ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Financial Economics

Title: The impact of energy efficiency on the market value in the residential real estate market

Name student: Stijn van der Ven

Student number: 384899

Supervisor: drs. Jeroen van Haaren

Second assessor: Prof. dr. F.G. (Frank) van Oort

Date final version: 12-06-2019

Abstract

This paper adds to the debate on the carbon footprint of the built environment by assessing the market capitalization of energy efficiency in the Dutch residential real estate market. This is a crucial component for owners that want to invest in more sustainable homes, as this gives them insight into the question whether this impacts its value. The study assesses the capitalization with different methods, among them the hedonic pricing model estimated by Ordinary Least Squares (OLS) method, the Repeat Sales model and the Instrumental Variable (IV) approach. The analysis indicates a potential for bias in these types of models. As the most advanced and elaborated IV method yields different coefficients, suggesting that there is a downwards endogeneity bias in the OLS. The IV approach finds per 10% increase in energy efficiency a price premium of approximately 0.7% on average. This is relatively low compared to the earlier findings of the literature, but an explanation for this could be the large explanatory value (R^2 = 0.90) in this research and the more recent dataset (2009-2018). Moreover, this dataset contains data from after the mandatory inclusion of the Energy Performance Certificates (EPCs) in 2015. In prior research that only uses data before 2015 this may have caused a sample selection bias. At last, this research makes the results more tangible in four energy label transformation business cases. The findings of this research are relevant to public organizations and house owners as they present investment inefficiencies and therefore potential investment opportunities.

Keywords: Green, Residential housing market, Energy efficiency, Price premium

JEL codes: R30, R31, Q40, Q41

Acknowledgements

First, I would like to thank Jeroen van Haaren for supervising me throughout my thesis. His vast knowledge about the housing market and his personal guidance helped me in the writing process. Further, I would like to express my gratitude to the real estate department of the company Rebel, and especially Koen Mulder, as I wrote my thesis during an internship there. Also, I would like to thank Nederlandse Vereniging van Makelaars (NVM) and SpringCo for providing me with relevant data for my research. At last, I am also very grateful to my family and girlfriend for their support, and in particular Merel van der Ven for her critical remarks on my work.

Table of Contents

1. Introduction

At the Paris climate conference (COP21) in December 2015, the first-ever universal legally binding global climate deal was established (European Commission, 2015). The deal illustrates the broad level of agreement amongst the scientific and political community on three aspects. First, to acknowledge that climate change is real. Second, that action needs to be taken. Last, that we need to prevent the global temperatures from rising more than two degrees above pre-industrial levels (IPCC, 2007). Furthermore, the European Union (EU) has committed itself in the Europe 2020 growth strategy to reduce the greenhouse gas emissions to 20% below 1990 levels, as is part of the Climate and Energy package (European Commission, 2016). More in-depth research into the carbon emission by Wilson et al. (2013) and Hertwich and Peters (2009) shows that household consumption comprises about 72% of the total carbon emission. Moreover, Ivanova et al. (2017) explores the household carbon footprint in Europe and finds that 22% is associated to housing of which 28% is related to direct shelter (e.g. due to combustion of fuel for heating at home). Conclusively, these findings indicate that direct shelter comprises of about 4.44% of the total carbon emission.

In recent years, a general consensus has taken shape that we need to reduce the negative impact our built environment has on our planet (European Commission, 2015). Consequently, there is an increasing interest in sustainability within the field of real estate finance. A basic premise in finance is that if you want to attract investors (e.g. for an investment in energy efficiency) there should be a financial justification and viability concerning the risk and/or uncertainty of the investment. Additionally, a traditional theory in the field of economics is the neoclassical theory of capital. According to this theory investors (private and corporate) aim to maximize utility. These two fundamental theories exemplify the following: if an investment aims to attract investors it should asses the corresponding risk and uncertainty and present opportunities to maximize utility. For these reasons, valuators have an important advisory role towards investors in the real estate market. Their financial analysis has the ultimate legal responsibility to provide an assessment about the market value of a property based on market evidence, their own knowledge and market trends. The valuation of the real estate object is in general a leading indicator for the investors. This demonstrates the economic importance of including all relevant variables, e.g., energy efficiency, in the valuation process. If not, there is no economic incentive for investments in these areas.

In the field of real estate there are two pronounced types: commercial real estate and residential real estate. The Oxford Dictionary accurately points out the difference between the two. Residential is defined as 'designed for people to live in' and commercial is defined as 'making or intended to make a profit' (Oxford Dictionary, 2019). In recent literature commercial real estate is covered extensively (see among others Krause and Bitter, 2012; Fuerst and McAllister, 2011; Wiley, Benefield and Johnson, 2010). In this field large economic incentives exist to make profits. The commercial real estate field is arguably a more efficient and rational market than the residential real estate market (Fuerst and McAllister, 2011). For this reason many large investors, e.g., pension funds and institutional investors are active in commercial real estate markets. On the other hand, residential real estate is relevant for every person that needs shelter and its idiosyncrasy creates challenges of its own in fostering sustainability. Moreover, three types of actors exist in the residential real estate market, namely: owner-occupied, private rented, and social housing (Wilkinson and Sayce, 2019). Each group with their own ability and challenges with respect to more sustainable practices. This research zooms into the residential real estate market and focuses on the owner-occupied houses. This group indicates the intrinsic need of the owner to invest and optimize their property, but also represents the investors with the most heterogeneous characteristics and preferences.

Horne and Hayles (2008) show that a reduction of the carbon footprint of residential homes is crucial in the pursuit of sustainability. Energy efficiency is an omnipresent theme in sustainability, but also functions as an important indicator of the degree of sustainable residential investments (Krause and Bitter, 2012; Lorenz, Trück and Lützkendorf 2006, 2007; Lorenz and Lützkendorf 2008; Aroul and Rodriguez, 2017; Warren-Myers, 2012). Hence, this study zooms into the role of energy efficiency in residential real estate value. Consequently, this paper explores the financial impact of energy efficiency - the so-called pricing premium - on the market value of residential houses.

Research question: Does energy efficiency of residential real estate have an impact on the market value of houses in the Dutch housing market?

This research continues with an exploration of the existing literature concerning both sustainability and energy efficiency its relation to the market value of residential real estate. Prior research into sustainability and real estate value is important to explore. This gives insights into the methodology and potentially relevant economic theories, but also legitimizes the decision to research energy efficiency. In continuation, the literature review focuses on the relationship between energy efficiency and residential real estate value. The theoretical findings are used as a foundation to develop hypotheses and these are visualized in a conceptual framework. Subsequently, a model using the hedonic pricing method estimated by ordinary least square (OLS) is presented. The model portrays the relationship between different energy efficiency estimators, like the energy performance certificates[1] (EPCs) and the energy performance index[2] (EPI), to the transaction price of a house. In addition, an instrumental variable (IV) approach is used to control for potential endogeneity. Furthermore, a repeat sales model is used as an alternative way to analysis the effect of increased energy efficiency measures over time on the transaction price. Lastly, robustness checks to the findings are performed.

The results of this study indicate that there is a price premium for energy efficiency within residential real estate. The different methods and energy efficiency measures show similar results, an additional 10% decrease in the EPI score leads to a price premium ranging from 0.46 to 0.7%. The EPCs are found to have a price premium between positive 3.19% (label A) and negative 2.02% effect (label G). Additionally, consumers would be willing to pay an additional 2.46% for a green labels (label A, B, and C), however this also works the other way around. A negative price premium of 1.76% exists for bad labels (label E, F, and G). All in all, the results are in line with the expectations, but seem to be lower than described in earlier literature. Nonetheless, potential explanations for this are the large explanatory value of this research (R^2 around 0.9) and the mandatory inclusion of the EPCs in 2015 (and therefore a potential sample selection bias in prior research that only uses data from before 2015).

The remainder of the paper is organized as follows: Chapter 2 consists of a review of existing literature, in particular with respect to findings and methodology. In chapter 3 the

---

[1] The EPCs give a label score to buildings from A (most efficient) to G (least efficient). In the Netherlands the EPCs are also referred to as energy labels. Moreover the EPCs in section 2.1.2.

[2] The EPI score is a continuous number that scales negatively, meaning that a lower EPI scores equals to more energy efficient. The EPI is a more precise measure than the EPCs. Moreover in section 2.1.2.

conceptual framework is presented, with the corresponding conceptual model, hypotheses, and the research question. In chapter 4 the chosen methodology and the model are introduced. Chapter 5 discusses the dataset and the corresponding matching process. Chapter 6 includes the results and the relevant outputs but also the synthesis, in which the findings are compared to the conceptual framework. Lastly, in chapter 7 the conclusion, limitations, and the recommendations for further research are presented.

2. Literature review

During the first years of the twenty-first century, sustainability emerged to be a prominent topic of interest among urban scholars. Different movements (e.g. New Urbanism) now take into account the benefits of energy efficiency (Jepson and Edwards, 2010). Moreover, more compact and connected forms of development are universally accepted to be superior to the original forms of development (Krause and Bitter, 2012). However, it seems that these sustainable factors slowly percolate into the marketplace. In response, researchers in the real estate field attempt to demonstrate that green and energy efficient buildings can command additional pricing premiums. The earlier literature such as Lorenz, Trück and Lützkendorf (2006, 2007), Lorenz and Lützkendorf (2008), and Warren-Myers (2012) focuses on the role of valuating sustainability in residential real estate. The most recent residential real estate literature has developed a focus on energy efficiency instead of sustainability as a whole, as it is extremely difficult to quantify and precisely define sustainability in this privacy-sensitive field. Furthermore, the initiative of the European Commission to introduce mandatory EPCs has given a boost to research into the price premium of energy efficiency (Mudgal et al. 2013). Moreover, findings in international studies indicate that there is a positive price premium between 2.8% to 12.8% for better energy labels or more energy efficient houses. However, recent research in the Netherlands by Aydin, Brounen and Kok (2018) illustrates that there is no price premium for the actual energy label, but only a price premium for energy efficiency itself. Conclusively, more research is needed with different methodologies to explore the price effect of energy efficiency on residential real estate value. The literature review begins with a concise exploration of the earlier research into the price premium of sustainability in the residential market, in order to present relevant takeaways and directions for this research. In continuation, the remainder of the literature review focuses on studies that explore the relationship between energy efficiency and residential real estate value and their corresponding methodologies.

2.1 Learning from research on sustainability

One of the first to empirically investigate the relationship between sustainability and the market value of residential real estate were Lorenz, Trück and Lützkendorf (2006, 2007) and Lorenz and Lützkendorf (2008). These researchers performed multiple studies on the valuation process in the German residential market. They covered various relevant themes in regard to sustainability: the impact of risk and uncertainty factors in property valuations, the relationship between the sustainability of construction and market value, and the role of sustainability in property valuation. The findings of Lorenz, Trück and Lützkendorf (2006) are consistent with the theoretical suggestions of Krause and Bitter (2012). If the perceived riskiness or uncertainty of real estate decreases (e.g., a dwelling becomes more sustainable) this should be taken into account in a risk premium and ultimately in the real estate value. Moreover, sustainability would be an appropriate instrument for risk and uncertainty reduction. The second research they conducted is one of the first to propose a hedonic pricing model as a way to incorporate different sustainability factors (among them the energy certificate and energy demand) in the valuation process of real estate. The last study of Lorenz and Lützkendorf (2008) identifies the financial benefits and risk reduction potential of sustainable design. The researchers explicitly stress the importance and necessity to integrate sustainability designs into the valuation process. The study concludes with key considerations for further empirical studies in the residential real estate field. Namely, the results are heavily reliant on the quality of the transaction data and better clarification of sustainability levels in the dataset would improve the results. Additional suggestions for research into residential real estate come from the study of Aroul and Rodriguez (2017). The researchers study the pricing premiums of green amenities in Frisco (Texas, United States of America). They find that the price premiums associated to green amenities are not stationary, the premium actually increased over their sample (2002-2009). The main drivers of the pricing premium would be energy efficiency and increased awareness of economic and non-financial benefits of energy efficiency. Furthermore, their research concludes that you cannot generalize findings between different international residential real estate markets as the attitude towards green amenities varies per country. For this reason, it is valuable to study domestic markets thoroughly in order to enhance the understanding of these particular markets.

A notable paper about the role of sustainability in real estate valuation is the meta-analysis of the literature by Warren-Myers (2012). She focuses on sustainability from a valuation perspective in both (commercial and residential) real estate markets. Her research indicates that the lack of historical evidence, data and information about sustainability leaves the relationship between sustainability and market value ambiguous. The evolving body of research focuses on normative effects of sustainability on market value. Exemplary of this phenomenon are the findings of Krause and Bitter (2012). These prove how sustainability 'should effect' market value rather than how it actually does (Kats, 2003). She concludes that there is a need for more quantitative studies of unbiased, evidence-based research in individual and broader markets to explore the implications of sustainability and valuation of real estate (Warren-Myers, 2012).

2.1.1 From sustainability awareness to pricing energy efficiency in the Netherlands

Conclusively, this paper takes into account the findings from the literature about sustainability and real estate value. The quantification process of sustainability is a complex process (Lorenz, Trück and Lützkendorf, 2007; Lorenz and Lützkendorf, 2008; Warren-Myers, 2012). For this reason (following the advice of Lorenz and Lützkendorf, 2008) the focus of this research lays on one aspect of sustainability that has been an omnipresent theme in the sustainable real estate literature, namely energy efficiency. Furthermore, this study focuses on high-quality transaction data instead of administrative taxation reports, as this is subjective to the appraiser and not necessarily reflects the supply and demand of the market. Additionally, as exemplified by Aroul and Rodriguez (2017), you cannot make cross-country generalization about the pricing premium of sustainability in residential real estate. Therefore this research focuses more thoroughly on an individual market: the Dutch housing market. The Dutch housing market is applicable for the following reasons: there is considerable amount of publicly available data (or by request) in the Netherlands, together with an existing body of research into the price premium of energy efficiency in the residential real estate market (Brounen and Kok, 2011; Aydin, Brounen and Kok, 2018; Havlınová et al., 2018).

2.1.2 The different facets of energy efficiency

In this research energy efficiency is measured by the use of EPCs and the EPI score. The EPCs are an international initiative in order to generate awareness for energy efficiency in the residential real estate market as an attempt to overcome information asymmetry. In essence, the EPCs give a label score (based on 10 variables) ranging from A (being the best) to G (being the worst) and this energy label (as how it is referred to in the Netherlands) should give an intuition of how energy efficient a house is.[3] In this manner, energy efficiency becomes more tangible and transparent. The EPCs were first introduced in the Netherlands in 2008, nevertheless the EPCs are only mandatory since 2015. An alternative measure of energy efficiency is the EPI score. This is a score given by a professional who calculates an energy indexation based on 150 different variables, for this reason the EPI score is more precise and accurate than the EPC rating. Moreover, the EPI score is presented as a continuous number that scales negatively. Therefore, when the number gets smaller the dwelling is more energy efficient. (Rijksdienst voor Ondernemend Nederland, 2019, Energie-Index). Additionally, if a house has a known EPI score the RVO will automatically estimate a corresponding label based on the EPI score.[4] However, this only works one way, if a dwelling has a label this dwelling will not get an EPI score. Conclusively, both energy efficiency indicators are researched individually as they both yield interesting research possibilities. Namely, a relatively small number of houses with a continuous number as estimate and a larger sample with the categorized labels.

In the field of finance, a considerable amount of studies have been conducted into investment inefficiencies, these are forces that may cause consumers to not maximize their profit. According to Allcott and Greenstone (2012) an illustrative investment inefficiency is the Energy Efficiency Gap. Essentially, this gap illustrates the difference between the cost-minimizing level of energy efficiency and the actual level of energy efficiency. This financial concept is also applicable to the residential real estate market. The energy efficiency gap theory states that the present value of the energy cost savings outweigh the needed energy efficiency investments, as illustrated in equation [2.1]. Following this argument consumers would underinvest and not

---

[3] See Appendix A, figure A1, A2 for an example of the composition of an energy label in the Netherlands (Institute for Market Transformation, 2019) and the variables that compose the energy score in the Netherlands.

[4] See Appendix A, figure A3 for a conversion Table of the EPI score to energy labels.

maximize their utility. The most probable reason for this is the lack of transparency and thus the information asymmetry in the real estate market (Aydin, Brounen and Kok, 2015, 2018).

[2.1] $$PV\ (energy\ cost\ savings) > PV\ (energy\ efficiency\ investments)$$

2.2 Measuring the effect of energy efficiency on real estate value

As described in the literature, the results from other countries might not necessarily be proportionate to the Dutch housing market (Aroul and Rodriguez, 2017). However, it is important to address the international literature in order to give an indication of the methodology and the potential results. In the remainder of the literature review, the research in the Dutch housing market is discussed, the case-study of the research is formulated, and lastly, the potential methodologies are explored.

2.2.1 International research

Mudgal et al. (2013) performed a widespread analysis on the EPCs and their impact on transaction prices on behalf of the European Commission. The researchers used a hedonic pricing method to explore the effect of EPCs in different parts of Austria, Belgium, France and the United Kingdom (UK). Overall, the EPCs seem to have a positive effect between 2.8 to 8% per one-letter label improvement on transaction prices. The only exception in their research are the results from the UK, a negative 4% price premium per one-letter label improvement. Additionally, Mudgal et al. (2013) find that in Belgium the effect is smaller in cities than in non-city areas. Nevertheless, the findings in Austria contradict these results as the effect there is greater in cities than in non-city areas. For this reason, the effect of city areas is ambiguous.

In response to the findings of Mudgal et al. (2013) from the UK, Fuerst et al. (2015) performed more research into the price premium of EPCs in England with a larger dataset (over 330.000 observations). The researchers use a hedonic pricing method and a cross-sectional augmented repeat sales method. The findings show that high labels (A and B) have a 5% price premium and low labels (G) sold for nearly 7% less. These findings do contradict the findings of Mudgal et al. (2013) about the UK. Additionally, Fuerst et al. (2016) performed the same research in Wales. Wherein they find even larger positive effects of high labels, namely 12.8% (A and B), and low labels (F) a negative effect of 6.5%. The researchers did not find a significant effect of label G on the transaction price, which is not in line with their findings in England. Stanley, Lyons and Lyons (2016) did also build further upon the research of Mudgal et al. (2013). They found that a one-letter improvement within the EPC system equals to a transaction price increase of 1% in Ireland.

More research into the price premium of EPCs is done by Fregonara, Rolando and Semeraro (2017). They explore the price effect of EPCs in the Turin real estate market. Their research uses a hedonic pricing method and focuses on different aspects of the selling procedure (listing price, transaction price, time-on-market, and bargaining outcome). The researchers find that the EPCs do not impact the prices in the Turin real estate market. In Sweden, Cerin, Hassel and Semenova (2014) explored the effect of EPCs on transaction prices. The researchers use a hedonic pricing method and perform univariate analyses. Their findings indicate that 1% increase in energy performance leads to a significant 0.06% increase of the transaction price.

The findings with respect to the relationship between EPCs and transaction prices (Mudgal et al., 2013; Fuerst et al. 2015, 2016; Fregonara, Rolando and Semeraro 2017; Cerin, Hassel and Semenova 2014; Stanley, Lyons and Lyons 2016) are in line with the remark of Aroul and Rodriguez (2017). The energy efficiency price premium is not generalizable and is dependent on the geographical location. However, to critically reflect upon the EPCs, they are designed based on benchmarks, for this reason scores are not continuous but portrayed in categories (A, B, C etc.). This classification is broad, and therefore not the most accurate estimation of energy efficiency.

The main approach to research the price premium of energy efficiency is the quantification of the effect of the EPCs on the transaction price. However, Ayala, Galarraga and Spadaro (2016) and Pride, Little and Mueller-Stoffels (2018) used alternative approaches to estimate the effect of energy efficiency. Ayala, Galarraga and Spadaro (2016) examined the price of energy efficiency in the Spanish housing market. They applied a hedonic pricing model and used data of surveys in order to determine the energy efficiency of households. The researchers found that more energy efficient houses have a price premium between 5.4 and 9.8%. On a critical note the sample size is relatively small (1507 observations) and the variables are subjective to a questionnaire, in which people might give socially desirable answers. This could potentially explain the high price premium that is found by the researchers. Pride, Little and Mueller-Stoffels (2017) researched the effect of a residential energy efficiency subsidy program (Rebate) on transaction prices in Alaska (United States). The primary goal of the policy is to reduce household energy costs by investing more into residential energy efficiency improvements. The researchers used a hedonic pricing model, repeat sales model and a difference-in-difference approach. Their results indicate that people are willing to pay a price premium of 4.2% for the houses that joined the Rebate program, thus more energy-efficient invested houses. The study shows the difference between the treatment group (Rebate program) and the control group (non-Rebate program). For this reason, it does not illustrate an exact comparable relationship between energy performance and transaction prices. However, the findings indicate that incentivizing subsidies have a pronounced effect. A recent literature review of Wilkinson and Sayce (2019) portrays the existing body of literature on energy efficiency and residential value in the European Union. Their exploration of the literature distinguishes between articles that discuss different types of houses (owner-occupied, private rented, and social housing). They recommend more research into the owner-occupied houses and recognize the complexity of an appropriate business case for energy efficiency investments in this field. This can be seen as a support for the existence of the energy efficiency gap as introduced by Allcott and Greenstone (2012).

