90)

*Label variables*
label variable lntransactie "Transaction price (log)"
label variable obj_hid_transactieprijs "Transaction price €"
label variable energieklasse "Energy label class"
label variable lnoppervlakte "Floor space (log)"
label variable obj_hid_m2 "Floor space square meters"
label variable obj_hid_bwper "Construction period"
label variable typewoning "Dwelling type"
label variable dakterras "Roof terrace"
label variable balkon "Balcony"
label variable tuin "Garden"
label variable buitenruimte "Outdoor space"
label variable parkeergelegenheid "Parking space"
label variable jaar "Transaction year"
label variable woonplaats "Residence"
label variable lnoad "Address density (log)"
label variable oad "Address density"
label variable sted "Urbanity index"
label variable coropcode "COROP-region"
label variable obj_hid_nkamers "Rooms"
label variable lnziekenhuis "Distance to hospital (log)"
label variable ziekenhuis "Distance to hospital"
label variable lnsupermarkt "Distance to large supermarket (log)"
label variable supermarkt "Distance to large supermarket"
label variable lnzwembad "Distance to public swimming pool (log)"
label variable zwembad "Distance to public swimming pool"
label variable lnoprit "Distance to freeway entrance (log)"
label variable oprit "Distance to freeway entrance"
label variable lntrein "Distance to train station (log)"
label variable trein "Distance to train station"
label variable lndays_listed "Days listed (log)"
label variable days_listed "Days listed"
label variable jaarcat "Year index"
label variable labelplicht "After 2015 Enrgy Label policy change"
label variable transactieprijs_procverschil "Relative discount transaction price"


**Descriptive statistics**

*Continous variables*
fsum lntransactieprijs lndays_listed dakterras tuin balkon parkeergelegenheid labelplicht
lnoppervlakte lnoad lnziekenhuis lnsupermarkt lnzwembad lnoprit lntrein, uselabel
fsum obj_hid_transactieprijs days_listed dakterras tuin balkon parkeergelegenheid
labelplicht obj_hid_m2 oad ziekenhuis supermarkt zwembad oprit trein, uselabel
format(%18.2f)

*Categorical variables*

tabulate jaar energieklasse
tabulate coropcode energieklasse
tabulate obj_hid_nkamers energieklasse
tabulate obj_hid_bwper energieklasse
tabulate typewoning energieklasse
```

```
tabulate sted energieklasse
tabulate dakterras energieklasse
tabulate tuin energieklasse
tabulate balkon energieklasse
tabulate parkeergelegenheid energieklasse
tabulate coropcode sted

*Mean transaction price total*
mean obj_hid_transactieprijs
*Mean transaction prices per year*
mean obj_hid_transactieprijs, over(jaar)
mean vierkantemeterprijs, over(jaar)
*Mean transaction prices per COROP-region*
mean obj_hid_transactieprijs, over(coropcode)
mean vierkantemeterprijs, over(coropcode)

**Descriptive statistics*
*tabulate obj_hid_bwper
*sum ibn.obj_hid_bwper ibn.typewoning ibn.coropcode
*tabulate coropcode
*tabulate typewoning
*tabulate typehouse
*tabulate typeapartment
*tabulate energieklasse
*sum ibn.energieklasse
*tabulate jaar
*sum lntransactieprijs
*sum lnoad lnsupermarkt lnrestaurant lnoprit lntrein

*tabstat ibn.energieklasse i.obj_hid_nkamers ib9.obj_hid_bwper i.typewoning ibn.jaar
ibn.sted, separator(0) allbaselevels

*Model assumptions*
*Correlation matrix*
correlate lntransactieprijs energieklasse lnoppervlakte obj_hid_nkamers obj_hid_bwper
typewoning dakterras tuin balkon parkeergelegenheid jaar woonplaats lnoad sted coropcode
lnziekenhuis lnsupermarkt lnzwembad lnoprit lntrein

**Distribution transformed variables**
*Transaction price*
hist obj_hid_transactieprijs, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/transactieprijs.png", replace
as(png)

*Log transaction price*
hist lntransactieprijs, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/lntransactieprijs.png", replace
as(png)

*Floor space*
hist obj_hid_woonopp, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/oppervlakte.png", replace
as(png)

*Log floor space*
hist lnoppervlakte, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/lnoppervlakte.png", replace
as(png)

*Address density*
hist oad, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/oad.png", replace as(png)

*Log address density*
hist lnoad, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/lnoad.png", replace as(png)

*Days listed*
hist days_listed, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/dayslisted.png", replace as(png)

*Log days listed*
hist lndays_listed, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/lndayslisted.png", replace
as(png)
```

