Erasmus University of Rotterdam

Erasmus School of Economics
MSc International Economics

Master Thesis

The Value of Energy Performance in the Dutch Housing Market

Author:
J. Havlínová

Supervisors:
Prof. Dr. B. Jacobs
Mw. L. Berkvens
Dhr. D. W. van Dijk

Co-reader:
Dr. A. A. F. Gerritsen

July 23, 2018

Abstract

To prevent global warming from rising above two degrees celsius, the European Union introduced Energy Performance Certificates (EPCs). The EPCs were introduced in the Netherlands in 2008 with the objective of creating awareness and solving the information asymmetry on energy performance in the housing market. Hence, as house buyers will be better informed about the future savings of energy costs, *'green'* dwellings should transact for a premium. This study investigates the premium paid for green dwellings. It thoroughly examines the mechanism at stake, which is followed by an empirical estimation of the premium. I find that people are willing to pay a premium of around 3% on average for a green dwelling. Moreover, this study analyses how a change in policy influenced the premium. The findings show a significant change in premium after the policy changed.

Keywords: green, premium, housing market, EPBD

Acknowledgements

Writing this thesis would not have been possible without the support and help of several people. Foremost, I want to thank Bas Jacobs for the valuable supervision of my thesis. I greatly enjoyed writing my thesis under your guidance. I want to thank you for your patience and time. Specifically, I truly appreciate your critical remarks regarding the theoretical part of my thesis which brought my thesis to a more ambitious level.

I am also very grateful to my two other supervisors from *De Nederlandsche Bank*. Firstly, Laurien Berkvens, with who I met almost weekly to discuss the progress of my thesis and who shared her vast knowledge on the Dutch energy transition with me. Secondly, Dorinth van Dijk who helped me with obtaining the dataset and advising me on the empirical part of my thesis. Moreover, I would like to acknowledge several people within the bank for their contributions regarding my thesis. A big thank you to Jante Parlevliet and the rest of the AEB team for the great opportunity to write my thesis at your department. It was great to have so many inspiring people around me while writing my thesis, always willing to help when I had questions. A special thanks to Maurice Bun, who helped me several times with my questions concerning the empirical part of my thesis.

Additionally, I want to express my gratitude to Edwin Marquart (RVO) and Frank Harleman (NVM) for helping me with receiving all the data that was necessary to do this research.

Finally, I would like to thank my parents and my friends for supporting me throughout the process of writing my thesis, which has proven to be difficult but rewarding at the same time. A special thank you to Lucas for patiently helping me to match the two datasets, and Nikki for critically reviewing my thesis.

Contents

List of Tables

List of Figures

1 Introduction

To prevent global warming from rising above two degrees Celsius, the world is in the midst of an energy transition. Apart from finding more sustainable energy sources, there is consensus that the energy efficiency needs to be increased in order to mitigate climate change. Related to this aspect of energy efficiency, the European Commission pointed out that 40% of the energy consumption and 36% of the CO2 emissions in the European Union originates from buildings (European-Commission, 2018a). Therefore, the EU has set a target to cut the emissions from buildings with 90% (compared to 1990) by 2050. In order to achieve this target Energy Performance Certificates (EPCs) were introduced as part of the Energy Performance Building Directive. The EPCs, used for labeling buildings according to their energy performance, have the objective of creating awareness. Subsequently this would lead to a more correct evaluation of energy performance in the housing market, as people would be better informed. This would then stimulate people to invest more in energy-saving measures. According to the EU, however, the effectiveness of the energy labels depends highly on the way they are implemented by the member states (Mudgal, Lyons, Cohen, Lyons & Fedrigo-Fazio, 2013, p. 18).