All in all, the international literature predominately indicates a positive price effect of higher EPCs on real estate value. The only exceptions are the negative results of Mudgal et al. (2013) in Oxford (England) and no results at all of Fregonara, Rolando and Semeraro (2017) in Turin (Italy). However both studies did have relatively small sample sizes (respectively 1344 and 879 observations) so this might have affected the results. An interesting fact is that most of the studies are conducted in an economic volatile period due to the credit crisis. Furthermore, the European debt crisis might have had an impact on the results. Especially, in Ireland[5] (Mudgal et al., 2013; Stanley, Lyons and Lyons, 2016) and Spain[6] (Ayala, Galarraga and Spadaro, 2016) as emergency measures have been taken during the observational period. All in all, the studies with larger samples did find positive price premiums for higher EPCs or better energy efficiency in general. In conclusion, based on the international literature a positive price premium between 2.8 and 12.8% is expected.

---

[5] The findings in Ireland of Mudgal et al (2013) are relatively low, 2.8%, while their findings from other countries were all above 4%.

[6] Ayala, Galarraga and Spadaro (2016) find a price premium of 5.4 to 9.8%, this is high in comparison to the other findings of the literature.

2.2.2 Research in the Dutch residential real estate market

Brounen and Kok (2011) are the first to explore the relationship between energy labels and residential real estate value in the Dutch housing market. The researchers use a hedonic pricing model and their dataset consists of early adaptors of the energy labels and the corresponding transaction prices. They find that homebuyers are willing to pay a price premium for houses that are labeled as greener. Although there are no immediate financial benefits from owning a greener house, this seems to be taken into account in the acquisition process. Nevertheless, their paper concludes that they were not able to distinguish between the effects of labeling and the economic benefits of energy savings per se.

Additional research is done by Havlınová et al. (2018). She has studied the effect of EPCs on the residential housing market. Her study focuses on the mandatory inclusion of EPCs by the government in 2015, making it compulsory to incorporate a label in the selling procedure. The results of her research indicate a price premium of 3.1% for green dwellings (label A, B, C). In addition, the pricing premium increases by 2.5% after the implementation of the legislation in 2015. Alternatively, Havlınová et al. (2018) uses a different energy efficiency estimator, and find that 1% decrease in the EPI score leads to an increase of 2.3% in the transaction price. However, the results in this research might be on the higher end as there are no controls for potential endogeneity. The only study that controls for endogenous variables in the valuation of energy efficiency is the study of Aydin, Brounen and Kok (2018). Their research focuses on information asymmetry and energy efficiency in the Dutch housing market. The researchers perform an instrumental variable approach and a repeat sales method. The instrumental variable instrument used in their research is the discontinuity in energy efficiency levels in 1973-1974 oil crisis. They find that if the energy requirements of a household are reduced by 10% the market price of the house increases by around 2.2%. Furthermore, the researchers find that energy labels do not lead to a significant change in the buyers' valuation. This conflicts with the findings of Brounen and Kok (2011) and Havlınová et al. (2018). Their recommendations indicate that more research needs to be done in order to understand the homeowners' investment decisions.

The studies from Netherlands are the most relevant for this research as they use similar datasets and the results can be compared without the complexity of cross-cultural differences. However, the observational period of two of the studies in the Netherlands are during the Dutch housing market crisis[7] (Brounen and Kok, 2011: 2008-2009 and Aydin, Brounen and Kok, 2018: 2008-2011). This potentially makes the results less representative for a more stable economy. All in all, the Dutch literature predominately indicates a positive price premium of the energy labels. However, the study with arguably the most advanced methodology does not find a price premium for energy labels. Nevertheless, the research indicates that there is a price premium for energy efficiency itself. For this reason, the expected effect of energy labels on the residential real estate value is ambiguous and the expected effect of a decrease in EPI score is positive.

2.2.3 Case study of the research

The transaction and housing characteristics data, that is available for research purposes, is that of a single Dutch province. For this reason, this study attempts to focus on the most representative Dutch province. The province South-Holland is chosen for the following reasons. First, the province has the most households (Centraal Bureau van Statistiek, Regionale kerncijfers Nederland, 2019). Second, the corresponding public information about the housing market is most complete for South-Holland (Rijksdienst voor Ondernemend Nederland, EP-online, 2019). Last, following the findings of Pride, Little and Mueller-Stoffels (2017) there is a great variety of subsidies concerning energy efficiency that might stimulate investment in energy efficiency in South-Holland (Rijksdienst voor Ondernemend Nederland, Energie Subsidie Wijzer, 2019).

2.2.4 Summary of literature

In conclusion, the international literature on the price premium of energy efficiency in residential real estate value has predominantly shown us that there is a positive effect of EPCs and energy efficiency on real estate value. Moreover, research in the Netherlands indicates a positive price premium for energy efficiency. Nevertheless, it is ambiguous about the price premium of energy labels. Furthermore, prior research has shown the statistical value of performing different methods in order to valuate energy efficiency in real estate. For this reason, this research uses three different methodologies from the energy efficiency valuation literature. Namely, the hedonic pricing model, the repeat sales model and lastly, and instrumental variable approach.

---

[7] Appendix A, figure A5 and A6 show the price index and the volume sold in the Dutch housing market over the period 2008-2013.

Conclusively, Table 2.1 summarizes the papers that research the pricing premium of the EPCs and energy efficiency from the literature review. The table portrays the following: observational period, sample size, results, and corresponding methodology in a clear overview. In addition, appendix B contains a thorough review of the methodologies as described in the real estate valuation literature, namely: hedonic pricing model, repeat sales model and the instrumental variable approach.

3. Conceptual framework

This section briefly reflects upon the literature and constructs corresponding hypotheses for the remainder of the research. These hypotheses serve as guidance throughout the research and the results ultimately attempt to formulate an answer to the hypotheses.

3.1 Hypothesis development and corresponding conceptual model

In continuation, the literature predominately seems to find a positive effect of high EPC ratings on real estate value (Mudgal et al. 2013, Fuerst et al. 2015, 2016; Cerin, Hassel and Semenova, 2014; Brounen and Kok, 2011; Stanley, Lyons and Lyons 2016). In addition, other energy efficiency estimators also appear to have a positive effect (Ayala, Galarraga and Spadaro, 2016; Pride, Little and Mueller-Stoffels, 2017). In contrast, Fregonara, Rolando and Semeraro (2017) and Aydin, Brounen and Kok (2018) did not find a price premium for EPCs. Notably, Aydin Brounen and Kok (2018) is the only research that tries to control for endogenous variables, and therefore arguably uses the most advanced methodology. Opposing all other findings, Mudgal et al. (2013) found a negative pricing premium for high EPC ratings in Oxford (England). All in all, the literature is primarily positive about the price premium of high EPC ratings[8], although there are some studies that indicate otherwise. For this reason this study explores:

Hypothesis 1 (H1): a high Energy Performance Certificate rating (label A, B, C) has a positive effect on residential real estate value.

---

[8] The high EPC ratings are also referred to as label A, B, and C, or known in this study as the green labels.

Characteristics of study				Findings of study	Methodolgy of study				
Article	Country	Period	Sample size	Effect energy efficiency on real estate value	Hedonic pricing model	Repeated sales	Instrumental variable	Difference in Difference	Univariate analysis
Aroul and Rodriguez (2017)	United States	2002 - 2009	25,272	Increasing interest in energy efficiency, ambiguous positive effect	x			x	
Ayala, Galarraga and Sparadro (2016)	Spain	2013	1,507	More energy efficient houses have a price premium of 5.4% to 9.8%	x				
Aydin, Brounen and Kok (2018)	The Netherlands	2008 - 2011	30,036	Find that EPCs do not have an effect. However, if energy requirements decrease by 10% the market price goes up by 2.2%	x	x	x		
Brounen and Kok (2011)	The Netherlands	2008 - 2009	194,379	Positive price premium, however varies with label category	x				
Cerin, Hassel, Semenova (2014)	Sweden	2009- 2010	67,559	1% increase in energy performance leads to 0.06% increase of transaction price	x				x
Fregonara, Rolando and Semeraro (2017)	Italy	2011 - 2014	879	No effect of the EPCs on the transaction price	x				
Fuerst et al. (2015)	England	1995 - 2012	330,095	Positive 5% price premium for high EPC score (A, B); negative 7% for low EPC score (G)	x	x			
Fuerst et al. (2016)	Wales	2003 - 2014	192,000	Positive 12.8% price premium for high EPC score (A, B); negative 6.5% for low EPC score (F)	x	x			
Mudgal et al (2013)	Austria	2012	3,000	One-letter improvement in EPC score estimated at 8%	x				
	Belgium	2012	26,000	One-letter improvement in EPC score estimated at 4.3%	x				
	France	2011 - 2012	3,400	One-letter improvement in EPC score estimated at 3.2% - 4.3%	x				
	Ireland	2008 -2012	20,000	One-letter improvement in EPC score estimated at 2.8%	x				
	United Kingdom	2012	1,344	One-letter improvement in EPC score estimated at negative 4%	x				
Pride, Little and Mueller-Stoffels (2017)	United States	2008 - 2015	18,724	4.2% price premium for energy efficient homes (Rebate program)	x	x		x	
Stanley, Lyons and Lyons (2016)	Ireland	2009 - 2014	2,792	One-letter improvement in EPC score estimated at positive 1%	x				

Table 2.1: shows the articles from the literature review. It contains the country, period and sample size, a short summary of the findings and the methodolgy that is used in these articles.

Building further upon H1, the existing literature indicates that there is not only a positive price premium for high EPC ratings. But also a negative price premium for low EPC ratings[9] (Mudgal et al. 2013; Fuerst et al. 2015, 2016). Therefore, this paper also researches the price premium of low EPC ratings.

Hypothesis 2 (H2): a low Energy Performance Certificate rating (label E, F, G) has a negative effect on residential real estate value.

The database of RVO contains the EPCs, but for some houses it also contains the EPI score.[10] The difference between the two measurements of energy efficiency is explained in section 2.1.2. In essence, the EPI score is more accurate than the EPC rating because it is a continuous energy efficiency estimator of a house.[11]

Hypothesis 3 (H3): a low Energy Index score has a positive effect on residential real estate value.

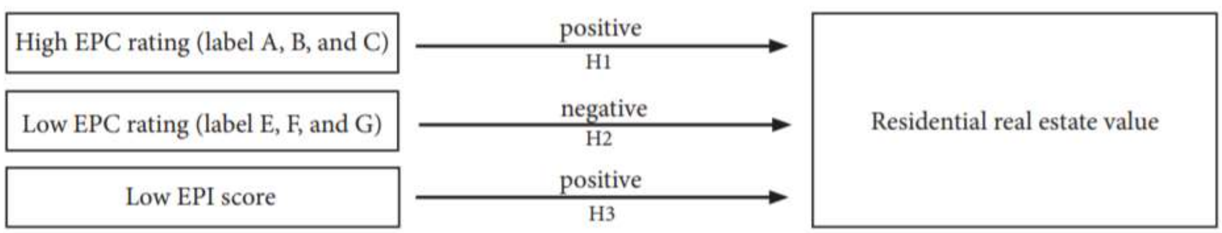

Figure 3.1: visualizes the hypotheses H1, H2, and H3 and the expected relationship between the independent variables and the dependent variable

The concept of the Energy Efficiency Gap as introduced by Allcott and Greenstone (2012) indicates an interesting potential investment inefficiency. Further research by Wilkinson and Sayce (2019) indicates the complexity of the corresponding business case. Hence, this research attempts

---

9 The low EPC ratings are seen as label E, F, and G, or known in this research as the bad labels

10 Appendix A, figure A4 shows the elaboration of the RVO about the Energy-Index (or known as EPI score). For more detailed information about the composition of the Energy Index: https://www.rvo.nl.

11 As the energy index (EPI) is a continuous number its effect applicable in two ways, positively and negatively. Therefore, H3 covers the price effect of both low and high energy index.

to study the existence of the potential energy efficiency gap and makes an effort to construct the corresponding business case.

Hypothesis 4 (H4): An energy efficiency gap exists in the residential real estate market

H4 = $PV\ (Energy\ cost\ savings) > PV\ (Energy\ efficiency\ investments)$

Conclusively, H1 and H2 focuses on the positive and negative price premiums of high and low EPC ratings on the residential real estate price. H3 uses the more precise Energy Index (EPI) as an estimate for energy efficiency. Lastly, H4 essentially tests if investors do not fully realize all benefits of potential energy efficiency investments.

4. Methodology

In this part the dataset and methodology of the research is described. The research method of the thesis is discussed and elaborated upon. First, the hedonic pricing model. Second, the repeat sales method. Third, the instrumental variable approach.

4.1 Hedonic pricing model

The literature seems to be unanimous about the approach to valuation in residential real estate.[12] For this reason, this research uses a hedonic pricing model based on the paper of Malpezzi (2003). The hedonic pricing literature indicates that variables have to be carefully selected.[13] In order to structure this process the eight categories as designed by Sirmans, Macpherson & Zietz (2005) will be used to classify the variables in this study.[14] Their meta-analysis composed a list of the twenty most frequently appearing variables in hedonic pricing models and their expected effect.[15] Due to data limitations this research does not have variables on two of the categories as identified by Sirmans, Macpherson & Zietz (2005). Therefore, the six categories relevant for this study are: structural characteristics(S), internal features (I), external features (Ef), environmental - neighborhood & location (El), environmental – natural (En), and lastly, marketing, occupancy

---

[12] Section 2, Table 2.1 portrays an extensive overview of the methodology of the earlier literature.
[13] Appendix B, section B1-B3 elaborates upon the commonly used methods in real estate valuation.
[14] Appendix C, figure C1 shows the eight categories as identified by Sirmans, Macpherson & Zietz (2005).
[15] Appendix C, figure C2 demonstrates the twenty most frequent variables in hedonic pricing models.

and selling factors (M). The transaction price (V) is the dependent variable. The categories represent variables of interest that are incorporated in the regression. The hedonic pricing model will therefore take an extended form of the original model of Malpezzi (2003).

[4.1] $V = f(S, I, Ef, El, En, M)$

Furthermore, the corresponding regression formula of the extended model is:

[4.2] $Log(Price_i) = \alpha_i + \beta S_i + \gamma I_i + \delta Ef_i + \epsilon El_i + \theta En_i + \vartheta M_i + \varepsilon_i$

The aim of the research is to observe the price premium of energy efficiency on the transaction price. In order to make the regression formula more comprehendible the categories of interest from regression [4.2] are joined together in the variable $Control_i$. The variable $Energy_i$ is added to the regression function, this variable represents the energy efficiency of the dwelling. The corresponding regression formula with inclusion of the energy efficiency variable and corresponding control variables is:

[4.3] $Log(Price_i) = \alpha + \beta Log(Energy_i) + \gamma Control_i + \varepsilon_i$

Equation [4.2] and [4.3] are estimated on the basis of an OLS, assuming that $Energy_i$ is independent of the error term ε_i. The data will be portrayed in a scatter plot in order to get a grasp of the data and the potential relations, and if needed winsorized. The regression is checked for multicollinearity and the error term ε_i examined for heteroscedasticity (Wooldridge, 2016). The dependent variable in [4.2] and [4.3] are logarithmic functions, for this reason the independent variables can be interpreted percentages. Furthermore, the OLS becomes less sensitive to extreme values (Wooldridge, 2016). On a critical note, the literature indicates that hedonic framework may contain irrelevant variables or omit relevant variables (Chau and Chin, 2003). The omitted variables may correlate to $Energy_i$ and therefore influence the coefficients and the outcome. The effect of the endogeneity on the β is most likely an underestimation of the true effect. On the grounds that, for example, a variable like aesthetics can influence the transaction price positively but make the house less energy efficient. Exemplary for this phenomena are dwellings with stained-glass windows or a classic facade. For this reason, this research aims to control for endogeneity with the use of an instrumental variable approach and therefore enhance the statistical credibility.

4.2 Repeat sales model

The repeat sales model is used in order to study houses that did become more energy efficient over time. However, the dataset does not contain this kind of information on the EPCs or EPI score. Nevertheless, in the EPC calculations four different types of isolations are mentioned as important characteristics for the determination of the label score.[16] For this reason, the repeat sales method focuses on the change of isolation over time and its corresponding price premium. The mathematical approach to this is based on Malpezzi (2003):

[4.4] $Log(V_x^i) - Log(V_y^i) = \beta_x T_x - \beta_y T_y$

Wherein: Price (V), House unit (i), Time dummy (T_i), sold in period x and y.

As expected, the amount of isolation methods is strongly correlated to the EPCs by 0.56 and correlated to the log (EPI) variable by 0.46.[17] For this reason, this can be used as an alternative method to approach the price premium of energy efficiency. In order to check whether the findings support the results from the OLS and instrumental variable approach. Nevertheless, the results from this method do not give an exact number about the price premium of energy efficiency, because it researches the effect of isolation.[18]

4.3 Instrumental variable approach

Aydin, Brounen and Kok (2018) and Havlınová et al. (2018) indicate that potential endogenous effects influence the relation between energy efficiency and transaction prices in the Dutch housing market. For this reason, Aydin, Brounen and Kok (2018) implement an instrumental variable approach in order to control for endogeneity and overcome this bias. Following this idea, this research explores three different potential instrumental variables. Namely,

---

[16] Appendix A, figure A2 shows four types of isolation in the calculation of the EPCs.

[17] Appendix D, figure D1 portrays the correlation results for isolation and inverse log (EPI). Figure D2 portrays the correlation results for isolation to the EPC.

[18] The method of researching the change in isolation is suboptimal compared to researching the change in EPI score or the EPCs over time. Only due to privacy regulations in the Netherland this data cannot be shared.

the oil crisis price shock in 1974 (also used by Aydin, Brounen and Kok, 2018), the estimated energy consumption of a dwelling and the estimated gas consumption by a dwelling per m^2. The instrumental variables are checked for the strength in the first stage (F-test > 10), the endogeneity (Durbin-Watson test) and the over identifying assumption (Sargan-Hansen test).

As described in the instrumental variable literature[19] the variables need to satisfy the following assumptions: first, variables oil price shock, estimated energy consumption and estimated gas consumption per m^2 (z) should not have a partial effect on transaction prices of houses (y). Second, oil price shock, estimated energy consumption and gas consumption per m^2 should not be correlated with the confounding variables. Last, oil price shock, estimated energy consumption and gas consumption per m^2 must be related to variable energy efficiency (x).

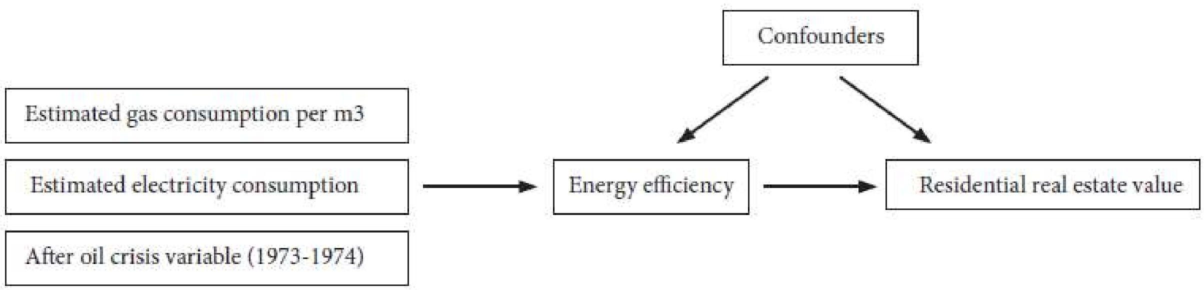

Figure 4.1: visualizes the relationship between the instrumental, dependent and independent variables.