```
save "/Users/Cassee/Documents/Master Thesis/STATA/CBS NVM RVO.dta", replace

**Regression models**
**Model 1**
*Hypothesis 1*
*Regression model*
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats ibn.sted, allbaselevels

outreg2 using "/Users/Cassee/Documents/Master Thesis/Excel/Regressions 1-4.xls", dec(3)
label(insert) replace excel ctitle(Model 1) sideway stats(coef se) e(r2_a) addtext(Address
density, No, Distance to amenities, No, Number of Rooms, Yes, Municipality, Yes, Year *
Energy Label interactions, No, COROP-Region * Energy Label interactions, No, Energy Label
policy enforcement interaction, No) drop(lntransactieprijs ibn.woonplaats)

**Model 2**
*Hypothesis 1*
*Regression model*
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats lnoad ibn.sted lnziekenhuis lnsupermarkt lnzwembad lnoprit lntrein,
allbaselevels

outreg2 using "/Users/Cassee/Documents/Master Thesis/Excel/Regressions 1-4.xls", dec(3)
label(insert) append ctitle(Model 2) sideway stats(coef se) excel e(r2_a) addtext(Address
density, Yes, Distance to amenities, Yes, Number of Rooms, Yes, Municipality, Yes, Year *
Energy Label interactions, No, COROP-Region * Energy Label interactions, No, Energy Label
policy enforcement interaction, No) drop(lntransactieprijs ibn.woonplaats)

**Model 3**
*Hypothesis 2 & 3*
*Regression model*
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats ibn.sted i.jaar#energieklasse ibn.coropcode#energieklasse, allbaselevels

outreg2 using "/Users/Cassee/Documents/Master Thesis/Excel/Regressions 1-4.xls", dec(3)
label(insert) append ctitle(Model 3) sideway stats(coef se) excel e(r2_a) addtext(Address
density, No, Distance to amenities, No, Number of Rooms, Yes, Municipality, Year * Energy
Label interactions, Yes, COROP-Region * Energy Label interactions, Yes, Energy Label
policy enforcement interaction, No) drop(lntransactieprijs ibn.woonplaats)


*Distribution residuals*
predict ndresiduals, residuals
histogram ndresiduals, frequency normal
graph export "/Users/Cassee/Documents/Master Thesis/STATA/Distribution Residuals.png",
replace as(png)


*Error term average of 0*
pnorm lntransactieprijs, grid
graph export "/Users/Cassee/Documents/Master Thesis/STATA/Error term average.png", replace
as(png)

*Homoscedasticity*
rvfplot, recast(scatter)
graph export "/Users/Cassee/Documents/Master Thesis/STATA/Homoscedasticity.png", replace
as(png)


*Model 4**
*Hypothesis 4*
*Regression model*
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats ibn.sted i.labelplicht#energieklasse, allbaselevels
outreg2 using "/Users/Cassee/Documents/Master Thesis/Excel/Regressions 1-4.xls", dec(3)
label(insert) append ctitle(Model 4) sideway stats(coef se) excel e(r2_a) addtext(Address
density, No, Distance to amenities, No, Number of Rooms, Yes, Municipality, Year * Energy
Label interactions, No, COROP-Region * Energy Label interactions, No, Energy Label policy
enforcement interaction, Yes) drop(lntransactieprijs ibn.woonplaats)

*Chow test*
```

```
*Subsamble is labelplicht==0 if <2015 or labelplicht==1 if 2015>
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats ibn.sted i.labelplicht#energieklasse if labelplicht==0, allbaselevels
reg lntransactieprijs ibn.energieklasse lnoppervlakte ibn.obj_hid_nkamers
ibn.obj_hid_bwper ibn.typewoning i.dakterras i.tuin i.balkon i.parkeergelegenheid i.jaar
ibn.woonplaats ibn.sted i.labelplicht#energieklasse if labelplicht==1, allbaselevels


*Chow regression results*
* Number of regressors + constant = k = 15
*Total*
*RSS 4630,02*
*<2015*
*RSS 1335,26*
*2015>*
*RSS 3139,34*

*RSS1 + RSS2
di 3139.24+1334.26
*=4473.5

*N1 + N2 - 2(k+1)
di (33565+62741-(2*15))
* 96276


*Chow statistic*
di ((4630.02-4473.5)/15)/((4473.5)/(33565+62741-(2*15)))
*222.93562