Although the EPCs were already introduced in the Netherlands in 2008, it was not until 2015 that the government included the EPBD in their legislation, making it mandatory for house-owners to include an EPC when selling a house. Consequently, the Dutch government was officially criticised several times by the European Commission on their delayed implementation[1]. This discussion led to a policy change in 2015 (see BOX 1 for more information on the EPBD and the change in policy). One way of investigating whether the EPCs were implemented adequately and whether the change in policy improved their effectiveness, is by looking how the EPCs are valued in the Dutch housing market. Therefore, this paper investigates the effect of EPCs on the transaction prices of residential dwellings

[1]Tweede Kamer, 2015-2016, 30 196, nrs. 469 & 549

in the Netherlands. Not only will this give an indication of how energy performance is valued in the Dutch housing market, it also allows us to evaluate the effectiveness of the policy. To analyse the developments in the implementation of the EPCs in the Netherlands, this thesis estimates the *'premium'* and the change in premium due to the change in policy. The premium is the extra amount paid (in percentages) for an energy efficient, or *'green'* dwelling, which is in this paper defined as a dwelling with labels A, B or C. The magnitude of this premium can give an useful insight into the effectiveness of the EPCs. Moreover, the developments in the magnitude of this premium gives the opportunity to evaluate the Dutch policy on the implementation of the EPCs.

This research is done by a thorough theoretical examination of the mechanism that works from assessing the energy performance of a dwelling to the transaction price paid for the dwelling. This shows how EPCs are supposed to solve the information asymmetry when it comes to the energy performance in the housing market, how this creates a premium and how supply and demand influence this premium. A hedonic price model is then used to estimate the effect of energy labels (as a proxy for energy performance) on house transaction prices (as a proxy for the value) through OLS. This model is commonly used for analyzing the willingness-to-pay for houses and their underlying characteristics (Palmquist, 2005; O'Sullivan & Gibb, 2008). To empirically estimate the premium I use a comprehensive dataset of more than 300.000 observations over the period of 2008-2017. However, when estimating the premium, I run into issues of endogeneity due to omitted variables and self-selection into the sample. These limitations to the model are thoroughly discussed. Even though not all issues can be solved, the findings still give a useful insight into the causal effect of EPCs on housing values.

The findings of this thesis show a significant positive effect of the EPCs on the transaction prices of houses in the Netherlands. More precisely, I find that green dwellings transact for a premium of 3.1% compared to non-green dwellings. This implies that the EPCs are effective. Furthermore, I find that this premium was only 1.2% on average before 2015 and increased by another 2.5 percentage points after 2015. This indicates that the effectiveness increased substantially due to the policy change. This is supportive of the

European Commission's reasoning that the effectiveness of the EPCs is sensitive to their implementation.

BOX 1 The implementation of the EPCs in the Netherlands

The EPCs have been part of the EU's Energy Performance Building Directive (EPBD), which was already initiated in 2002, and had the objective of stimulating energy efficiency of buildings by creating awareness amongst house owners. The Directive mandated member states to set minimal requirements and inspections for energy performance. Moreover, it introduced the EPCs that label buildings from A (most efficient) to G (least efficient). The EPCs are based on energy-saving characteristics of the building such as the quality of insulation, ventilation, solar systems etc. Besides, an EPC also provides recommendations on how to improve the energy efficiency of the building. An example of an EPC can be found in Appendix A.

Member states were supposed to implement the EPCs by 2006, but were allowed to postpone the implementation for another three years. Since 2009 the inclusion of an EPC when renting out or selling a house has been made mandatory by the EU. However, the EPCs were not mandatory in the Netherlands before 2015. Brounen & Kok (2011) mention several aspects that contributed to the slow implementation of the EPCs in the Netherlands. Firstly, there was an option for house buyers to sign a waiver, which meant that an EPC did not have to be included in the sale or rental of a dwelling. Other factors were the lack of unskilled assessors and unclear requirements (Brounen & Kok, 2011, p. 4). The Dutch government was therefore criticised by the European Commission for not including the EPBD in their own legislation as well as wrongly executing it.

Moreover, before the introduction of the EPBD, the Netherlands already had their own legislation (since 1995) on the energy performance of buildings in the Netherlands:

the Energy Performance Standard (EPN). The EPN uses a different measure of energy performance of the housing market, namely the Energy Index (EI). The EI is an index comparing a buildings' energy performance to the energy performance of the average building in 1990, which makes it a much more specific proxy for a building's energy performance (van Ekerschot & Heinemans, 2008). There has been some confusion with the combination of the two legislations, as the energy labels and the Energy Index do not coincide in 8 percent of the cases[a].