In continuation, the instrumental variables should hence satisfy the following two statistical assumptions [6], [7] and an argumentation is needed why the instrumental variables only affects the transaction price through energy efficiency:

[6] $Cov(z, u) = 0;$ *[7]* $Cov(z, x) \neq 0$

In more detail, the rationale that the instrumental variables (log of electricity consumption and log of gas consumption) only affect the dependent variable (log of transaction price) through the independent variable (energy efficiency). After the acquisition of a house, the dweller needs to find an energy or gas distributor, therefore this is a follow-up procedure of the transaction and

[19] Appendix B, section B3, describes the general literature of the instrumental variable approach.

hence should not be included in the valuation process. One might argue that the energy or gas consumption could be an indicator of expected costs. Nevertheless, the former consumption is heavily reliant on the consumption behavior and household characteristics of the old dweller and the level of energy efficiency of the house.[20] Based on this argument one can say that in this model electricity and gas consumption only affects transaction prices through the level of energy efficiency of a house. Alternatively, energy prices are one of most important drivers for investments in energy efficiency. For this reason, an exogenous shock in the energy prices (like the Oil price shock in 1973-1974) would cause a discontinuity that can be exploited as an instrumental variable.[21] The oil prices increased by 252% in 1974 (BP, 2017), for this reason houses anticipate on this new price and therefore adapted themselves and became more energy efficient.[22] In order to make use of this exogenous shock a dummy variable is constructed that indicates if the dwelling is constructed after the oil price crisis.

A Two-Stage Least Squares (2SLS) regression model is used for the IV estimation. The corresponding first stage regression formulas are:

[4.5] $Log(Energy\ _i) = \alpha + \beta Log(Energyconsumption)_i + \gamma Control_i + \varepsilon_i$

[4.6] $Log(Energy\ _i) = \alpha + \beta Log(Gasconsumptionperm^2)_i + \gamma Control_i + \varepsilon_i$

[4.7] $Log(Energy\ _i) = \alpha + \beta Dummy(Constructedafter1974)_i + \gamma Control_i + \varepsilon_i$

The second-stage regression [4.8] will use the outcomes of variable energy efficiency as presented by regressions [4.5], [4.6], and [4.7].

[4.8] $Log(Price\ _i) = \alpha + \beta \widehat{Energy_i} + \gamma Control_i + \varepsilon_i$

By the use of the first-stage regression [4.5], [4.6] and [4.7] a coefficient β is estimated for $Energy_i$ that is exogenous and independent of the error term. For this reason, we can perform the second-stage regression [4.8] with the corresponding variable $\widehat{Energy_i}$. This alternative method

[20] There is no public data available concerning the attitude of the old dweller, for this reason this is not integrated in the model. This could be an interesting field of research in the future.
[21] Appendix D, figure D3 shows the oil prices for the time period 1900-2017.
[22] Appendix D, figure D4 shows the EPI performance score over the years 1960-2000

should increase the statistical credibility of the results and therefore give a better insight in the impact of energy efficiency on the transaction price.

In conclusion, this research employs three different methods to find the price premium of energy efficiency in the Dutch residential real estate market. First, the most commonly used method in the earlier literature, the hedonic pricing method. Second, the repeat sales model based on the change in amount of isolation manners. Last, the instrumental variable approach is explored for three distinct potential IVs: estimated gas consumption per m^2, estimated electricity consumption and the after oil crisis shock.

5. Data

This research makes use of data from three different dataset. The databases of Nederlandse Vereniging van Makelaars (NVM), Rijksdienst voor Ondernemend Nederland (RVO), and SpringCo are merged on the basis of a unique code per dwelling based on their postal code, house number, and suffix. This joined dataset contains house-specific data on transaction prices, corresponding important characteristics, the EPCs, EPI score, the estimated electricity, and gas consumption.

5.1 Data sources

5.1.1 Nederlandse Vereniging van Makelaars

The NVM is the largest organization of realtors in the Netherlands. NVM provides panel data on transaction prices of residential real estate for research purposes to students. The dataset contains all transactions from NVM registered realtors in the province South-Holland from 2008 to 2018. The data is house-specific and describes important characteristics of the houses and the transaction price.[23] All the observations are specified on postal code, house number and suffix.

---

[23] Appendix E portrays all the important variables categorized based on the themes of Sirmans, Macpherson & Zietz (2005)

5.1.2. Rijksdienst voor Ondernemend Nederland – Energy efficiency data

The RVO is a governmental organization that supports entrepreneurs in the Netherlands. The RVO has an open database with over three million houses and their corresponding EPC rating (energy label) and (if known) the EPI score. The dataset only contains the EPCs and EPI score of the year 2018. However, it contains the registration dates of the EPCs, so there is the possibility to filter non-active labels during the transaction. The observations are specified on postal code, house number and suffix.

5.1.3 SpringCo – estimated electricity and gas consumption data

The company SpringCo provided a combined dataset from information of CBS (Centraal Bureau voor Statistiek) and BAG (Basisregistratie Adressen en Gebouwen) data in which the company estimated the electricity and gas consumption per household. The dataset relates the aggregated electricity and gas consumption data from CBS to the household surface data from BAG. This gives an estimate of the electricity and gas consumption on household level. The dataset is from 2017 and is specified on postal code, house number, and suffix.

5.2 Data matching procedure

The three distinctive datasets are checked for potential duplicates and inconsistent variables, furthermore the dataset is filtered geographically based on the province South-Holland. In continuation, the data of the research is matched based on the postal code, house number and the abbreviations. A unique number code is established for the three components above. For this reason the letters of the postal code are quantified (A=01, B=02, etc.). The sequence of the numbers is based on: postal code + house number + abbreviation. The abbreviations formed a challenge as there are many different abbreviations (e.g. II, IIA, and II1) in the different datasets.[24] The research has chosen to only quantify the abbreviations that are letters of the alphabet. These abbreviations did have the largest frequency and consistency throughout the three datasets. The letter abbreviations are quantified the same way as the postal code letters. An illustration of the quantification procedure: Postal code 1010AB number 2 abbreviation C has a unique identification

[24] Exemplary, in the dataset of SpringCo there are 4531 unique abbreviations and in the dataset of RVO 3614.

number (ID) of: 1010-0102-0002-03. Moreover, the complex letter and non-letter abbreviations (example: B2 or bis) are all quantified with the number 50. For example: Postal code 1010 BC number 3 abbreviation B2 becomes: 1010-0203-0003-50. The IDs with a complex abbreviation are dropped at first because of the high likelihood of wrongly matched houses. However, a robustness check is performed that takes these houses into consideration to see the impact of the inclusion of these alternative abbreviations.

5.3 Data matching and description

The dataset of NVM is used as the foundation for the merging process with the other datasets. Table 5.1 portrays the results of the matching procedure. The house-specific information of the RVO and SpringCo datasets is matched with a success rate of about 60%.[25]

	Total Observations	Matched with NVM	Percentage of NVM
NVM	284,473	x	x
RVO	876,201	172,510	61%
SpringCo	1,042,410	170,615	60%
RVO or SpringCo	x	238,096	84%

Table 5.1: Matching procedure of the RVO, SpringCo and NVM databases.

The main interest of this research the price premium of the EPCs and the EPI score on the transaction price. For this reasons alternations are made to the data. The data from RVO is from 01/01/2019 and indicates the current EPCs and EPI score with the corresponding registration date of the data. For this reason, we compare the transaction date to this registration date and drop the observation if the label was not registered during the transaction.[26] Furthermore, all continuous variables are inspected for skewness and kurtosis and if necessary winsorization is used. Additionally, all duplicates and inconsistent variables are dropped. After these procedures the dataset ‘After filter’ is constructed, this dataset contains 84,738 observations and is used in

[25] Note: There is a total of 46.377 houses (about 16% of the NVM database) that are not matched to either the RVO or SpringCo datasets. The similar numbers and percentages (both around 170.000 houses and 60% matching score) in Table 5.1 might imply that there are structural shortcomings in the ID creation process and that in both datasets the same houses are not matched. Nevertheless, this is not the case as illustrated with the total match to either RVO or SpringCo of 84%.

[26] The first EPC in the RVO dataset was registered on the date 06-01-2009, this means that all data before this period in time is dropped in the filtered version.

continuation of the research. Table 5.2 shows the summary statistics of the most relevant characteristics in the total sample and the filtered dataset.

	Total sample (284,473)		After filter (84,738)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Transaction price (in EUR)	247,671	150,214	250,945	144,421
Living area (in m2)	111	42	107	41
Size (in m3)	341	153	341	146
Number of floors	2.1	0.9	2.1	0.9
Number of rooms	4.3	1.5	4.3	1.4
Maintenance inside	6.9	1.1	6.9	1.1
Maintenance outside	7	0.9	7	0.8
Number of isolation methods	1.9	1.8	2.1	1.8
Estimated electricity consumption	2,883	1,027	2,852	993
Estimated gas consumption	1,178	602	1,121	574
Energy Performance Index	197	60	195	59

Table 5.2: Summary statistics of the combined dataset with the most important variables over the total and filtered sample.

The matching results based on the EPCs, EPI score of RVO, and estimated gas and electricity consumption of SpringCo is described in Table 5.3. This table shows the composition of the data in detail; summed up it shows that 20,969 houses have both EPI score and EPCs, and that the information for all aspects is complete for 12,140 dwellings in the dataset.

	Observations	Matched	Matched (after filter)
NVM	284,473	x	x
NVM x RVO labels	x	172,510	84,738
NVM x RVO EPI score	x	34,372	20,969
NVM x Label x Gas x Electricity	x	103,072	50,789
NVM x Labels x EI score x Electricity x Gas	x	20,019	12,140

Table 5.3: The exact matches of the most important variables from the RVO and SpringCo datasets.

Moreover the EPCs, Table 5.4 shows the label distribution over the total and filtered sample. The RVO dataset contains A+ (5) and A++ (3) labels, the numbers are too small for a representative analysis. For this reason these labels are transformed to the A label.[27] In both samples label C and D are the largest groups covering around 43 to 44%, in the filtered dataset there are relatively more green labels and less bad labels than in the total sample.

	Total sample		After filter	
EPC label	%	N	%	N
Label A	12.45%	21,485	14.36%	12,170
Label B	12.67%	21,857	13.42%	11,370
Label C	25.49%	43,973	26.69%	22,615
Label D	18.27%	31,518	17.98%	15,234
Label E	14.22%	24,531	13.04%	11,046
Label F	9.97%	17,199	8.78%	7,440
Label G	6.93%	11,955	5.74%	4,865
Total labels		172,510		84,738

Table 5.4: The composition of the label distribution in the total and filtered dataset.

In conclusion, an overview of all relevant variables and their types with the corresponding categories for this research is constructed.[28] Additionally, different types of relevant bar charts, distributions per building year, and scatterplots are.[29]

6. Results and Discussion

This section elaborates upon the findings of this research and the corresponding discussion. The three different methodologies that are discussed are the hedonic pricing model, the repeat sales model, and lastly, the instrumental variable approach. All in all, the different methods indicate a positive price premium for energy efficiency measures. The findings are checked for robustness with a variety of different tests. At last business cases are developed in order to make the results more tangible. Moreover, the themes as identified in the paper of Sirmans, Macpherson & Zietz (2005) are used to structure the control variables in all the regressions. This categorization enhances the readability of the regressions, as there are a considerable amount of control variables.

---

[27] This is in line with the conversion Table as found in appendix A, figure A3.

[28] Appendix E, Table E1 shows all the relevant variables for the regressions categorized in a thematic framework.

[29] Appendix E, figure E2-E9, shows eight relevant figures about the dataset.

6.1 Hedonic pricing method

The hedonic pricing method estimated by OLS as portrayed in equation [4.3] is the first method applied to explore the price premium of energy efficiency in residential real estate. The results from the OLS regression can be found in Table 6.1. The dependent variable in all six regressions is the logarithmic function of the transaction price. The independent variable in column 1 and 2 is the inverse of the logarithmic function of the EPI score. The inverse of the variable is taken to make the interpretation of the regressions more comprehendible and better comparable to the other energy efficiency measures. The EPI scales negatively, thus traditionally a lower EPI score illustrates that a dwelling is more energy efficient. Hence, the inverse of the EPI score indicates that a larger EPI score equals to more energy efficient.

Column 1 of Table 6.1 shows that a decrease of 10% in EPI leads to a significant 2.94% price premium. However, the model in column 1 does not control for other variables yet. In column 2 the control variables are added to the model and the price premium decreases to about 0.46% per 10% decrease in EPI.[30] The addition of the control variables increase the R^2 from 0.03 to 0.91. The R^2 is high, but in line with the expectations, because the hedonic pricing model is a method used to explain the price composition of a house. Additionally, column 3 performs tests with other indicators of energy efficiency, namely the sole energy labels of the houses. The difference between the two energy efficiency measures lays in the fact that the energy labels are not continuous but categorized. The dataset of the EPCs contains about 64.000 additional observations with regards to the EPI dataset. The regression in column 3 is performed without any control variables and label D (generally seen as the neutral label) as reference point. All energy labels indicate a significant effect, however interestingly the ‘bad’ labels F and G have a positive price effect with regards to label D. This contradicts the expectations, nonetheless this most likely has to do with omitted variables (e.g. building year or type of the building) that are not taken into account in the model of column 3. In column 4 the additional control variables are added, the R^2 increases from 0.11 to about 0.91 and the corresponding price premiums decrease substantially as expected. The results are now in line with the theory, a better energy label has a greater price premium. Interestingly, the results indicate that the one-letter label jump from D to C has relatively

---

[30] Appendix F, Table F1 shows the complete regressions with the inclusion of all variables.

the largest effect (about 1.91%). A potential explanation is that the labels A, B, C are generally seen as green labels. Therefore the transaction price might respond to the transition from a neutral label to a green label. The individual labels have a price premium ranging from positive 3.19% (label A) to negative 2.02% (label G). The results indicate that the price premium consistently scales according to the energy labels, meaning that better energy labels have larger price premiums.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Inverse Log(Energy Performance)	0.294***	0.0460***				
	(0.0114)	(0.00620)				
Label A			0.476***	0.0319***		
			(0.00569)	(0.00332)		
Label B			0.309***	0.0285***		
			(0.00580)	(0.00253)		
Label C			0.121***	0.0191***		
			(0.00491)	(0.00179)		
Label E			-0.0277***	-0.00585***		
			(0.00584)	(0.00198)		
Label F			0.0585***	-0.0108***		
			(0.00661)	(0.00239)		
Label G			0.0829***	-0.0202***		
			(0.00768)	(0.00331)		
Green label					0.0246***	
					(0.00163)	
Bad label						-0.0176***
						(0.00160)
Control variables:						
Structural characteristics	No	Yes	No	Yes	Yes	Yes
Internal features	No	Yes	No	Yes	Yes	Yes
External features	No	Yes	No	Yes	Yes	Yes
Environmental - Natural	No	Yes	No	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	No	Yes	No	Yes	Yes	Yes
Marketing, occupancy and selling factors	No	Yes	No	Yes	Yes	Yes
Constant	13.61***	8.271***	12.15***	8.222***	8.188***	8.206***
	(0.0597)	(0.168)	(0.00379)	(0.0590)	(0.0587)	(0.0590)
Observations	21,273	21,049	85,475	84,554	84,554	84,554
R-squared	0.030	0.908	0.113	0.909	0.909	0.909

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

Table 6.1: Four different measures for energy efficiency are used as independent variables: EPI score, energy labels, green labels and bad labels (dependent variable log of transaction price).

Additionally, the last two measures of energy efficiency are two groups based on the EPCs: the high labels, also known as the green labels (A, B, and C) and the low labels, known as the bad labels (E, F, and G). The results in column 5 and 6 indicate that there is a price premium of 2.46% for a green label and a negative price premium of 1.76% for a bad label. The results are slightly different than the earlier results from column 3 and 4. This is the case as the reference group in column 5 and 6 is the opposing group of energy labels (so either green or bad label) and the neutral label D. Nonetheless, the results indicate a significant price premium (positive and negative) for having either a good or a bad label.

The results are in line with hypotheses 1, 2 and 3 as formulated in section 3.1. However, the residential real estate literature (especially the international literature) seems to present a larger price premium for energy efficiency. This effect is most evident in the price premium of the energy labels in residential real estate. Exemplary for this phenomenon is the research of Mudgal et al. (2013) on account of the European Union that explored the price premium in different countries. Their results indicate that a one-letter improvement of the energy label leads to a positive price premium between 2.8 and 8%. Contrastingly, this research finds that a jump from label D to A (three-letters improvement) equals to a 3.19% price premium.

Other research that has studied the price premium of green labels and bad labels is that of Fuerst et al. (2015, 2016). In Wales and England the researchers find a positive price premium of respectively, 5 and 12.8% for green labels. Additionally, they find a negative price premium of 7 and 6.5% for bad labels. Likewise, the research of Havlınová et al. (2018) finds a 3.1% price premium for green labels in the Netherlands. These results, and especially those from Fuerst et al. (2015, 2016), are larger than the results of this paper (a positive 2.46% effect for green labels and a negative 1.76% for bad labels). At last, Cerin, Hassel and Semenova (2014) found that an additional 1% increase in energy efficiency leads to a transaction price increase of about 0.06%. The pricing premium is again larger than the results from this study (1% increase in energy efficiency leads to 0.05% increase in transaction price), but more comparable than the other studies.

There are two potential explanations for the variation in the results of this research with regard to the earlier literature. First, this research has a relatively high R^2 as it includes a great variety of control variables. These control variables have drastically decreased the price premium as clearly portrayed in column 1 and 2, or column 3 and 4 from Table 6.1. Second, the EPCs have become mandatory in 2015. For this reason, there might be a sample selection bias in the datasets of the earlier studies. As only more energy efficient houses would have requested an energy label or EPI score before 2015. This makes the sample less representative for the complete housing market and therefore these studies might overestimate the price premium of the EPCs and the EPI.

Conclusively, the results from the hedonic pricing method estimated by OLS indicate that there is a positive price premium for being more energy efficient. An additional 10% decrease in EPI score has a price effect of about 0.46%. In order to make this number more tangible, a shift from of 0.75 EPI[31] (same as 38.46% of 195) equals to a label jump from label D to A, or a label decrease from label D to G. Accordingly, the decrease in EPI representing a label jump from label D to A has a positive price premium of about 1.77% (vice versa, a label decrease from label D to G has a negative price premium of 1.77%). In continuation, the second energy efficiency measure, namely the individual energy labels, show that label A has the largest positive price premium with 3.19%. In addition, the results also indicate a negative price premium for bad labels, label G has the largest negative price effect of 2.02%. The last measure of energy efficiency categorized the labels more general in either green labels or bad labels. The green labels are valuated with a positive price premium of 2.46% and bad labels with a negative price premium of 1.76%. This verifies the earlier findings from the individual labels, namely that the positive price premium for a green label seems to be larger than the negative price premium for a bad label. The three different estimators of energy efficiency indicate similar results. However the continuous (and therefore arguably more precise) EPI score does show a lower price premium. The individual energy labels and the clustered groups show relatively high price premiums for good labels and lower negative price premiums for bad labels. All in all, the three measures indicate that there is an actual price premium for energy efficiency in residential real estate. In conclusion, the findings in the international and the Dutch housing market are pointing in the same direction but seem rather high compared to the findings in this study.

[31] Appendix A, figure A3 shows the conversion Table for an EPI score to an energy label.

6.2 Repeat sales method

The second method that evaluates the price premium of energy efficiency in residential real estate is the repeat sales model as formulated in equation [4.4]. The variable of interest is the change in the amount of isolation methods of a dwelling. The variable is an important variable in the EPI score and EPC rating calculation. Hence, the correlation of 0.46 with inverse log (EPI) and the strong correlation of 0.56 to the EPCs is expected.[32] This is the most appropriate variable with data over time to research with an absence of EPI and EPC data. The dataset for this analysis contains 34,511 houses that were sold two times in the period 2008-2018. Table 6.2 portrays the most important statistics for the total sample with the fixed variables (think of size or number of floors), and the non-fixed variables that may change overtime. Additionally, the table shows the average change in transaction price, maintenance (inside and outside) and the number of isolation methods.

	Total sample (34,511)	
Variable	Mean	Std. Dev.
Living area in m2	100	37
Size in m3	311	136
Number of floors	1.9	0.9
Number of rooms	4.0	1.4

	First sale		Second sale		
	Mean	Std. Dev.	Mean	Std. Dev.	Δ
Transaction price in EUR	209,066	122,391	234,244	140,318	12.0%
Maintenance inside	6.8	1.3	7.2	0.9	5.4%
Maintenance outside	7.0	0.9	7.1	0.7	1.7%
Number of isolation methods	1.7	1.7	1.8	1.6	6.3%

Table 6.2: Summary statistics of the repeat sales sample.