*Critical value with 90% confidence
di invFtail(15,96276,.1)
* 1.487214

*Critical value with 95% confidence
di invFtail(15,96276,.05)
*1.6664901

*Critical value with 99% confidence
di invFtail(15,96276,.01)
* 2.0387142
```

Appendix 7 Chow-test results

	Label enforcement
Residuals squared sum (number of observations) pooled model	4630,02 (96.306)
Residuals squared sum (number of observations) <2015	3139,24 (33.565)
Residuals squared sum (number of observations) 2015>	1334,26 (62.741)
Number of regressors + constant (k)	15
Degrees of freedom (<2015+2015>-(2*k)	96.276
Chow-test	
F-statistic(15,96276)	222.93***
Critical F-value (15,96276,99%)	2.038

*** p<0.01

REFERENCES:

Aalders, R., Lennartz, C. and Vrieselaar, N. (2018), House Price Growth Reaches Dizzying Heights, While Sales Start to Decline, RaboResearch - Economic Research, available at: https://economics.rabobank.com/publications/2018/may/house-price-growth-reaches-dizzying-heights-while-sales-start-to-decline-/ (accessed 12 July 2018).

Boelhouwer, P., Conijn, J. and de Vries, P. (1996), "DEVELOPMENT OF HOUSE PRICES IN THE NETHERLANDS", Netherlands Journal of Housing and the Built Environment, Vol. 11 No. 4, pp. 381–400.

Boelhouwer, P.J. (2000), "Development of house prices in the Netherlands: An international perspective", Journal of Housing and the Built Environment, Vol. 15 No. 1, pp. 11–28.

Bokeloh, P. (2018), Woningmarktmonitor, ABN AMRO, Economic Bureau.

Boyle, M.A. and Kiel, K.A. (2001), "A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities", Journal of Real Estate Literature, Vol. 9 No. 2, pp. 117–144.

Braanker, W. (2015), Unknown Makes Unloved, Vastgoed Actueel, available at: https://www.vastgoedactueel.nl/sites/default/files/Consumentenonderzoek%20Energielabel%20.pdf (accessed 27 February 2018).

Brooks, C. and Tsolacos, S. (2010), "Real Estate Modelling and Forecasting by Chris Brooks", Cambridge Core, April, available at:https://doi.org/10.1017/CBO9780511814235.

Brounen, D. (2018), "Unfavorable energy label causes little pain in large cities", Tilburg University, 5 July, available at: https://www.tilburguniversity.edu%2Fnl%2Factueel%2Fnieuws%2Fpersbericht-effect-rode-energielabels%2F (accessed 12 July 2018).

Brounen, D. and Kok, N. (2011), "On the economics of energy labels in the housing market", Journal of Environmental Economics and Management, Vol. 62 No. 2, pp. 166–179.

Brueckner, J.K., Thisse, J.-F. and Zenou, Y. (1999), "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory", European Economic Review, Vol. 43 No. 1, pp. 91–107.

Case, K.E. and Shiller, R.J. (1988), The Behavior of Home Buyers in Boom and Post-Boom Markets, Working Paper No. 2748, National Bureau of Economic Research, available at:https://doi.org/10.3386/w2748.

Chow, G.C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions on JSTOR", Econometrica, Vol. 28 No. 3, pp. 591–605.

Daams, M.N., Sijtsma, F.J. and Vlist, A.J. van der. (2016), "The Effect of Natural Space on Nearby Property Prices: Accounting for Perceived Attractiveness", Land Economics, Vol. 92 No. 3, pp. 389–410.

van Dam, F. and Visser, P. (2006), Prijs van de plek - Woonomgeving en woningprijs, PBL Netherlands Environmental Assessment Agency, Den Haag, p. 81.

De Ayala, A., Galarraga, I. and Spadaro, J.V. (2016), “The price of energy efficiency in the Spanish housing market”, Energy Policy, Vol. 94, pp. 16–24.

DiPasquale, D. and Wheaton, W.C. (1994), “Housing Market Dynamics and the Future of Housing Prices”, Journal of Urban Economics, Vol. 35 No. 1, pp. 1–27.

Dutch Senate. (2008), “Europese Berichtgeving Eerste Kamer - E080134 - Voorstel voor een herschikking van richtlijn 2002/91/EG met betrekking tot de energieprestatie van gebouwen”, 13 November, available at: https://www.eerstekamer.nl/eu/edossier/e080134_voorstel_voor_een (accessed 3 July 2018).

Eskinasi, M. (2011), Houdini: Een systeemdynamische modellering van regionale woningmarkten, PBL Netherlands Environmental Assessment Agency, Den Haag.

Evans, A.W. (2004), Economics, Real Estate and the Supply of Land, John Wiley & Sons.

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, Vol. 48, pp. 145–156.