The variable isolation is defined as the amount of isolation methods, ranging from 0 to 5. Exemplary, an increase in one point of isolation equals to one additional isolation method (e.g. insulated glazing or facade insulation). The variables maintenance inside and maintenance outside range from 1 (poor) to 9 (very good). Column 1 of Table 6.3 shows that there is a significant positive price premium of 12.4% for one additional isolation method. Column 2 performs the same

[32] Appendix D, figure D1 shows the correlation output between inverse log (EPI) and isolation. In addition, figure D2 shows the correlation output between the EPCs and isolation.

regression on houses that have an unchanged amount of isolation methods in the first and second sale. As expected this does not present any price premium. However, column 1 and 2 both only control for year fixed effects. For this reason, additional control variables are added, namely the change in the maintenance inside and the change in maintenance outside. Column 3 shows a significant positive price premium of 8.86% per extra isolation method. Moreover, column 4 shows again that an unchanged amount of isolation still does not have an effect.

VARIABLES	(1)	(2)	(3)	(4)
Isolation (positive change)	0.124***		0.0886***	
	(0.00389)		(0.00370)	
Isolation (no change)		-0.00153		-0.00586
		(0.0112)		(0.0108)
Control variables				
Year fixed effects	Yes	Yes	Yes	Yes
Δ Maintenance inside			0.104***	0.109***
			(0.00562)	(0.00565)
Δ Maintenance outside			0.0993***	0.115***
			(0.00973)	(0.00966)
Observations	34,511	34,511	34,511	34,511
R-squared	0.055	0.028	0.114	0.101

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

Table 6.3: The regression over the repeat sales model that focuses on the amount of isolation ways (dependent variable log over the price change between the first and second transaction)

The results from the repeat sales model indicate that additional ways of isolation have a positive price premium. However, the finding 8.86% is over the price change between the first transaction and second transaction. This is in line with the expectations, because most types of additional isolation are expensive investments to the house and thus should be incorporated in the transaction price. In addition, extra types of isolation present energy cost savings, for this reason isolation has a large effect on the EPI score and the EPCs.[33] The results from the repeat sales model indicate a positive price premium for additional ways of isolation. Therefore, this supports the

[33] See Appendix E, figure E8 and E9, this portrays the relationship between the amount of isolation ways and the EPI score and EPCs.

earlier findings of a price premium for energy efficiency. Nonetheless, exact statements about the price premium of energy efficiency cannot be made based on Table 6.3. However, the results from the repeat sales model should be seen as a directive, and thus an alternative method to confirm the existence of a price premium of energy efficiency.

6.3 Instrumental variable

The last approach to identify the price premium of energy efficiency on residential real estate is the instrumental variable approach. The research tested for three different instrumental variables. Namely, the logarithmic function of estimated gas consumption per m2, electricity consumption, and the discontinuity of the oil crisis in 1974 (as used by Aydin, Brounen and Kok, 2018). The estimated gas and electricity consumption are both continuous numbers and the after oil crisis is a dummy variable representing the years after the crisis (due to categorization of the NVM data the after oil crisis dummy entails houses from 1980 onwards). Table 6.4 shows the first and second stage regressions of the three potential instrumental variables. The main statistical conditions for an instrumental variable to be credible is to look at its significance (F-test > 10, as rule of thumb) and if the variable is exogenous (Durbin-Watson p value should not be significant).

Table 6.4 shows that all three variables have an F-stat of above 10, but only after oil crisis (with as reference group before the oil crisis years 1945-60 and after the oil crisis 1981-1990) is exogenous as can be seen from the Durbin-Watson p value.[34] This research has further explored all potential combinations between the three different instrumental variables. Nonetheless, the combinations were either endogenous or over identifying according to the Sargan-Hansen test. For this reason only the after oil crisis instrumental variable is discussed in the remainder of the paper.

The building year variable in the dataset is not discrete but categorized in time periods. Therefore, the after oil crisis variable represents houses constructed from 1980 onwards. Nevertheless, the building years play an important role in the regression and cannot be left out completely as control variables. Hence, two reference groups in the data should be chosen, one before and one after the oil crisis. As described above, the reference groups 1945-1960 and 1981-1990, the rougher economic times with the relatively cheap and lower quality houses.

[34] The variables log (gas consumption per m2) and log (electricity consumption) are also tested just for observations from the year 2017 (as this is their registration year). Log (electricity consumption) was endogenous but only had an F score of 9.95.

Alternatively, the reference groups 1930-1944 and after 2000, these houses are the better quality houses with a high demand. However, when the latter one is used as a reference group the Durbin-Watson p value falls to around zero. This means that the after oil crisis is not a good instrumental variable with these reference groups. This is most likely the case as there is no control variable for the newest and most energy efficient houses (after 2001). The houses from 1930-1944 and after 2000 are both popular in demand but in characteristics very different. For this reason, the after oil crisis variable has 1945-1960 and 1981-1990 as reference periods.

VARIABLES	(1) first inv_logep	(2) second logtransactie	(3) first inv_logep	(4) second logtransactie	(5) first inv_logep	(6) second logtransactie
After oil crisis	0.233***					
	(0.00705)					
Log(gas consumption per m2)			-0.0446***			
			(0.00384)			
Log(electricity consumption)					-0.0401***	
					(0.00947)	
Inverse Log(Energy Performance)		0.0400*		-0.117*		-0.989***
		(0.0229)		(0.0642)		(0.297)
Control variables:						
Structural characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-5.446***	8.162***	-5.066***	7.726***	-5.083***	3.043*
	(0.243)	(0.223)	(0.284)	(0.398)	(0.283)	(1.601)
Observations	20,757	20,757	11,196	11,196	12,430	12,430
R-squared	0.589	0.911	0.605	0.913	0.611	0.748
IV F-stat		1095		134.5		17.95
Durbin pval		0.770		0.00422		2.95e-09

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

Table 6.4: IV estimation of the three potential instrumental variables (dependent variable is the log of the transaction price)

Column 2 of Table 6.4 shows that the price premium of an additional 10% decrease in EPI score equals to 0.4% transaction price increase. However, this finding is only significant at 10%, whilst the prior results have been consistently significant at the 1% benchmark. This result would imply that endogenous variables in the earlier models (section 6.1) influenced the price premium and that the price effect is less evident when using an exogenous estimator. In response to this finding, one might argue that unobserved housing characteristics or preferences might have changed over time. The instrumental variable would be vulnerable to these changes as it is over a long time period (before 1905 to 2018). For this reason an additional regression is done for the time period 1945-2000 (about 30 years before and after the oil price crisis), again using 1945-1960 and 1981-1990 as reference groups. Table 6.5 shows the results from the whole sample in column 1 and 2, and the sample with only buildings from 1945-2000 in column 3 and 4.[35] Both models control for the same variables as the earlier OLS regressions. The results indicate that the oil price crisis is a valid instrument in the time period 1945-2000 (see Durbin-Watson test and the F-stat). Interestingly the new results show that a decrease of the EPI score by 10% leads to a significant 0.7% price premium. These results indicate that potential unobserved changes in housing characteristics or preferences have influenced the regression in column 1 and 2.

Conclusively, the instrumental variable approach has shown varying results. The most appropriate instrumental variable in this research is the oil price crisis in 1974 with as reference points: 1945-1960 and 1981-1990. The first regression over the whole dataset indicates that a 10% decrease in EPI score has a 0.4% price premium, but only at 10% significance. This could potentially mean that endogenous factors have influenced the earlier results. However, one could argue that unobserved housing characteristics might have changed over time and influenced the results. For this reason, the same analysis is performed on houses solely from 1945-2000. This way the timeframe for potential change of unobserved housing characteristics is limited. The results showed that if the EPI score decreases by 10% this yields a price premium effect of 0.7%. This price premium is larger than the initial findings from section 6.1. An approach (also used in section 6.1) to make the results of the EPI score more tangible is to show the jump in EPI score from energy label D to A (about 38.64%). The results from the IV estimation indicate a price premium of 2.7% for the label D to A transformation.

---

[35] Appendix F, Table F2 demonstrates the complete results of the IV regression of Table 6.5.

VARIABLES	(1) Full sample first inv_logep	(2) Full sample second logtransactie	(3) 1945-2000 first inv_logep	(4) 1945-2000 second logtransactie
After oil crisis	0.233***		0.231***	
	(0.00705)		(0.00695)	
Inverse Log(Energy Performance)		0.0400*		0.0699***
		(0.0229)		(0.0231)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-5.446***	8.162***	-5.079***	8.635***
	(0.243)	(0.223)	(0.242)	(0.222)
Observations	20,757	20,757	15,885	15,885
R-squared	0.589	0.911	0.542	0.896
IV F-stat		1095		1110
Durbin pval		0.770		0.160

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

Table 6.5: IV estimation over the full sample and the sample from 1945-2000 (dependent variable log transaction price)

The research of Aydin, Brounen and Kok (2018) used the same instrumental variable but did find a price premium of about 2.2% per 10% increase in energy efficiency. This price premium is substantially higher than the 0.7% price premium in this research. A potential explanation might be that their dataset is from 2008 – 2011, in this time period the EPCs were not mandatory yet (from 2015 onwards). Therefore, there might be a sample selection bias in this dataset, because only more energy efficient houses would request a label or EPI score in 2008-2011. For this reason, the dataset of this research (2009-2018) would be more representative for the effect in the housing market. In addition, the research of Aydin, Brounen and Kok (2018) has less control variables and therefore a lower explanatory value. Presumably this is the case because they use a dataset of all Dutch provinces. Nevertheless, this lower explanatory value might have affected the results upwards.

All in all, the instrumental variable approach indicates a higher price premium for an additional 10% decrease in EPI score than the hedonic pricing model estimated by OLS. This most likely has to do with the endogenous variables[36] that influence both the transaction price and EPI score in the hedonic pricing model estimated by OLS. In conclusion, the IV estimation from this research presents a price premium of 0.7 for an additional 10% decrease in EPI score. This result is relatively low compared to the earlier literature. Nonetheless the non-obligatory inclusion of EPCs in a transaction prior to 2015 might have created a sample selection bias that overvalues the price premium of energy efficient houses. Hence, the sample in this research is arguably more representative for the calculation of the price premium in the housing market.

6.4 Robustness check

In order to check the robustness of the results from section 6.1 and 6.3 this section discusses different alterations to the models. All in all, the results in this study have been robust for most additional checks.

First, in section 6.1 alternative measures for energy efficiency have been used, namely: EPI score, energy labels and the categorization of green and bad labels. The similar results show the robustness of these findings in the hedonic pricing model estimated by OLS. Second, as mentioned in the data matching procedure the alternative abbreviations were dropped. As a robustness check all the alternative abbreviations (e.g. B2, bis- etc.) were numbered with 50 and the datasets was merged again with these new IDs. The same regressions as in Table 6.1 and 6.3 were performed with this new and larger dataset.[37] The results indicate a lower price premium and R^2, this is as expected as the chance of wrongly matched dwellings is greater. Hence, the effect would be lower as faulty matches reduce the explanatory value. Conclusively, the results are robust and the decision to drop the alternative abbreviations seems to be the right one. Third, one might argue that the energy efficiency of a house is stationary. For this reason, the argument to drop observations if the label was not registered at the time of transaction might suboptimal, as it makes the dataset smaller. The same regressions are performed without dropping these 'non-active'

---

[36] Potential examples could be: stained-glass or a classic façade.

[37] Appendix G, Table G1 and G2 shows the OLS regression from Table 6.1 and the IV regression from Table 6.3 with the inclusion of the alternative abbreviations.

labels.[38] The sample size almost doubles, however the results become smaller and some of the significance decreases. This legitimizes the decision of this study to drop the non-active labels. This robustness check indicates that labels are not stationary and may change over time. Thus one should control for this by only taking active labels at the moment of transaction.

Fourth, the control variables are checked for in various ways. The individual variables were checked beforehand on outliers via the kurtosis and skewness. However the amount of dummies for the year-month and postal code could be a point of discussion.[39] For this reason the same regressions are performed with dummies for the three-digit postal code and the two-digit postal code. Additionally, the time dummies are changed to year-quarter and year.[40] The results are robust for the postal codes, without barely any differences. However, when the year-quarter or year dummies are used the significance of the model and the explanatory value decreases. This makes sense as the price fluctuations from a whole year are very hard to capture in one dummy and thus month-year dummies should be used.

Fifth, in order to explore the impact of the mandatory inclusion of EPCs in 2015. The dataset is split up in two different time periods, the first with transactions before the inclusion, thus 2009-2014. The second with transactions from after the inclusion, 2015-2018.[41] The same IV analysis is performed as in section 6.3. Interestingly, the price premium from 2009-2014 is 0.89% per 10% decrease in EPI score. This is a larger price premium than the findings from the whole sample and thus of this study. The regression over the period 2015-2018 portrays the surprising results that the significance of the price premium for EPI score vanishes. This could potentially mean that the findings in this research are driven by the period before the mandatory inclusion of EPCs. For this reason, this finding confirms the belief that a potential sample selection bias exists in studies that use datasets from before 2015. Nonetheless, to bring nuance to these findings over 2015-2018, the dataset only contains just over 6.000 observations, and the timeframe is relatively

[38] Appendix G, Table G3 performs the robustness check with non-active labels covering the OLS regression from 6.1

[39] The corresponding numbers of dummies are: four-digit postal code: 478; three-digit postal code: 114; two-digit postal code: 15; year-month: 108; year-quarter: 36; year: 9

[40] Appendix G, Table G4 shows the robustness check covering the OLS regression from 5.1 with different forms of postal code and time series dummies.

[41] Appendix G, Table G5 and G6 show the regression output of the two different time periods.

small. Therefore, definite statements about this finding cannot be made. Nevertheless, this finding is engaging and very interesting for further research.

Sixth, an additional check is performed with the green labels (A, B, and C) as independent variable in the IV estimation instead of the inverse of log (EPI)[42]. The results indicate a positive price premium of around 3.3% for a green label. However, the Durbin-Watson test indicates that the after oil crisis variable is less exogenous than in the IV estimations of section 6.3.[43] Nevertheless, the robustness check confirms the existence of a price premium for energy efficiency. At last, an additional robustness check is performed for the instrumental variable approach. The after oil crisis variable as explained in section 6.3 did represent houses constructed after 1980, because the data of the building years has been categorized in time periods. However, one might argue that this excludes the years 1974-1980. For this reason a robustness check is performed wherein the after oil crisis dummy presents houses from 1970 onwards (instead of 1980).[44] The findings from this additional IV estimation are robust with the findings in section 6.3. The same pattern occurs, the whole sample indicates that an additional 10% decrease in EPI has a price premium of 0.4% (significant at the 10%). The sample from 1945-2000 shows a price premium of 0.7% (significant at 1%) per 10% additional decrease in EPI score.

6.5 Investing in energy efficiency

In order to make the results of this research more tangible and put them into perspective, different cases are discussed in this section. Table 6.6 provides an overview of four different business cases based on information from Van Hoek and Koning (2018) and the extension by Van Hoek, Koning and Stiemer (2018) who performed a research in name of the EIB (Economisch Instituut voor de Bouw). The articles estimate the energy efficiency investments needed for a label transformation, but also the corresponding yearly energy cost savings in the Netherlands.

The business cases contain the energy efficiency investment and the corresponding yearly savings for the label transformation in an owner-occupied housing market. The investment costs

---

[42] Appendix G, Table G7 performs the IV estimation as in Table 6.3 with green label as independent variable.

[43] The Durbin-Watson tests indicates that over the whole sample the value is 0.06, and for the period 1945-2000 the value is 0.05 (rounded upwards). Hence the credibility of the instrument is less then with the inverse log (EPI) as dependent variable.

[44] Appendix G, Table G8 performs the IV estimation as in Table 6.3 with the alternative estimation of the after oil crisis variable.

are seen as a loan and therefore calculated as an annuity, called the yearly costs.[45] In addition, the transaction price premium is calculated based on the change of EPI score and the results from section 6.3 (per 10% energy efficiency increase a price premium of 0.7%). The average price of a dwelling is taken over the whole dataset (see Table 5.2). Last, the present value of the cost savings is calculated.[46] The four different cases are analyzed for three components: first, the investment gap between yearly costs and savings. Second, the energy efficiency investment compared to the transaction price change. Last, the capitalization of the present value of the yearly savings in relation to the transaction price premium.

	Label D to label A	Label G to label A	Label G to label C	Label G to label D
Energy efficiency investment	€ 12,000	€ 23,000	€ 14,000	€ 11,000
Yearly costs	€ 940	€ 1,799	€ 1,095	€ 860
Yearly savings	€ 550	€ 1,650	€ 1,300	€ 1,100
Investment gap per year	€ (390)	€ (149)	€ 205	€ 240
Average house price	€ 247,671	€ 247,671	€ 247,671	€ 247,671
Transaction price premium	2.7%	4.2%	3.0%	2.1%
Transaction price change	€ 6,687	€ 10,511	€ 7,420	€ 5,255
Present value cost savings	€ 7,031	€ 21,093	€ 16,618	€ 14,062
%	95.1%	49.8%	44.6%	37.4%

Table 6.6: This overview presents the business cases for the specific label jumps.

6.5.1 Business case synthesis

The first finding from Table 6.6 is the variation in the investment gap per year (wherein the brackets indicate a negative number). The results indicate that the label transformations from label D to A and label G to A provides a negative investment gap per year. On the other hand, the label transformation from label G to C or label G to D show a positive investment gap of above 200 EUR per year. In essence, this means that the yearly savings outweigh the yearly costs, and therefore it is in general financially favorable to invest in these label transformations. This is evidence for the energy efficiency gap as introduced by Allcott and Greenstone (2012) and hence confirms hypothesis 4 as constructed in the conceptual framework. Furthermore, this might be an

---

[45] Van Hoek and Koning (2018) propose an interest rate of 6% and a time span of 25 years. This is used in all the annuity calculations.

[46] For the present value calculation the recommendations of van Hoek and Koning (2018) are used again, an interest rate of 6% and a time span of 25 years.

explanation to why label C and D are the most frequently occurring energy labels (see Table 5.4). Alternatively, an investor could buy a house, transform it to a better label and then sell it again. However, the results from Table 6.6 show that this is not profitable. None of the four cases indicate that the transaction price premium for more energy efficient houses outweigh the energy efficiency investment. Hence, this business model is not interesting from a financial standpoint based on these results. Last, the difference between the present value of the energy efficient cost savings and the transaction price premium show that it is not fully capitalized in the transaction price. From label D to A there is a (high) 95.1% of market capitalization, however the transformation from label G to A, C, or D the market capitalization is only about 37.4 to 49.8%. This finding illustrates that the market only partially capitalizes on the energy cost savings. Building further upon this result, this implies that there exists an investment opportunity in houses that typically would have label G but are transformed to label A, C, or D.[47] However, to bring nuance to these findings, van Hoek and Koning (2018) state in their paper that their indications are a best case scenario. As potentially the price of energy might not be stable, or the energy cost savings might be lower than expected. On the other hand, the time span (25 years[48]) and interest rate (6%[49]) are relatively conservative. Nevertheless, van Hoek and Koning (2018) state that the numbers of their calculations are more likely to be lower in practice. This is the complexity of an energy efficiency business case as described by Wilkinson and Sayce (2019). These findings should therefore only be seen as a rough indication, this has to do with the heterogeneity of the residential housing market. Every house is unique in its characteristics and therefore you cannot generalize the financial feasibility of certain label transformation for all houses.

[47] Only four cases have been researched in this study, hence other label transformation might also yield similar effects.

[48] One could potentially argue that 25 years is a relatively long time span. However, many sustainable investments have a very long duration. Think of double glass or facade isolation.

[49] This is currently relatively high for sustainable investments. However, these numbers are taken in order to be more conservative. Moreover, if the interest rate would actually go down this would mean that the investment gap is even larger, and the market capitalization rate even lower, than in the current calculations of Table 6.6.

7. Conclusion

The political and scientific debate is currently ongoing about actions that need to be taken in order to decrease our carbon emission footprint. An important discussion within this topic is how we can make the built environment more energy efficient. Recent initiatives like the EPCs and the EPI score have increased the tangibility and transparency of energy efficiency. However, the exact capitalization of energy efficiency in the housing market is ambiguous. Moreover, the valuation of energy efficiency by the market could potentially bring important insights for the carbon emission discussion. For this reason, this research explores the price premium of energy efficiency in the residential real estate market in the Netherlands.

Most of the literature on this topic did use datasets prior to the regulation (2015) that has made the inclusion of EPCs mandatory in the transaction of a house. This political regulation highlights the relevance for more research into the price premium of energy efficiency. In order to enhance the statistical validity and credibility this study makes use of three different methodologies. First, the traditional method in the housing market, the hedonic pricing model. Second, the repeat sales model. Last, the instrumental variable approach. All three methods separately indicate a positive price premium for more energy efficient houses. The hedonic pricing model estimated by OLS indicates a 0.46 price premium for a 10% decrease in EPI, a positive price premium of 3.19% for a house with label A and a negative price premium of 2.02% for a dwelling with label G.[50] Moreover, dwellings with a green label could command a positive price premium of 2.64% and dwellings with bad label a negative price premium of 1.76%. Furthermore, the research into the EPCs indicates that the market yields a larger positive price premium for more energy efficient houses than it has a negative price premium for less energy efficient houses. In addition, an alternative approach to the hedonic pricing model estimated by OLS is to research the effect of more isolation methods by the use of a repeat sales model. Isolation is an important component of energy efficiency and one of the drivers of the EPI score and EPCs. The findings indicate an 8.86% price premium per additional isolation method. However, to nuance this finding this is only a percentage over the price change between the first and second transaction. Nevertheless, definite conclusions cannot be made with the repeat sales method about the price

[50] The reference group for the price premiums of the labels A and G is the neutral label D.

premium of energy efficiency. This should be seen as a directive and therefore the existence of a positive price premium. At last the instrumental variable approach is used, this in order to control for potential endogeneity in the model. The findings from the instrumental variable approach shows that an additional 10% decrease in EPI score yields a positive price premium of approximately 0.7% on average. This means that the results from the hedonic pricing model estimated by OLS seen a bias towards zero compared to the IV estimation. A potential explanation for this result might be that omitted variables, like stained-glass or classic facades, would have a positive effect on the transaction price and at the same time have a negative effect on energy efficiency.

All in all, the results are uniform about the existence of a positive price premium. The most advanced methodology that controls for endogeneity shows that a 10% increase of energy efficiency has a price premium of approximately 0.7% on average. The results of this study seem to be on the lower end compared to the earlier literature. However, this study has a relatively high explanatory value (R^2 of 0.90) and makes use of a more recent dataset (including years after the mandatory inclusion of the EPCs). The findings of this research could be used to further promote energy efficiency investments in the residential housing market. This study shows that energy efficiency investments not only lead to the energy cost savings over time, but that it also yields a positive price premium on the transaction price. At last, the results have been put into perspective in four different label transformation business cases. The findings indicate that the market does not fully capitalizes on the energy cost savings and that investment inefficiencies exist in the residential real estate market. This information could stimulate dwellings on an individual level to increase their energy efficiency and therefore reduce their own carbon emission footprint. Potential public or private promotion campaigns should accentuate these direct and indirect financial effects of the investment in energy efficiency in the residential real estate sector.

The limitations of this research are mainly found in the data compartment. As mentioned in section 5.2 the complexity to merge the alternative abbreviations across the different datasets. This research did have merge rate with the individual datasets of about 60%, thus there is room for improvement in this field. Furthermore, this research did not have access to data over time of the EPCs and EPI scores due to privacy regulations. However, if this data would be available the repeat sales model could be a great alternative way to present an accurate estimation of the price premium.

This research did only have limited data over time, therefore only an indication (based on isolation manners) could be given instead of a definite number.

In the upcoming years this topic will most likely remain of high relevance. For this reason more research is needed in this particular field. As mentioned the EPCs did become mandatory in 2015, therefore it might be interesting to perform the same analysis with the same research methods in a couple of years. The robustness check has shown the large price premium disparity between the years before 2015 and after. This is a very interesting phenomenon and a great direction for further research. Additionally, an interesting field of research in the hedonic pricing methodology might be the price premium of the water level on residential real estate in the upcoming 100 years in the Netherlands. Does the water level need to be taken into account in the hedonic pricing method at one point?

8. Bibliography

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

Aroul, R. R., & Rodriguez, M. (2017). The Increasing Value of Green for Residential Real Estate. *Journal of Sustainable Real Estate*, *9*(1), 112-130.

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

Aydin, E., Brounen, D., & Kok, N. (2018). Information provision and energy consumption: Evidence from a field experiment. *Energy Economics*, *71*, 403-410.

BP (2017). Statistical Review of World Energy 2017, retrieved 23 March 2019, from https://bp.com/statisticalreview/

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

Centraal Bureau voor de Statistiek (2019). Regionale kerncijfers Nederland, retrieved 11 March 2019, from https://opendata.cbs.nl/statline/

Centraal Bureau voor de Statistiek (2019). Bestaande koopwoningen; verkoopprijzen prijsindex, retrieved 18 March 2019, from https://opendata.cbs.nl/statline/

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

Chau, K. W., & Chin, T. L. (2002). A critical review of literature on the hedonic price model.

European Commission (2015). Paris Agreement, retrieved 14 August 2018, from https://ec.europa.eu/clima/policies/international/negotiations/paris_en

European Commission (2016). Climate strategies and targets: 2020 climate and energy package, retrieved 14 August 2018, from https://ec.europa.eu/clima/policies/strategies/2020_en

Fregonara, E., Rolando, D., & Semeraro, P. (2017). Energy performance certificates in the Turin real estate market. *Journal of European Real Estate Research*, *10*(2), 149-169.

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

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

Fuerst, F., McAllister, P., Nanda, A., & Wyatt, P. (2016). Energy performance ratings and house prices in Wales: An empirical study. *Energy Policy*, *92*, 20-33.

Havlınová, J., Jacobs, B., Berkvens, M. L., van Dijk, D. D., & Gerritsen, A. A. F. (2018). The Value of Energy Performance in the Dutch Housing Market.

Hertwich, E. G., & Peters, G. P. (2009). Carbon footprint of nations: A global, trade-linked analysis. *Environmental Science & Technology, 43*; 6414–6420

Horne, R. and Hayles, C. (2008). 'Towards global benchmarking for sustainable homes: an international comparison of the energy performance of housing', *Journal of Housing and the Built Environment, 23*; 119-130.

IPCC (2007). Climate Change 2007: Mitigation of Climate Change: Working Group III Contribution to the Fourth Assessment, Report of the Intergovernmental Panel on Climate Change, Cambridge University Press

Institute for Market Transformation (2019). Netherland's Energy Performance Certificate, retrieved 13 March 2019, from https://www.buildingrating.org/

Ivanova, D., Vita, G., Steen-Olsen, K., Stadler, K., Melo, P. C., Wood, R., & Hertwich, E. G. (2017). Mapping the carbon footprint of EU regions. *Environmental Research Letters*, *12*(5), 054013

Jepson Jr, E. J., & Edwards, M. M. (2010). How possible is sustainable urban development? An analysis of planners' perceptions about new urbanism, smart growth and the ecological city. *Planning Practice & Research*, *25*(4), 417-437.

Kats, G. (2003). Green building costs and financial benefits (p. 1). Boston: Massachusetts Technology Collaborative.

Krause, A. L., & Bitter, C. (2012). Spatial econometrics, land values and sustainability: Trends in real estate valuation research. *Cities, 29*, S19-S25.

Lorenz, D., Trück, S., & Lützkendorf, T. (2006). Addressing risk and uncertainty in property valuations: a viewpoint from Germany. *Journal of property investment & finance, 24*(5); 400-433.

Lorenz, D. P., Trück, S., & Lützkendorf, T. (2007). Exploring the relationship between the sustainability of construction and market value: Theoretical basics and initial empirical results from the residential property sector. *Property Management, 25(*2); 119-149.

Lorenz, D., & Lützkendorf, T. (2008). Sustainability in property valuation: theory and practice. *Journal of Property Investment & Finance*, *26*(6), 482-521.

Malpezzi, S. (2003). Hedonic Pricing Models: A Selective and Applied Review, w Housing Economics and Public Policy: Essays in honor of Duncan Maclennan, red. T. O'Sullivan, K. Gibb.

Milieu Centraal (2019). Verkopers en Energielabel, retrieved 21 March 2019, from https://www.energielabel.nl/

Mudgal, S., Lyons, L., Cohen, F., Lyons, R., & Fedrigo-Fazio, D. (2013). Energy performance certificates in buildings and their impact on transaction prices and rents in selected EU countries. *European Commission (DG Energy), Paris*.

Oxford Dictionary (2019). *Oxford University Press*, retrieved 26 February 2019, from https://en.oxforddictionaries.com/

Pride, D., Little, J., & Mueller-Stoffels, M. (2018). The value of residential energy efficiency in interior Alaska: A hedonic pricing analysis. *Energy Policy*, *123*, 450-460.

Rijksdienst voor Ondernemend Nederland (2019). Energie Subsidie Wijzer, retrieved 11 March 2019, from https://www.energiesubsidiewijzer.nl/

Rijksdienst voor Ondernemend Nederland (2019). EP-online, retrieved 11 March 2019, from https://www.rvo.nl/

Rijksdienst voor Ondernemend Nederland (2019). Energie-index, retrieved on 18 March 2019, from https://www.rvo.nl/

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

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, *13*(1), 1-44.

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

Van Hoek, T & Koning, M (2018). Klimaatbeleid en de gebouwde omgeving. *Economisch Instituut voor de Bouw*

Van Hoek, T, Koning, M & Stiemer, D. (2018). Klimaatbeleid en de gebouwde omgeving (2). *Economisch Instituut voor de Bouw*

Warren-Myers, G. (2012). The value of sustainability in real estate: a review from a valuation perspective. *Journal of Property Investment & Finance, 30*(2); 115-144.

Wiley, J. A., Benefield, J. D., & Johnson, K. H. (2010). Green design and the market for commercial office space. *The Journal of Real Estate Finance and Economics, 41*(2); 228-243.

Wilkinson, S., & Sayce, S. (2019). Energy efficiency and residential values: a changing European landscape.

Wilson, J., Tyedmers, P., & Spinney, J. E. L. (2013). An exploration of the relationship between socioeconomic and well-being variables and household greenhouse gas emissions. *Journal of Industrial Ecology, 17*; 880–891.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

9. Appendices

Abbreviations

EIB = Economisch Instituut voor de Bouw

EPC = Energy performance certificate

EPI = Energy Performance Index

ID = Identification number

IV = Instrumental variable

OLS = Ordinary least square

RVO = Rijksdienst voor Ondernemend Nederland

UK = United Kingdom

2SLS = Two-Stage Least Squares

Appendix A – General information

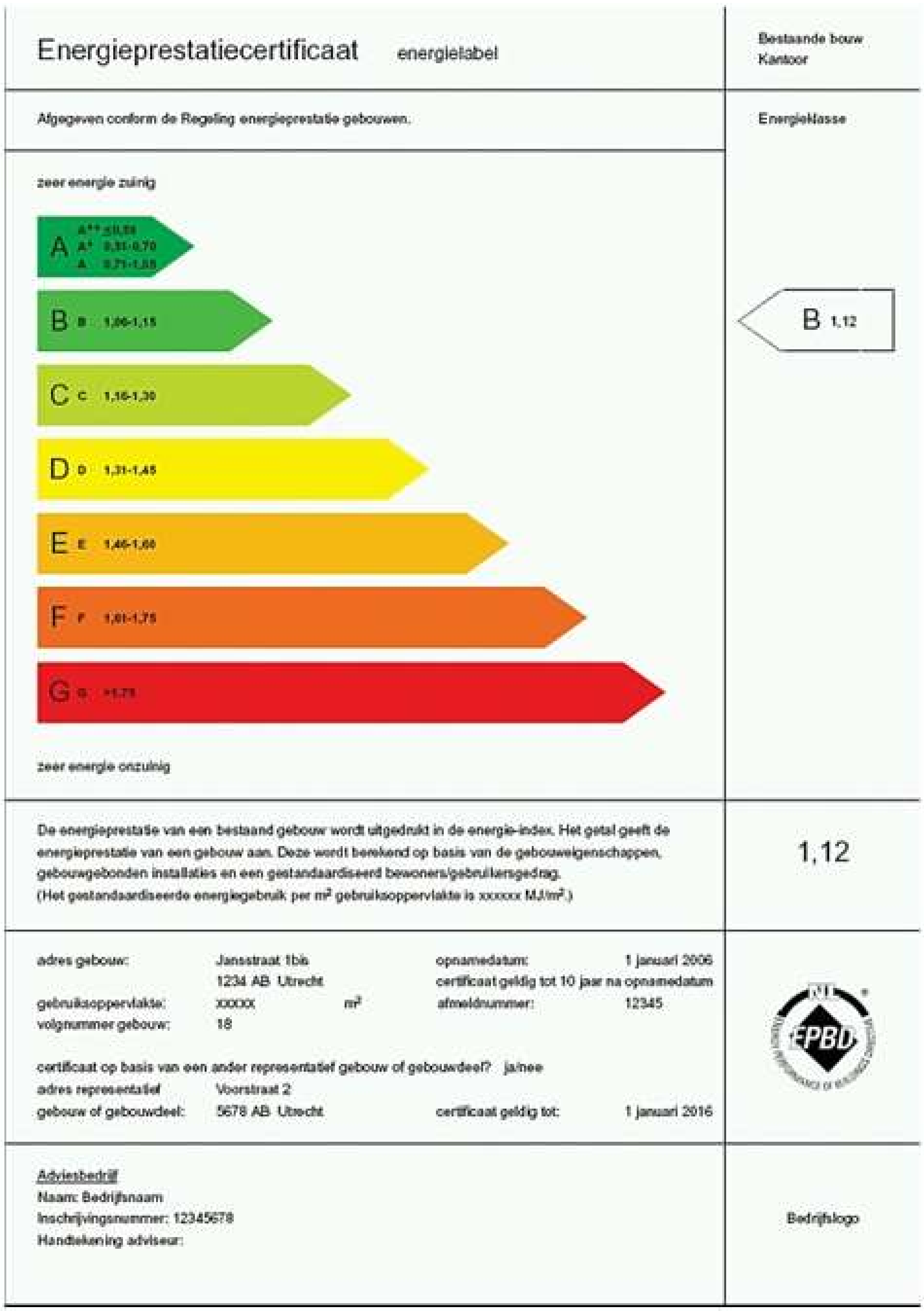

Figure A1. Energy Performance Certificate (EPC) label

(Source: Institute for Market Transformation, 2019)

Het energielabel van je woning wordt bepaald aan de hand van bouwjaar, woningtype, woningoppervlakte en andere woningkenmerken die van belang zijn voor het energieverbruik. Deze kenmerken zijn:

- Soort glas
- Gevelisolatie
- Dakisolatie
- Vloerisolatie
- Soort verwarming
- Aparte voorziening voor warm water badkamer naast CV ketel (geen - keukenboiler)
- Ventilatiesysteem
- Zonnepanelen en zonneboiler

Figure A2. Most important variables in the composition of the EPC label.

(Source: Milieu Centraal, 2019)

Labelletter	Grenswaarden Energie-Index (EI) woningen
A	Kleiner of gelijk aan 1.20
B	1.21 - 1.40
C	1.41 - 1.80
D	1.81 - 2.10
E	2.11 - 2.40
F	2.41 - 2.70
G	Groter dan 2.70

Figure A3: The threshold value for the conversion from an EPI score to an energy label.

(Rijksdienst voor Ondernemend Nederland, 2019, Energie-Index)

Energie-Index: 150 kenmerken

Een Energie-Index mag alleen opgesteld worden door een gecertificeerde energieadviseur (EPA-W, BRL9500-01). Een energieadviseur komt langs en neemt ongeveer 150 kenmerken van de woning op. Dit betekent in de praktijk dat de energieadviseur de afmetingen van de woning in detail opmeet, noteert wat de isolatiekwaliteit is en in kaart brengt welke installaties aanwezig zijn. Op basis van deze uitgebreide opname berekent hij vervolgens het Energie-Index-getal (geen letter).

Figure A4: Elaboration on the Energy-Index (or known as EPI) score.

(Rijksdienst voor Ondernemend Nederland, 2019, Energie-Index)

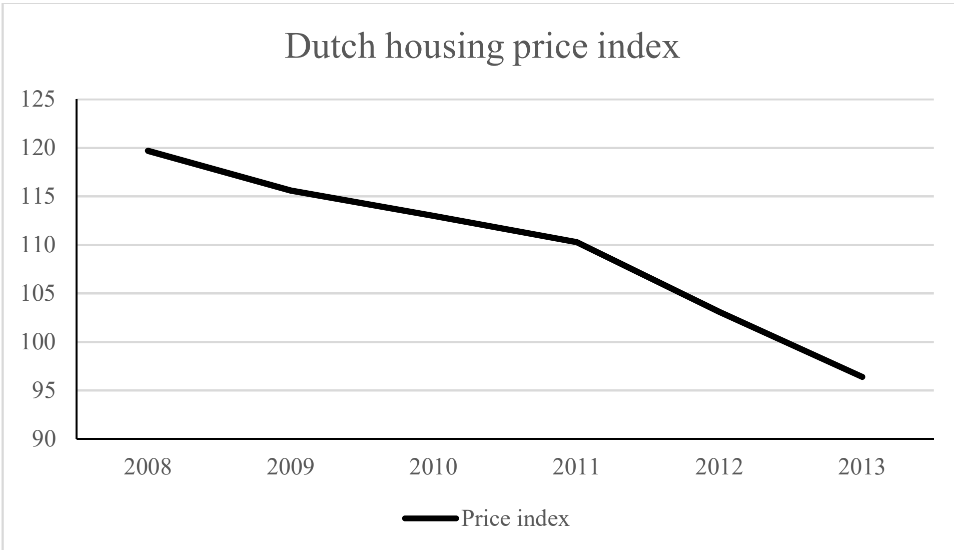

Figure A5. The Dutch housing price index over the period 2008 to 2013

(Centraal Bureau voor de Statistiek, 2019. Bestaande koopwoningen; verkoopprijzen prijsindex)

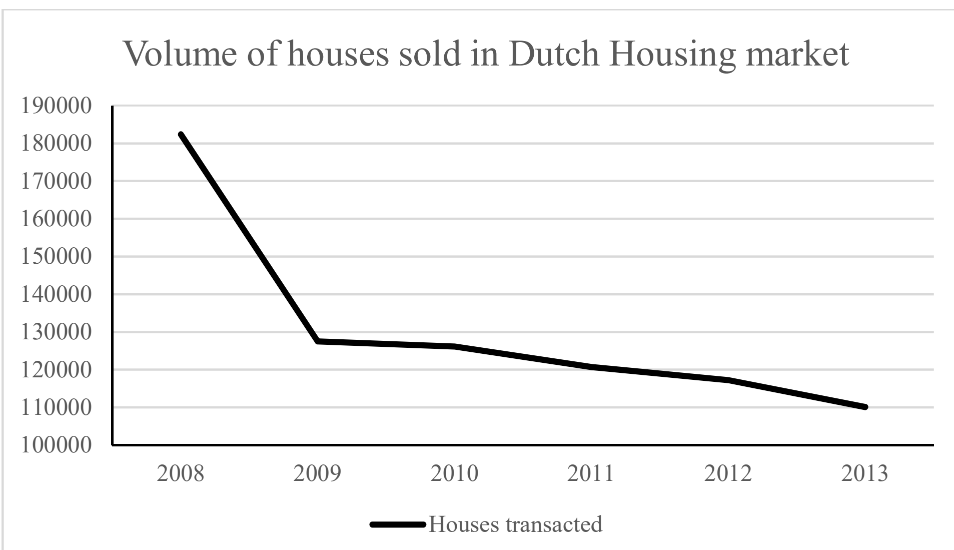

Figure A6. The volume of houses sold in the Dutch housing market over the period 2008 to 2013.