Galati, G., Teppa, F. and Alessie, R. (2011), “Macro and micro drivers of house price dynamics: An application to Dutch data”, DNB Dutch National Bank, p. 38.

Government of The Netherlands. (2017a), “Coalition agreement 2017: ‘Confidence in the future’ - Publicatie - Rijksoverheid.nl”, publicatie, , 10 October, available at:https://www.rijksoverheid.nl/documenten/publicaties/2017/10/10/regeerakkoord-2017-vertrouwen-in-de-toekomst.

Government of The Netherlands. (2017b), “Report ˝Study of effect energy labels for dwellings˝ - Rapport - Rijksoverheid.nl”, rapport, , 10 February, available at:https://www.rijksoverheid.nl/documenten/rapporten/2017/10/02/rapport-%CB%9Donderzoek-effect-energielabel-voor-woningen%CB%9D.

Government of The Netherlands. (2018), “Kabinet: einde aan gaswinning in Groningen - Nieuwsbericht - Rijksoverheid.nl”, nieuwsbericht, , 29 March, available at: https://www.rijksoverheid.nl/actueel/nieuws/2018/03/29/kabinet-einde-aan-gaswinning-in-groningen (accessed 31 July 2018).

Harding, J.P., Rosenthal, S.S. and Sirmans, C.F. (2007), “Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model”, Journal of Urban Economics, Vol. 61 No. 2, pp. 193–217.

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

Kok, N. and Jennen, M. (2012), “The impact of energy labels and accessibility on office rents”, Energy Policy, Vol. 46, pp. 489–497.

Koster, H.R.A. and Rouwendal, J. (2017), “Historic Amenities and Housing Externalities: Evidence from the Netherlands”, The Economic Journal, Vol. 127 No. 605, pp. F396–F420.

Lazrak, F., Nijkamp, P., Rietveld, P. and Rouwendal, J. (2014), “The market value of cultural heritage in urban areas: an application of spatial hedonic pricing”, Journal of Geographical Systems, Vol. 16 No. 1, pp. 89–114.

Milieu Centraal. (2019), “Verbeterjehuis - Milieu Centraal”, 20 January, available at: https://www.verbeterjehuis.nl/ (accessed 20 January 2019).

Netherlands Enterprise Agency. (2014), The Energy-Index and the Energy Label, DGMR.

Netherlands Enterprise Agency. (2016), “Energielabel C kantoren | RVO.nl”, April, available at: https://www.rvo.nl/onderwerpen/duurzaam-ondernemen/gebouwen/wetten-en-regels-gebouwen/bestaande-bouw/energielabel-c-kantoren (accessed 9 August 2018).

NOS.nl. (2018), “Wiebes in brief: bedrijven moeten versneld stoppen met Groningengas”, 22 January, available at: https://nos.nl/artikel/2213299-wiebes-in-brief-bedrijven-moeten-versneld-stoppen-met-groningengas.html (accessed 31 July 2018).

Renes, G., Thissen, M. and Segeren, A. (2006), Betaalbaarheid van koopwoningen en het ruimtelijk beleid, PBL Netherlands Environmental Assessment Agency, Den Haag, p. 46.

Rooijers, E. (2015), “‘We kunnen niet om de vastgoedwereld heen bij terugdringen CO2’”, Fd.Nl, 12 April, available at: https://fd.nl/ondernemen/1129909/we-kunnen-niet-om-de-vastgoedwereld-heen-bij-terugdringen-co2 (accessed 3 July 2018).

Rosen, S. (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, Journal of Political Economy, Vol. 82 No. 1, pp. 34–55.

Sirmans, G.S., Macpherson, D.A. and Zietz, E.N. (2005), “The Composition of Hedonic Pricing Models”, Journal of Real Estate Literature, Vol. 13 No. 1, pp. 3–43.

Statistics Netherlands. (2018a), “Landelijk dekkende indelingen”, Centraal Bureau voor de Statistiek, webpagina, , available at: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/gemeente/gemeenten-en-regionale-indelingen/landelijk-dekkende-indelingen (accessed 27 November 2018).

Statistics Netherlands. (2018b), “Regionale verschillen huizenprijzen”, webpagina, , available at: http://www.cbs.nl/nl-nl/nieuws/2018/06/regionale-verschillen-huizenprijzen (accessed 16 July 2018).

Statistics Netherlands and Land Registry Office. (2018), “House prices at all-time high in May”, webpagina, , 21 June, available at: http://www.cbs.nl/en-gb/news/2018/25/house-prices-at-all-time-high-in-may (accessed 3 July 2018).

Van Duijn, M., Rouwendal, J. and Boersema, R. (2016), “Redevelopment of industrial heritage: Insights into external effects on house prices”, Regional Science and Urban Economics, Vol. 57, pp. 91–107.

Wong, G. k. m. (2002), “A Conceptual Model of the Household’s Housing Decision–Making Process: The Economic Perspective”, Review of Urban & Regional Development Studies, Vol. 14 No. 3, pp. 217–234.

ons