(Centraal Bureau voor de Statistiek, 2019. Bestaande koopwoningen; verkoopprijzen prijsindex)

Appendix B - Common methods in real estate valuation

B1. Hedonic pricing model

To research the effect of energy efficiency on residential real estate the literature is explored for different methods. The literature of energy efficiency is consistent with the recommendations of Lorenz, Trück and Lützkendorf (2007) who were one of the first to propose the hedonic pricing method as an advanced technique to measure sustainability in residential real estate valuation. For this reason, the basis of this research is a hedonic pricing model. Historically, Sherwin Rosen (1974) was the first to introduce the theory of hedonic pricing. He defines hedonic prices as "the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them" (Rosen, 1974, p. 34). Malpezzi (2003) builds further upon this concept and elaborately explains the mathematical application of the hedonic pricing in regards to the housing market (this is more elaborately discussed in the methods section). The most basic hedonic pricing regression [1] with regards to value based on Malpezzi (2003) is:

[1] $$V = f(S, N, L)$$

Where V = value; S = structural characteristics; N = neighborhood characteristics; L= location within market;

This methodology provides a framework for analysis of the price formation of variables that do not have an obvious market value. The analysis is conducted by a multiple regression model in order to compare different characteristics, and evaluate the differences with respect to the price driving factors. Sirmans, Macpherson & Zietz (2005) extensively explored the existing body of literature about hedonic pricing methods and the corresponding relevant characteristics. In their findings the researchers present a list of the twenty characteristics appearing the most frequent in the literature. In addition to this findings they established eight overarching categories[51]. These categories and variables are used as a framework to identify potentially relevant variables in the dataset of this study.

---

[51] See Appendix C for the twenty most frequent variables and the eight categories

B2. Repeat sales model

Additionally Malpezzi (2003) discusses the repeat sales model. In essence this model analysis units that have sold at least twice in a given period of time. To mathematically grasp this method the following formula [2] can be used as interpretation.

[2] $\ln P = X\beta + \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + B_4 T_4$

Where P = value; X = vector for all relevant characteristics, including a constant term; and T_i = time dummies. To illustrate this effect, Malpezzi (2003) hypothesized that an item is sold in period 2 and 4. Therefore we can mathematically deduce the following from formula [2]

[3] $\ln P_4^A - \ln P_2^A = X\beta + \beta_4 T_4 - X\beta - \beta_2 T_2$

[4] $\ln P_4^A - \ln P_2^A = \beta_4 T_4 - \beta_2 T_2$

Formulas [3, 4] illustrate that the relevant characteristics drop out and this gives a feeling of how the function works. The main advantage of the method is that it does not require a database with detailed information on characteristics and therefore it can be used on widely available datasets. However this method also has its downsides, for the same reasoning this will only give an estimate of the price change. Furthermore, it only gives information about a fraction of the housing market and it assumes that no alterations have been made to the real estate object. Nonetheless, this method is interesting to perform complementary to the hedonic pricing model and potentially annualize the percentage growth over time.

B3. Instrumental variable approach

Following the methodology of Aydin, Brounen and Kok (2018) this research makes an extension in order to control for potential endogenous effects within the hedonic pricing method. For this reason to strengthen the statistical credibility a two stage least squares instrumental variable approach is applied based on the book of Wooldridge (2015). To illustrate the effect the simple regression [5] is used:

[5] $y = \beta_0 + \beta_1 x + u$

If the researcher suspects that x and u in the regression [5] are correlated this might create a bias in the regression results. For this reason, the researcher needs to find a β_0 and $\beta_1$where x and u are uncorrelated. Therefore we need to find an observable variable z that satisfies the following two assumptions [6, 7]:

[6] $Cov(z,u) = 0$ [7] $Cov(z,x) \neq 0$

This essentially means that: first, variable z should not have a partial effect on y. Second, z should not be correlated with the omitted variables. Last, z must be related to variable x.

Appendix C – Variables and categories from the literature

Categories within the hedonic pricing model
1. Construction and Structure
2. House internal features
3. House external features
4. Environmental - Natural
5. Environmental - Neighborhood and Location
6. Environmental - Public service
7. Marketing, Occupancy and Selling
8. Financial Issues

Figure C1: The categories within the hedonic pricing model used as framework in this paper. Note: overarching categories as identified by Sirmans, Macpherson and Zietz (2005). (Sirmans, Macpherson and Zietz, 2005)

Variables	Effect
Lot size	Positive
Ln Lot size	Positive
Square feet	Positive
Ln Square feet	Positive
Brick	Positive
Age	Negative
Number of stories	Negative
Number of Bathrooms	Positive
Number of Rooms	Positive
Bedrooms	Ambiguous
Full baths	Positive
Fireplace	Positive
Air-Conditioning	Positive
Basement	Positive
Garage Spaces	Positive
Deck	Positive
Pool	Positive
Distance	Ambiguous
Time on Market	Negative
Time Trend	Ambiguous

Figure C2: Expected effects of the variables from Sirmans, Macpherson and Zietz (2005). Note: The expected effect of the variables is based on the literature review of Sirmans, Macpherson and Zietz (2005). The higher in the list the more frequent the variable occurred in the literature.

(Sirmans, Macpherson and Zietz, 2005)

Appendix D – Results concerning methodology

	Inv_ln~I	isolat~n
Inv_lnEPI	**1.0000**	
isolation	**0.4605**	**1.0000**

Figure D1: shows the correlation matrix between inverse log (EPI) and isolation ways as constructed in Stata

	isolatie	labels
isolatie	**1.0000**	
labels	**0.5561**	**1.0000**

Figure D2: shows the correlation matrix between energy labels and isolation ways as constructed in Stata

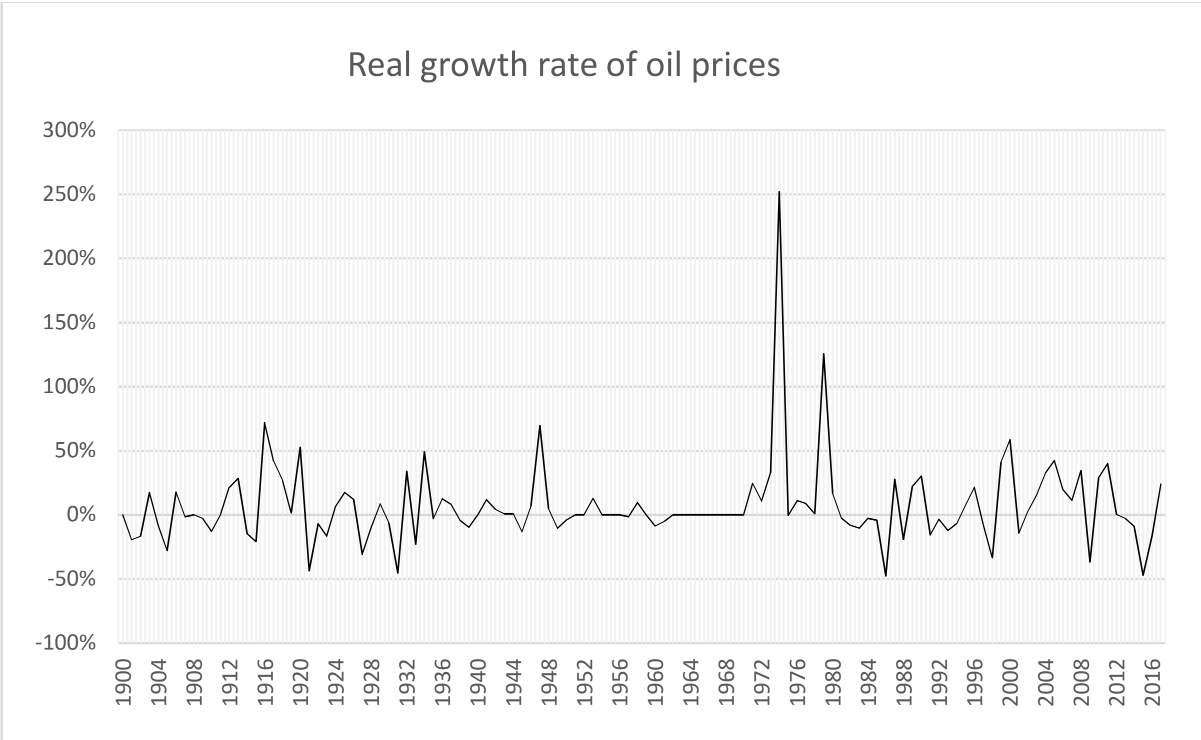

Figure D3: shows the real growth rate of oil prices (1974 has an increase of 252%)

Source: BP (2017), Statistical Review of World Energy 2017

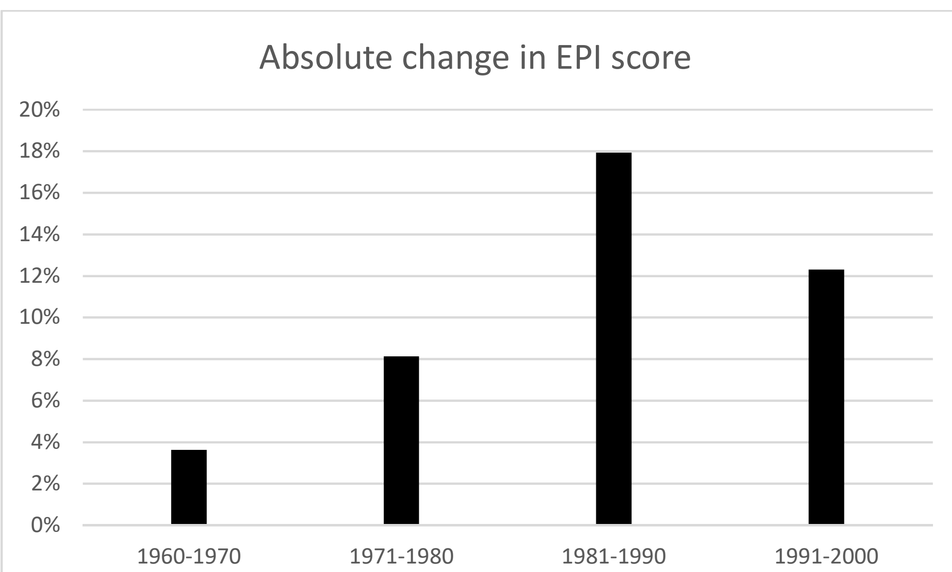

Figure D4: shows the absolute change in EPI score over the years 1960-2000

Appendix E – Variables, distribution and scatter plots

Themes		Variables	
Strucutural characteristics		Description	type
Ln lot size	=	The lot size in m3	continuous
Ln lot size	=	The lot size in m2	continuous
Age	=	The building year of the dwelling	dummies
Number of stories	=	The numers of stories	discrete
Type of dwelling	=	The type of dwelling (categorized by NVM)	dummies
Quality appartment	=	The quality of the apartment	dummies
Ratio m3/m2	=	Ratio between m3 and m2	continuous
Type of house	=	Type of house	dummies
Internal features			
Number of rooms	=	Number of rooms	discrete
Number of toilets	=	Number of toilets	discrete
Heating system	=	Type of heating system	dummies
Amount of isolation	=	Amount of isolation available	dummies
Number of domers	=	Is there a dormer?	discrete
Number of pantries	=	Is there a pantry	discrete
Attick	=	Is there an attick	dummy
External features			
Parking space	=	Type of parking spots	dummies
Environmental - Natural			
Location is beautiful	=	Is the location beautiful	dummies
Environmental - Neighborhood & Location			
Location postal code	=	Dummy for postal codes	dummies
Location road	=	How is the location compared to the road	dummies
Location city center	=	How is the location compared to the city center	dummies
Environmental - Public service			
Marketing, occupancy and Selling factors			
Condition of house outside	=	Maintenance condition outside	dummies
Condition of house inside	=	Maintenance condition inside	dummies
Time effects month	=	Year-month dummies	dummies
Financial issues			

Figure E1: presents the variables of the study categorized in the themes of Sirmans, Macpherson and Zietz (2005). This is the framework used for the regressions.

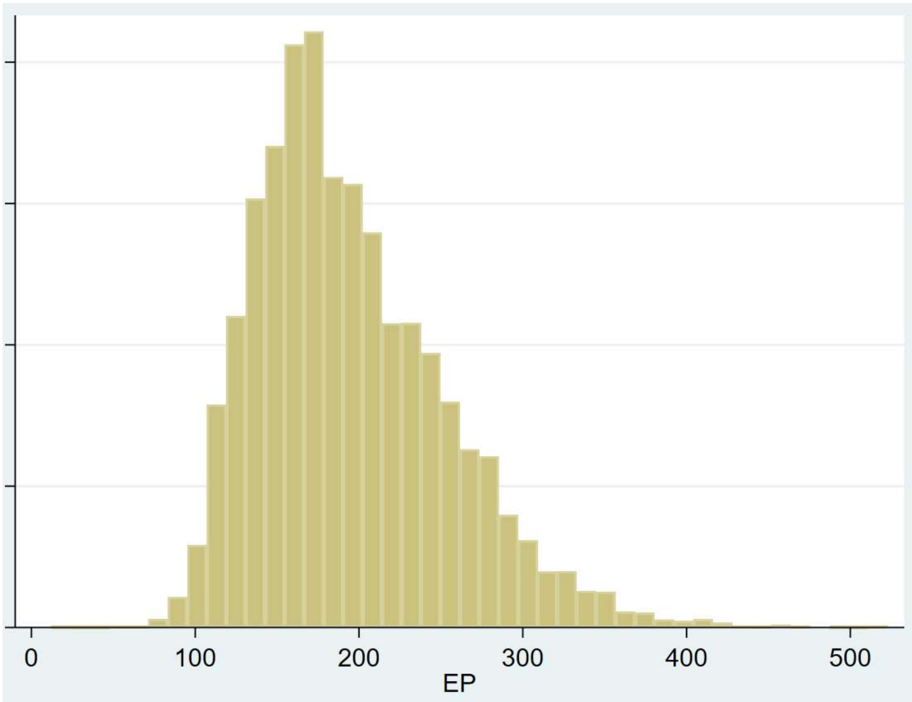

Figure E2: presents the distribution of the EPI scores (x-axis). The y-axis presents the density.

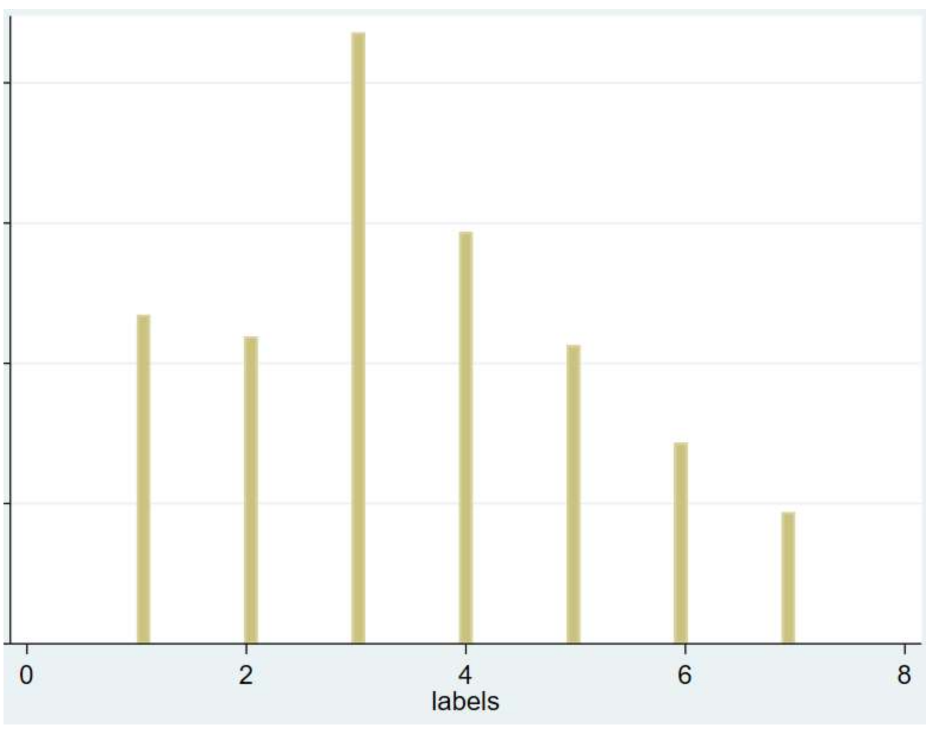

Figure E3: present the energy label distribution (x-axis) wherein A=1 …. G = 7. The y-axis presents the density.

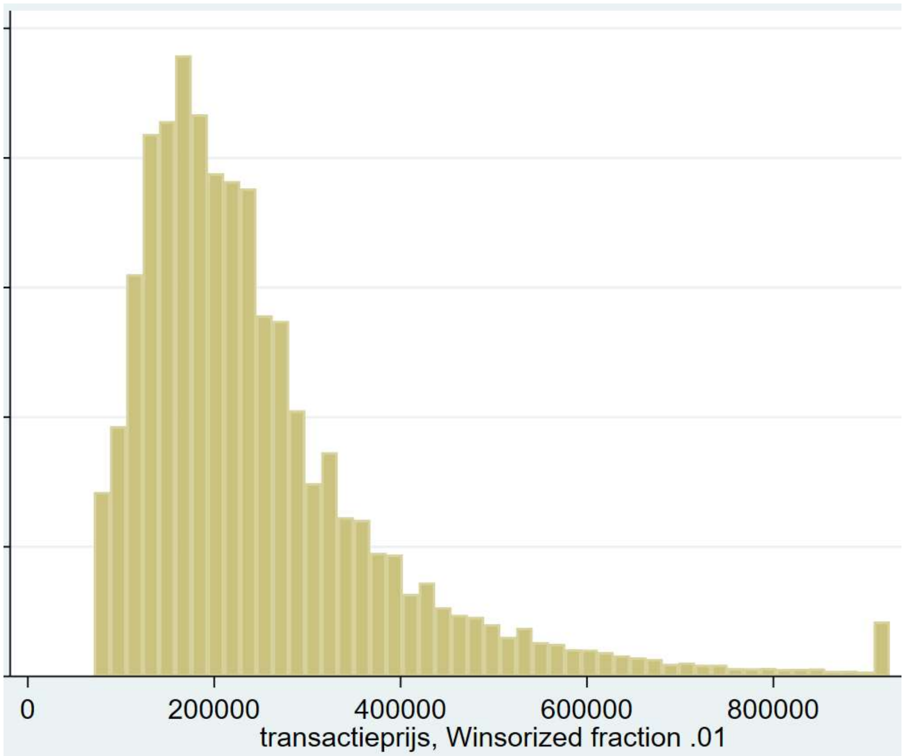

Figure E4: distribution of the winsorized transaction prices (x-axis). The y-axis presents the density.

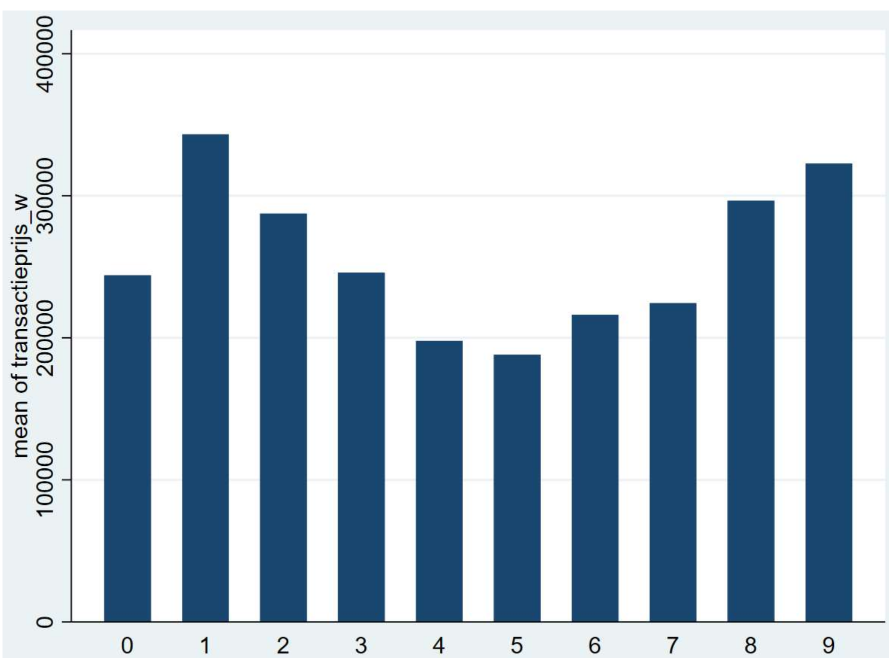

Figure E5: average transaction price (y-axis) per building year (x-axis) category 0 = before 1905, 9 = after 2001

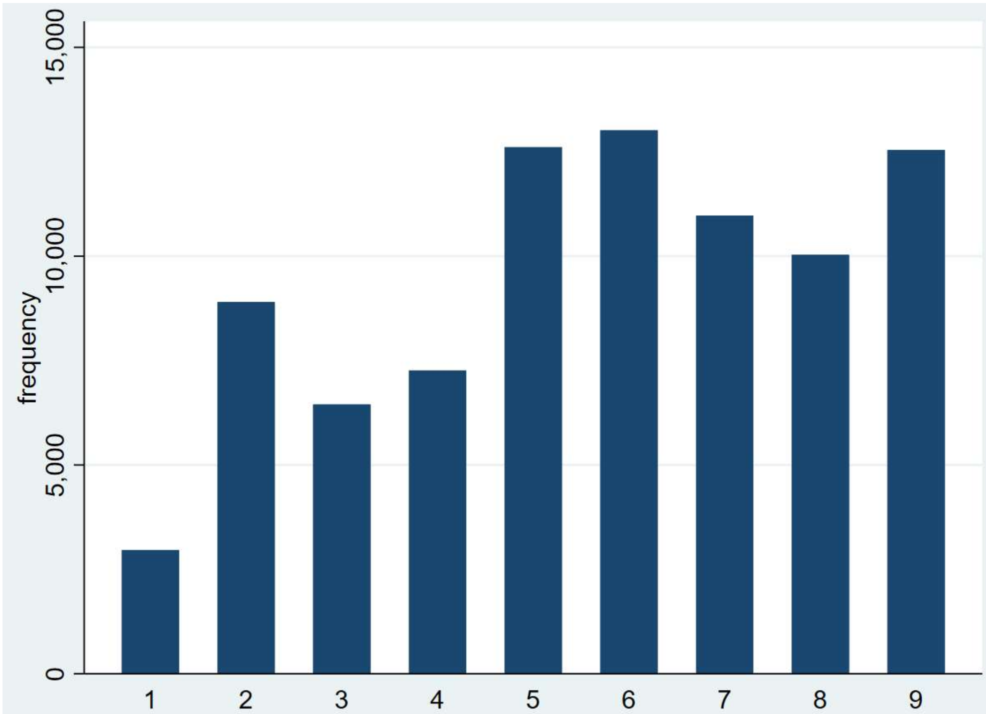

Figure E6: average volume (y-axis) per building year category (x-axis)

0 = before 1905, 9 = after 2001

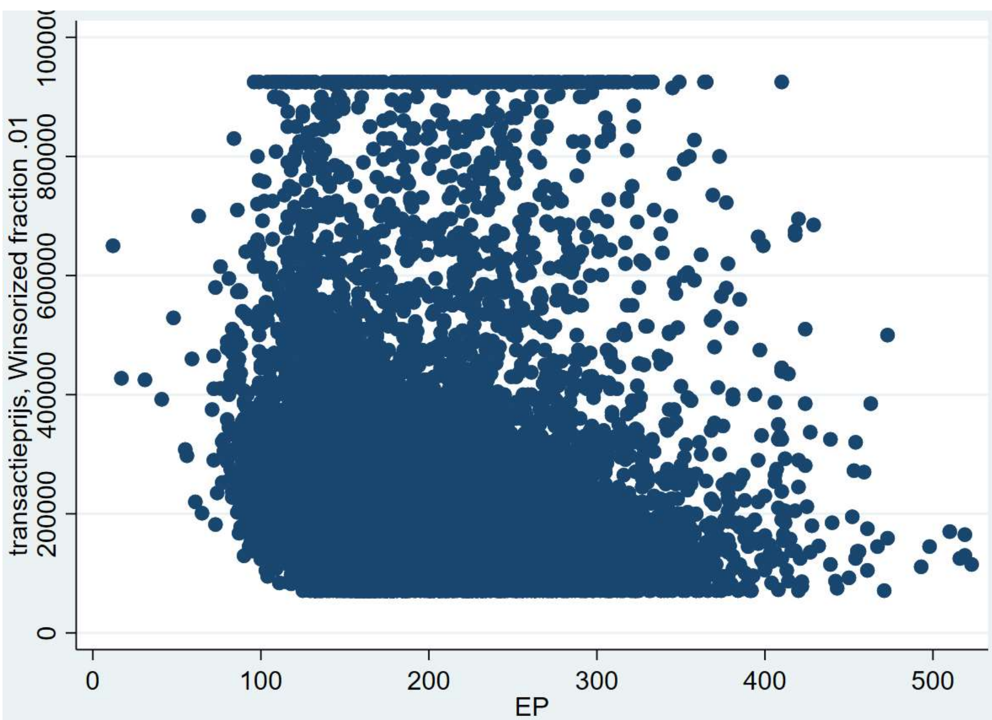

Figure E7: scatter plot of EP (x-axis) and transaction price (y-axis)

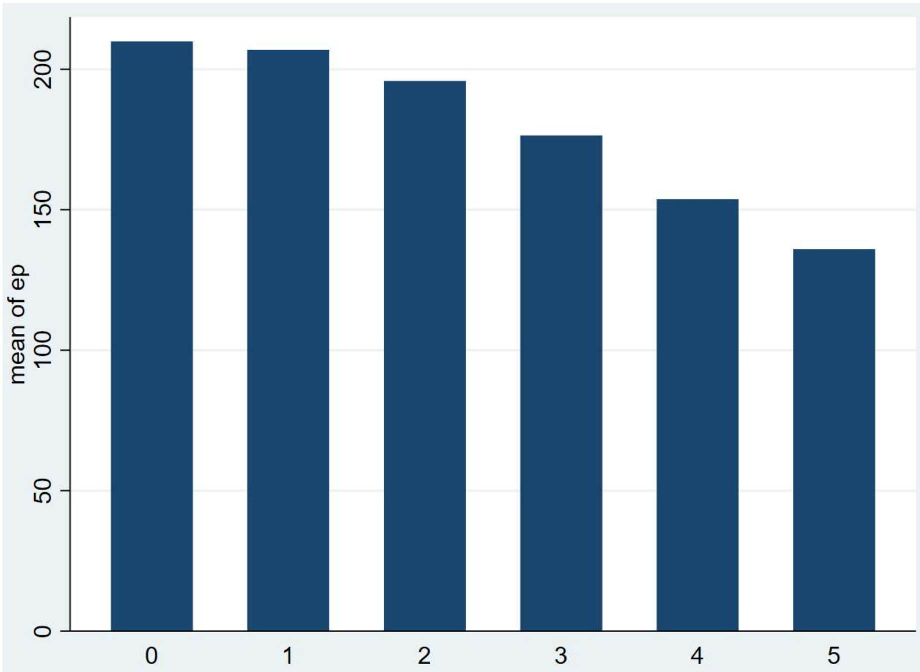

Figure E8: mean of the EPI score (y-axis) over the amount of isolation ways (x-axis).

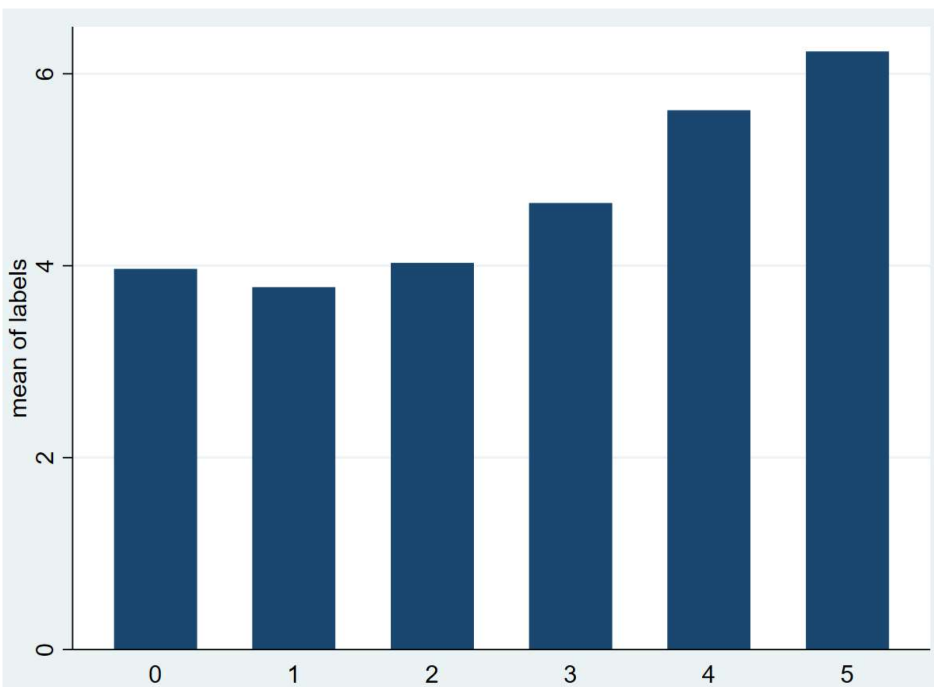

Figure E9: mean of the labels (y-axis) over the isolation ways (x-axis). Wherein Label A =7, label G = 1

Appendix F – Full regressions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Inverse Log(Energy Performance)	0.294***	0.0460***				
	(0.0123)	(0.00614)				
Label A			0.476***	0.0315***		
			(0.00531)	(0.00329)		
Label B			0.309***	0.0282***		
			(0.00557)	(0.00251)		
Label C			0.121***	0.0187***		
			(0.00476)	(0.00178)		
Label E			-0.0277***	-0.00626***		
			(0.00619)	(0.00198)		
Label F			0.0585***	-0.0108***		
			(0.00749)	(0.00238)		
Label G			0.0829***	-0.0203***		
			(0.00982)	(0.00329)		
Green label					0.0243***	
					(0.00162)	
Bad label						-0.0177***
						(0.00160)
Log size in m3		0.243***		0.227***	0.228***	0.228***
		(0.0171)		(0.00719)	(0.00717)	(0.00717)
Log living area in m2		0.527***		0.539***	0.538***	0.539***
		(0.0183)		(0.00783)	(0.00782)	(0.00782)
Unknown building year		-0.00115		0.01000	0.0127	0.0196
		(0.0618)		(0.0496)	(0.0501)	(0.0501)
before 1905		0.0569***		0.0364***	0.0370***	0.0373***
		(0.0113)		(0.00457)	(0.00457)	(0.00458)
1931-1944		0.0109		0.0154***	0.0153***	0.0150***
		(0.00833)		(0.00312)	(0.00312)	(0.00312)
1945-1959		-0.0385***		-0.0305***	-0.0275***	-0.0280***
		(0.00707)		(0.00320)	(0.00315)	(0.00316)
1960-1970		-0.0774***		-0.0768***	-0.0731***	-0.0729***
		(0.00707)		(0.00315)	(0.00310)	(0.00311)
1971-1980		-0.0706***		-0.0724***	-0.0681***	-0.0655***
		(0.00728)		(0.00326)	(0.00320)	(0.00319)
1981-1990		-0.0418***		-0.0271***	-0.0228***	-0.0166***
		(0.00741)		(0.00340)	(0.00333)	(0.00327)
1991-2000		0.0507***		0.0425***	0.0518***	0.0600***
		(0.00837)		(0.00382)	(0.00358)	(0.00348)
After 2001		0.111***		0.0802***	0.0932***	0.101***
		(0.00893)		(0.00432)	(0.00370)	(0.00361)
Amount of floors		-0.0243***		-0.0268***	-0.0264***	-0.0266***
		(0.00360)		(0.00159)	(0.00159)	(0.00159)
Apartment		-0.0545**		-0.0469***	-0.0450***	-0.0467***
		(0.0226)		(0.0104)	(0.0104)	(0.0103)
Between house		0.0596***		0.0500***	0.0497***	0.0491***
		(0.0194)		(0.00476)	(0.00476)	(0.00475)
Corner house		0.0345***		0.0334***	0.0327***	0.0318***
		(0.00328)		(0.00161)	(0.00160)	(0.00160)

Semi-detached	0.105***	0.118***	0.117***	0.116***
	(0.00999)	(0.00279)	(0.00279)	(0.00279)
Detached house	0.176***	0.232***	0.231***	0.229***
	(0.0195)	(0.00557)	(0.00558)	(0.00557)
Vacation home	-0.461***	-0.316***	-0.315***	-0.318***
	(0.142)	(0.0690)	(0.0687)	(0.0686)
Single family home	0.0123**	0.0130***	0.0136***	0.0142***
	(0.00489)	(0.00415)	(0.00415)	(0.00414)
Canal house	0.168***	0.132***	0.132***	0.132***
	(0.0371)	(0.0192)	(0.0192)	(0.0192)
Mansion	0.154***	0.0959***	0.0962***	0.0971***
	(0.0109)	(0.00513)	(0.00513)	(0.00513)
Farmhouse	0.249***	0.122***	0.122***	0.122***
	(0.0627)	(0.0243)	(0.0242)	(0.0243)
Bungalow	0.146***	0.148***	0.150***	0.150***
	(0.0265)	(0.00860)	(0.00861)	(0.00860)
Villa	0.190***	0.103***	0.104***	0.105***
	(0.0212)	(0.00721)	(0.00721)	(0.00721)
Large house	0.250**	0.130***	0.131***	0.132***
	(0.126)	(0.0216)	(0.0215)	(0.0215)
Ground floor apartement	-0.0652***	-0.0551***	-0.0547***	-0.0544***
	(0.0218)	(0.00889)	(0.00892)	(0.00885)
Upstairs apartement	-0.115***	-0.143***	-0.145***	-0.143***
	(0.0217)	(0.00877)	(0.00879)	(0.00872)
Maisonnette	-0.126***	-0.150***	-0.151***	-0.149***
	(0.0218)	(0.00884)	(0.00887)	(0.00880)
Staircase entrance flat	-0.118***	-0.135***	-0.137***	-0.136***
	(0.0217)	(0.00884)	(0.00886)	(0.00879)
Gallery flat	-0.126***	-0.147***	-0.150***	-0.148***
	(0.0218)	(0.00895)	(0.00897)	(0.00891)
Service flat	0.124***	0.195***	0.196***	0.200***
	(0.0372)	(0.0205)	(0.0201)	(0.0195)
Normal apartment	0.0173***	0.0464***	0.0474***	0.0464***
	(0.00643)	(0.00480)	(0.00480)	(0.00479)
Luxurious apartment	0.122***	0.141***	0.142***	0.142***
	(0.00877)	(0.00545)	(0.00545)	(0.00544)
Ratio (m2/m3)	-0.0166***	-0.0110***	-0.0112***	-0.0111***
	(0.00585)	(0.00249)	(0.00248)	(0.00248)
Number of rooms	0.000344	0.00503***	0.00484***	0.00481***
	(0.00190)	(0.000861)	(0.000860)	(0.000860)
Number of toilets	0.00366***	0.00447***	0.00451***	0.00454***
	(0.000884)	(0.000428)	(0.000428)	(0.000428)
Number of bathrooms	0.00686**	0.00771***	0.00769***	0.00761***
	(0.00280)	(0.00127)	(0.00127)	(0.00127)
Attick	0.0154***	0.0129***	0.0126***	0.0125***
	(0.00385)	(0.00176)	(0.00176)	(0.00177)
Gas heating	-0.0596***	-0.0719***	-0.0802***	-0.0765***
	(0.0125)	(0.00701)	(0.00679)	(0.00680)
One type of isolation	-0.00349	-0.00417**	-0.00378**	-0.00425**
	(0.00339)	(0.00181)	(0.00181)	(0.00181)
Two types of isolation	0.0184***	0.0163***	0.0173***	0.0171***
	(0.00524)	(0.00239)	(0.00238)	(0.00239)

Three types of isolation	0.0194***	0.0181***	0.0195***	0.0202***
	(0.00627)	(0.00266)	(0.00266)	(0.00266)
Four types of isolation	0.0130**	0.0111***	0.0127***	0.0136***
	(0.00642)	(0.00253)	(0.00253)	(0.00253)
Five types of isolation	0.000208	0.0144***	0.0159***	0.0166***
	(0.00545)	(0.00228)	(0.00227)	(0.00227)
Number of domers	0.0151***	0.0155***	0.0152***	0.0155***
	(0.00433)	(0.00137)	(0.00137)	(0.00137)
Number of pantries	0.0205***	0.0135***	0.0132***	0.0133***
	(0.00674)	(0.00241)	(0.00241)	(0.00241)
Parking place	0.0515***	0.0439***	0.0438***	0.0442***
	(0.00554)	(0.00203)	(0.00203)	(0.00203)
Carport	0.0892***	0.0785***	0.0785***	0.0790***
	(0.00716)	(0.00299)	(0.00299)	(0.00299)
Garage	0.106***	0.0868***	0.0869***	0.0875***
	(0.00689)	(0.00221)	(0.00221)	(0.00221)
Garage and Carport	0.138***	0.104***	0.105***	0.105***
	(0.0238)	(0.00875)	(0.00875)	(0.00877)
Garage for multiple cars	0.151***	0.110***	0.110***	0.110***
	(0.0203)	(0.00553)	(0.00553)	(0.00553)
Near forest	0.0904***	0.0770***	0.0765***	0.0774***
	(0.0182)	(0.00747)	(0.00745)	(0.00747)
Near water	0.0826***	0.0658***	0.0657***	0.0659***
	(0.00518)	(0.00208)	(0.00208)	(0.00208)
Near park	0.0405***	0.0313***	0.0311***	0.0313***
	(0.00573)	(0.00267)	(0.00267)	(0.00267)
Clear view	0.0279***	0.0224***	0.0221***	0.0223***
	(0.00303)	(0.00140)	(0.00140)	(0.00140)
Near calm road	0.00220	0.00191*	0.00192*	0.00191*
	(0.00253)	(0.00112)	(0.00112)	(0.00112)
Near busy road	-0.0220***	-0.0366***	-0.0367***	-0.0367***
	(0.00667)	(0.00364)	(0.00364)	(0.00363)
outside built-up area	0.0634	0.0678***	0.0678***	0.0674***
	(0.0532)	(0.0118)	(0.0118)	(0.0118)
residential area	-0.0181***	-0.0187***	-0.0187***	-0.0187***
	(0.00294)	(0.00146)	(0.00146)	(0.00146)
In centrum	0.0231***	0.0224***	0.0219***	0.0220***
	(0.00693)	(0.00308)	(0.00308)	(0.00308)
Maintenance outside: 2	0.0166	0.0377	0.0388	0.0382
	(0.0890)	(0.0676)	(0.0675)	(0.0673)
Maintenance outside: 3	0.166***	0.0853*	0.0881*	0.0851*
	(0.0641)	(0.0475)	(0.0475)	(0.0475)
Maintenance outside: 4	0.225***	0.136***	0.138***	0.135***
	(0.0682)	(0.0493)	(0.0494)	(0.0493)
Maintenance outside: 5	0.226***	0.141***	0.144***	0.142***
	(0.0627)	(0.0459)	(0.0460)	(0.0459)
Maintenance outside: 6	0.235***	0.167***	0.170***	0.168***
	(0.0629)	(0.0460)	(0.0460)	(0.0460)
Maintenance outside: 7	0.241***	0.175***	0.179***	0.176***
	(0.0628)	(0.0459)	(0.0459)	(0.0459)
Maintenance outside: 8	0.252***	0.179***	0.183***	0.182***
	(0.0634)	(0.0460)	(0.0461)	(0.0460)

Maintenance outside: 9		0.241***		0.173***	0.178***	0.175***
		(0.0632)		(0.0460)	(0.0460)	(0.0460)
Maintenance inside: 2		-0.0662		-0.0310	-0.0300	-0.0288
		(0.0510)		(0.0410)	(0.0408)	(0.0408)
Maintenance inside: 3		-0.000546		-0.0174	-0.0157	-0.0161
		(0.0211)		(0.0205)	(0.0206)	(0.0205)
Maintenance inside: 4		0.00362		0.00361	0.00557	0.00528
		(0.0230)		(0.0218)	(0.0219)	(0.0218)
Maintenance inside: 5		0.0348*		0.0166	0.0194	0.0184
		(0.0205)		(0.0198)	(0.0199)	(0.0198)
Maintenance inside: 6		0.0420**		0.0323	0.0352*	0.0340*
		(0.0208)		(0.0200)	(0.0201)	(0.0200)
Maintenance inside: 7		0.0958***		0.104***	0.108***	0.107***
		(0.0206)		(0.0199)	(0.0199)	(0.0199)
Maintenance inside: 8		0.170***		0.142***	0.145***	0.145***
		(0.0220)		(0.0201)	(0.0201)	(0.0201)
Maintenance inside: 9		0.203***		0.168***	0.171***	0.171***
		(0.0214)		(0.0200)	(0.0201)	(0.0200)
Year dummies	No	Yes	No	Yes	Yes	Yes
Postal code dummies	No	Yes	No	Yes	Yes	Yes
Constant	13.61***	8.225***	12.15***	8.180***	8.147***	8.166***
	(0.0639)	(0.153)	(0.00387)	(0.0578)	(0.0576)	(0.0578)
Observations	21,273	21,049	85,475	84,554	84,554	84,554
R-squared	0.030	0.909	0.113	0.911	0.910	0.910

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

Table F1: shows the full regressions as portrayed in Table 5.1(dependent variable: log transaction price). Reference categories of the regression are:

- *Energy labels: label D*
- *Building year: 1906-1930*
- *Isolation: zero ways of isolation*
- *Maintenance outside and inside: grade 1*
- *Type of house: town house*
- *Type of dwelling: simple house*
- *Quality of apartment: simple apartment*
- *Type of heating: no heating*
- *Location variables: not specified*
- *Parking variables: no parking space*

VARIABLES	(1) Full sample first inv_logep	(2) Full sample second logtransactie	(3) 1945-2000 first inv_logep	(4) 1945-2000 second logtransactie
After oil crisis	0.233***		0.231***	
	(0.00705)		(0.00695)	
Inverse Log(Energy Performance)		0.0400*		0.0699***
		(0.0229)		(0.0231)
before 1905	0.0802***	0.0907***	-	-
	(0.0114)	(0.00820)		
1905-1930	0.0493***	0.0357***	-	-
	(0.00771)	(0.00547)		
1931-1944	0.0106	0.0494***	-	-
	(0.00866)	(0.00647)		
1960-1970	0.0159**	-0.0379***	0.00437	-0.0309***
	(0.00627)	(0.00455)	(0.00604)	(0.00459)
1971-1980	0.0874***	-0.0280***	0.0812***	-0.0213***
	(0.00690)	(0.00422)	(0.00674)	(0.00416)
1991-2000	0.116***	0.0935***	0.125***	0.0969***
	(0.00760)	(0.00707)	(0.00754)	(0.00711)
After 2001	0.249***	0.154***	-	-
	(0.00841)	(0.00987)		
Log size in m3	0.00170	0.242***	-0.0456**	0.249***
	(0.0186)	(0.0141)	(0.0218)	(0.0169)
Log living area in m2	-0.0312	0.522***	-0.0142	0.470***
	(0.0195)	(0.0147)	(0.0225)	(0.0173)
Amount of floors	0.00650	-0.0242***	-0.00401	-0.0202***
	(0.00405)	(0.00307)	(0.00464)	(0.00357)
Apartment	0.0247	-0.0403**	0.00261	-0.0917***
	(0.0237)	(0.0180)	(0.0359)	(0.0277)
Between house	-0.0285	0.0622***	-0.0481**	0.0919***
	(0.0183)	(0.0139)	(0.0188)	(0.0146)
Corner house	-0.0238***	0.0345***	-0.0221***	0.0369***
	(0.00484)	(0.00371)	(0.00479)	(0.00373)
Semi-detached	-0.0536***	0.105***	-0.0482***	0.138***
	(0.0108)	(0.00829)	(0.0131)	(0.0102)
Detached house	-0.0915***	0.191***	-0.0806***	0.251***
	(0.0164)	(0.0126)	(0.0205)	(0.0159)
Vacation home	-0.00682	0.226***	-0.112	0.444***
	(0.0911)	(0.0689)	(0.0929)	(0.0715)
Single family home	0.0247***	0.0124**	0.0167**	0.0120**
	(0.00697)	(0.00532)	(0.00695)	(0.00537)
Canal house	-0.00642	0.187***	-	-
	(0.0487)	(0.0368)		
Mansion	0.0229*	0.156***	0.0296*	0.140***
	(0.0119)	(0.00900)	(0.0163)	(0.0126)
Farmhouse	0.0440	0.232***	-0.192**	0.400***
	(0.0492)	(0.0373)	(0.0924)	(0.0713)
Bungalow	0.129***	0.154***	0.110***	0.130***
	(0.0245)	(0.0189)	(0.0242)	(0.0188)

Villa	0.0434**	0.182***	0.0767***	0.176***
	(0.0204)	(0.0155)	(0.0245)	(0.0190)
Large house	0.157**	0.311***	0.286***	0.117
	(0.0790)	(0.0599)	(0.110)	(0.0853)
Ground floor apartement	-0.0326	-0.0756***	-0.0656*	-0.0321
	(0.0226)	(0.0171)	(0.0355)	(0.0274)
Upstairs apartement	-0.0449**	-0.128***	-0.0740**	-0.0729***
	(0.0225)	(0.0170)	(0.0354)	(0.0273)
Maisonnette	-0.0497**	-0.139***	-0.0478	-0.0840***
	(0.0228)	(0.0173)	(0.0352)	(0.0271)
Staircase entrance flat	-0.0551**	-0.130***	-0.0700**	-0.0820***
	(0.0225)	(0.0170)	(0.0351)	(0.0270)
Gallery flat	-0.0744***	-0.138***	-0.0909***	-0.0889***
	(0.0227)	(0.0172)	(0.0352)	(0.0272)
Service flat	0.0583	0.120	0.157	0.170
	(0.139)	(0.106)	(0.179)	(0.138)
Normal apartment	0.0255***	0.0133**	0.0222***	-0.00250
	(0.00691)	(0.00526)	(0.00691)	(0.00534)
Luxurious apartment	0.0496***	0.118***	0.0141	0.0896***
	(0.00977)	(0.00745)	(0.0114)	(0.00881)
Ratio (m2/m3)	-0.0170**	-0.0181***	-0.0107	-0.0149**
	(0.00663)	(0.00503)	(0.00817)	(0.00629)
Number of rooms	-0.00573***	0.00129	-0.00252	-0.00145
	(0.00205)	(0.00157)	(0.00234)	(0.00181)
Number of toilets	0.00346***	0.00387***	0.000855	0.00385***
	(0.00111)	(0.000849)	(0.00123)	(0.000944)
Number of bathrooms	-0.00465	0.00631**	-0.00266	0.00590**
	(0.00338)	(0.00256)	(0.00370)	(0.00285)
Attick	-0.0132**	0.0149***	-0.0132**	0.00980**
	(0.00542)	(0.00412)	(0.00546)	(0.00422)
Gas heating	-0.178***	-0.0536***	-0.144***	-0.0239
	(0.0148)	(0.0120)	(0.0190)	(0.0152)
CV heating	0.0362***	0.0103**	0.0121*	0.00673
	(0.00595)	(0.00462)	(0.00653)	(0.00504)
Solar panels	0.664***	-0.00272	0.471***	0.0877
	(0.0602)	(0.0479)	(0.102)	(0.0793)
One type of isolation	0.0259***	-0.00404	0.0278***	0.00229
	(0.00408)	(0.00314)	(0.00423)	(0.00331)
Two types of isolation	0.0531***	0.0174***	0.0518***	0.00965*
	(0.00662)	(0.00514)	(0.00694)	(0.00547)
Three types of isolation	0.0967***	0.0181***	0.0923***	0.0161**
	(0.00789)	(0.00636)	(0.00815)	(0.00662)
Four types of isolation	0.0966***	0.0113	0.0778***	0.00209
	(0.00856)	(0.00693)	(0.00930)	(0.00747)
Five types of isolation	0.101***	-0.000508	0.0898***	-0.000943
	(0.00651)	(0.00557)	(0.00739)	(0.00620)
Number of domers	-0.0150***	0.0144***	-0.00310	0.0157***
	(0.00574)	(0.00437)	(0.00631)	(0.00486)
Number of pantries	-0.0181**	0.0208***	0.000414	0.0330***
	(0.00726)	(0.00551)	(0.00821)	(0.00632)
Parking place	0.00512	0.0507***	0.0194***	0.0366***
	(0.00638)	(0.00483)	(0.00742)	(0.00574)

Carport	0.0264***	0.0871***	0.0457***	0.0786***
	(0.00847)	(0.00647)	(0.0100)	(0.00786)
Garage	0.0167**	0.104***	0.0284***	0.104***
	(0.00753)	(0.00571)	(0.00809)	(0.00627)
Garage and Carport	0.0709**	0.128***	0.106***	0.138***
	(0.0294)	(0.0223)	(0.0353)	(0.0273)
Garage for multiple cars	-0.0102	0.143***	0.00503	0.127***
	(0.0188)	(0.0142)	(0.0225)	(0.0173)
Near forest	-0.0276*	0.0938***	0.0174	0.0985***
	(0.0164)	(0.0124)	(0.0175)	(0.0135)
Near water	-0.00328	0.0808***	0.00654	0.0831***
	(0.00580)	(0.00439)	(0.00622)	(0.00479)
Near park	-0.0205***	0.0438***	-0.0146**	0.0439***
	(0.00746)	(0.00567)	(0.00738)	(0.00569)
Clear view	-0.0246***	0.0269***	-0.0197***	0.0305***
	(0.00391)	(0.00302)	(0.00403)	(0.00314)
Near calm road	-0.000971	0.00226	-0.00268	0.00786***
	(0.00330)	(0.00249)	(0.00348)	(0.00268)
Near busy road	-0.00210	-0.0249***	0.00183	-0.0256***
	(0.00767)	(0.00580)	(0.00862)	(0.00663)
outside built-up area	-0.0651**	0.0926***	0.0220	0.127***
	(0.0313)	(0.0237)	(0.0416)	(0.0320)
residential area	-0.0138***	-0.0193***	-0.0166***	-0.0220***
	(0.00370)	(0.00281)	(0.00392)	(0.00303)
In centrum	-0.0373***	0.0178***	-0.0291***	0.00962
	(0.00744)	(0.00572)	(0.00847)	(0.00657)
Maintenance outside: 2	0.0230	0.0264	-0.210	-0.0240
	(0.0774)	(0.0586)	(0.147)	(0.113)
Maintenance outside: 3	0.146***	0.178***	0.0548	0.00356
	(0.0486)	(0.0369)	(0.107)	(0.0823)
Maintenance outside: 4	0.163***	0.232***	0.0624	0.0762
	(0.0526)	(0.0400)	(0.109)	(0.0840)
Maintenance outside: 5	0.191***	0.233***	0.0699	0.0631
	(0.0463)	(0.0352)	(0.106)	(0.0812)
Maintenance outside: 6	0.170***	0.242***	0.0480	0.0685
	(0.0466)	(0.0355)	(0.106)	(0.0813)
Maintenance outside: 7	0.206***	0.245***	0.0687	0.0711
	(0.0463)	(0.0353)	(0.106)	(0.0812)
Maintenance outside: 8	0.315***	0.260***	0.127	0.0597
	(0.0475)	(0.0366)	(0.106)	(0.0819)
Maintenance outside: 9	0.252***	0.244***	0.132	0.0474
	(0.0470)	(0.0360)	(0.106)	(0.0817)
Maintenance inside: 2	0.0310	-0.0630*	0.0402	-0.0464
	(0.0479)	(0.0362)	(0.0653)	(0.0503)
Maintenance inside: 3	0.0883***	-0.00452	0.0193	-0.00641
	(0.0248)	(0.0189)	(0.0291)	(0.0224)
Maintenance inside: 4	0.103***	0.00248	0.0584*	0.00903
	(0.0274)	(0.0209)	(0.0313)	(0.0242)
Maintenance inside: 5	0.133***	0.0345*	0.0568**	0.0255
	(0.0242)	(0.0186)	(0.0284)	(0.0219)
Maintenance inside: 6	0.140***	0.0417**	0.0600**	0.0332
	(0.0247)	(0.0190)	(0.0288)	(0.0222)

Maintenance inside: 7	0.151***	0.0961***	0.0732**	0.0815***
	(0.0243)	(0.0188)	(0.0285)	(0.0220)
Maintenance inside: 8	0.141***	0.168***	0.0645**	0.168***
	(0.0263)	(0.0202)	(0.0311)	(0.0240)
Maintenance inside: 9	0.221***	0.205***	0.0710**	0.204***
	(0.0253)	(0.0199)	(0.0298)	(0.0230)
Month year dummies	Yes	Yes	Yes	Yes
Postal code dummies	Yes	Yes	Yes	Yes
Constant	-5.446***	8.162***	-5.079***	8.635***
	(0.243)	(0.223)	(0.242)	(0.222)
Observations	20,757	20,757	15,885	15,885
R-squared	0.589	0.911	0.542	0.896
IV F-stat		1095		1110
Durbin pval		0.770		0.160

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

Table F2: shows the full regressions as portrayed in Table 5.5 (dependent variable: log transaction price)

Reference categories of the regression are:

- *Energy labels: label D*
- *Building year: 1906-1930*
- *Isolation: zero ways of isolation*
- *Maintenance outside: 1*
- *Maintenance inside: 1*
- *Type of house: town house*
- *Type of dwelling: simple house*
- *Quality of apartment: simple apartment*
- *Type of heating: no heating*
- *Location variables: not specified*
- *Parking variables: no parking space*

Appendix G – Robustness checks

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Inverse Log(Energy Performance)	0.294***	0.0410***				
	(0.0114)	(0.00627)				
Label A			0.476***	0.0263***		
			(0.00569)	(0.00338)		
Label B			0.309***	0.0241***		
			(0.00580)	(0.00259)		
Label C			0.121***	0.0131***		
			(0.00491)	(0.00183)		
Label E			-0.0276***	-0.00753***		
			(0.00584)	(0.00201)		
Label F			0.0585***	-0.00941***		
			(0.00661)	(0.00244)		
Label G			0.0829***	-0.00871**		
			(0.00768)	(0.00339)		
Green label					0.0185***	
					(0.00166)	
Bad label						-0.0141***
						(0.00163)
Control variables:						
Structural characteristics	No	Yes	No	Yes	Yes	Yes
Internal features	No	Yes	No	Yes	Yes	Yes
External features	No	Yes	No	Yes	Yes	Yes
Environmental - Natural	No	Yes	No	Yes	Yes	Yes
Environmental - Neighborhood and Location	No	Yes	No	Yes	Yes	Yes
Marketing, occupancy and selling factors	No	Yes	No	Yes	Yes	Yes
Constant	13.61***	8.440***	12.15***	8.439***	8.425***	8.438***
	(0.0597)	(0.176)	(0.00379)	(0.0617)	(0.0615)	(0.0617)
Observations	21,275	21,050	85,481	84,558	84,558	84,558
R-squared	0.030	0.906	0.113	0.905	0.905	0.905

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

Table G1: Regression from Table 6.1 with the inclusion of the alternative abbreviations. (Dependent variable: log transaction price)

	(1)	(2)	(3)	(4)
	Total sample	Total sample	1945-2000	1945-2000
	First	Second	First	Second
VARIABLES	inv_logep	logtransactie	inv_logep	logtransactie
After oil crisis	0.238***		0.234***	
	(0.00705)		(0.00693)	
Inverted Log(Energy performance)		0.0268		0.0554**
		(0.0230)		(0.0235)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-5.606***	8.319***	-5.150***	8.757***
	(0.247)	(0.232)	(0.246)	(0.232)
Observations	21,050	21,050	15,971	15,971
R-squared	0.587	0.906	0.543	0.889
IV F-stat		1136		1145
Durbin pval		0.544		0.370

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

Table G2: Regression from Table 6.5 with the inclusion of the alternative abbreviations. (Dependent variable: log transaction price)

VARIABLES	(1)	(2)	(3)	(4)
Inverse Log(Energy Performance)	0.0161***			
	(0.00496)			
Label A		0.0211***		
		(0.00258)		
Label B		0.0171***		
		(0.00193)		
Label C		0.0161***		
		(0.00134)		
Label E		0.00260*		
		(0.00140)		
Label F		-0.000646		
		(0.00169)		
Label G		-0.00761***		
		(0.00218)		
Green label			0.0164***	
			(0.00121)	
Bad label				-0.00690***
				(0.00113)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	8.240***	8.111***	8.106***	8.110***
	(0.0941)	(0.0339)	(0.0338)	(0.0339)
Observations	33,946	170,730	170,730	170,730
R-squared	0.908	0.902	0.902	0.901

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

Table G3: Regression from Table 6.1 without dropping the non-active labels throughout the transaction. (Dependent variable: log transaction price)

VARIABLES	(1) Original model	(2)	(3)	(5)	(6)
Inverse Log(Energy Performance)	0.0475***	0.0475***	0.0482***	0.0313***	-0.0108
	(0.00604)	(0.00605)	(0.00609)	(0.00635)	(0.00730)
Control variables:					
Structural characteristics	Yes	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Locatior	Yes	Yes	Yes	Yes	Yes
- Postal code (4 digits) dummies	Yes	-	-	Yes	Yes
- Postal code (3 digits) dummies	-	Yes	-	-	-
- Postal code (2 digits) dummies	-	-	Yes	-	-
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes	Yes
- Year and month dummies	Yes	Yes	Yes	No	No
- Year and quarter dummies	-	-	-	Yes	-
- Year dummies	-	-	-	-	Yes
Constant	8.317***	8.536***	8.484***	8.196***	7.811***
	(0.161)	(0.150)	(0.152)	(0.116)	(0.115)
Observations	20,755	20,755	20,755	20,755	20,755
R-squared	0.910	0.910	0.907	0.880	0.838

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

Table G4: Regression with the breakdown of the Postal code and Time series dummies. (Dependent variable: log transaction price)

	2009-2014			
	Full sample		1945 - 2000	
	(1)	(2)	(3)	(4)
	first	second	first	second
VARIABLES	inv_logep	logtransactie	inv_logep	logtransactie
After oil cris	0.235***		0.227***	
	(0.00909)		(0.00912)	
Inverse Log(Energy Performance)		0.0554*		0.0892***
		(0.0283)		(0.0297)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-4.970***	8.186***	-4.606***	8.704***
	(0.209)	(0.209)	(0.220)	(0.216)
Observations	13,543	13,543	9,713	9,713
R-squared	0.603	0.924	0.538	0.911
IV F-stat		670.7		620.9
Durbin pval		0.861		0.0724

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

Table G5: Regression from Table 6.5 with only the transactions from year 2009-2014. (Dependent variable: log transaction price)

	2015-2018			
	Full sample		1945 -2000	
	(1)	(2)	(3)	(4)
	first	second	first	second
VARIABLES	inv_logep	logtransactie	inv_logep	logtransactie
After oil cris	0.235***		0.253***	
	(0.0112)		(0.0111)	
Inverse Log(Energy Performance)		0.00416		0.0171
		(0.0355)		(0.0335)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-5.692***	8.453***	-5.103***	8.914***
	(0.275)	(0.288)	(0.245)	(0.254)
Observations	7,210	7,210	6,171	6,171
R-squared	0.641	0.894	0.634	0.891
IV F-stat		438.3		517.1
Durbin pval		0.565		0.719

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

Table G6: Regression from Table 6.5 with only the transactions from year 2015-2018. (Dependent variable: log transaction price)

VARIABLES	(1) Full sample first inv_logep	(2) Full sample second logtransactie	(3) 1945-2000 first inv_logep	(4) 1945-200 second logtransactie
After oil crisis	0.583***		0.568***	
	(0.00605)		(0.00723)	
Green labels		0.0333***		0.0339***
		(0.00473)		(0.00484)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-0.225	8.119***	-0.398	8.386***
	(0.331)	(0.151)	(0.374)	(0.142)
Observations	83,850	83,850	53,536	53,536
R-squared	0.586	0.911	0.500	0.914
IV F-stat		9277		6166
Durbin pval		0.0583		0.0454

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

Table G7: IV regression from Table 6.5 with green label as independent variable instead of inverse log (EPI). (Dependent variable: log transaction price)

VARIABLES	(1) Full sample first inv_logep	(2) Full sample second logtransactie	(3) 1945-2000 first inv_logep	(4) 1945-200 second logtransactie
After oil crisis (from 1970 onwards)	0.233***		0.232***	
	(0.00705)		(0.00695)	
Inverse Log(Energy Performance)		0.0399*		0.0698***
		(0.0229)		(0.0231)
Control variables:				
Structural characteristics	Yes	Yes	Yes	Yes
Internal features	Yes	Yes	Yes	Yes
External features	Yes	Yes	Yes	Yes
Environmental - Natural	Yes	Yes	Yes	Yes
Enivornmental - Neighborhood and Location	Yes	Yes	Yes	Yes
Marketing, occupancy and selling factors	Yes	Yes	Yes	Yes
Constant	-5.448***	8.165***	-5.074***	8.637***
	(0.243)	(0.223)	(0.242)	(0.222)
Observations	20,756	20,756	15,885	15,885
R-squared	0.589	0.911	0.542	0.896
IV F-stat		1095		1110
Durbin pval		0.764		0.161

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

Table G8: IV regression from Table 6.5 with the after oil crisis containing observations from 1970 onwards instead of 1980. (Dependent variable: log transaction price)