All of this confusion and critique led to a new system that was implemented by the government in 2015 to increase its consumer friendliness (van Eck, 2015). As of January 1st all residential dwellings received a temporary EPC. House owners can change this into a final EPC, by providing information on the thermal characteristics of the dwelling including evidence via a web-tool. Afterwards an independent *'assessor'* uses this information to decide on the final EPC. Hence, requesting an EPC now happens entirely online and the assessor no longer needs to be physically present. The new policy also made the inclusion of an EPC during the sale of a house mandatory. When the EPC is not included in the sale or rental of the dwelling, the owner is fined with a maximum of 405euros (van Eck, 2015). However, a downside to the simplicity of the system is the difficulty in obtaining the Energy Index. This is due to the lack of detailed information compared to when the assessor is physically present.

[a] Tweede Kamer, 2017-2018, 30 196, nr. 563

This thesis relates to existing literature that estimate the premium for green buildings. Up to now, the only academic research done on Dutch level looking into EPCs has been the work of Brounen and Kok (2011), which dates back to the beginning of energy labels in the Netherlands. In recent years, however, mitigating climate change and promoting energy efficiency has gained a lot of importance. Examples of this are the Dutch Energy Accord that was reached in 2013 and the Global Paris agreement in 2015. Consequently, many targets have been set recently and policies changed accordingly (e.g. see Box 1). More importantly, the policy change of 2015 has never been academically investigated.

Besides giving a more recent estimation of the premium in the Netherlands, the extensive dataset over the relatively long observation period allows to give an useful insight in how the implementation of the EPCs influence their effectiveness. Moreover, it makes it possible to see the development of the premium throughout recent years where climate change has gained importance. Additionally, this thesis contributes to existing literature by giving a thorough discussion of the theoretical background and by explaining the mechanism. This theoretical analysis prior to the empirical analysis gives insight into the causal effect of EPCs on the transaction prices.

The remainder of this thesis is structured as follows: the next chapter reviews the existing literature on this topic. It analyses their findings and their different empirical methodologies used. Subsequently, Chapter 3 looks into the theory and discusses the role of imperfect information in energy performance of the housing market. It also reveals the mechanism behind the causal effect of EPCs on transaction prices and introduces the hedonic price model. This model is then used in Chapter 4 to obtain the empirical regression function used in this thesis to estimate the premium. Next, the extensive dataset is explained in Chapter 5 and is followed by a thorough discussion of the results of the model in Chapter 6. The results are checked for robustness in Chapter 7. Finally, Chapter 8 summarizes the findings of this thesis once more and concludes with several policy recommendations as well as suggestions for further research.

2 Literature Review

There is a variety of empirical researches focusing on energy-efficiency and the value of buildings. Since climate change is an issue that has recently gained a lot of attention, this chapter focuses on more recent researches that look for the effect of energy performance on real estate values. Table 2.1 summarizes the main effects found by academic researches. Almost all researchers seem to find significant results with the only exception being Oxford, most likely due to the small number of observations (Mudgal et al., 2013). The lowest premium is found in Ireland, where a one step increase of the label increases the transaction price of the dwelling with only 2.8% (Mudgal et al., 2013). In comparison, the highest premium is found in the United States where office buildings with a certification from the programs LEED or Energy Star are sold for a premium of almost 17% (Eichholtz, Kok & Quigley, 2010). Clearly, there are large differences between the premia, but the researchers also differ in many other aspects.

Table 2.1 gives a clear overview of the factors that play a role in estimating premia. Even though the premia give an indication of the amount people are willing to pay for energy performance, they are not easily comparable to each other due to these differences. Therefore, it is important to review all of these differences in regions, programs that measure energy performance, observation periods, type of buildings, number of observations and comparing methods.

To begin with the differences between researches in regions, the European Commission argued the following:

> *"The implementation and effectiveness of the EPCs vary from country to country and region to region, ... due to political and legal context as well as the characteristics of their property markets"* (Mudgal et al., 2013, p. 17).

The extensive research done by Mudgal et al. (2013) for the European Commission includes several countries (France, Belgium, UK, Ireland and Austria) that all report very different results. In France the results are stronger for Marseille than for Lille, which is contradictory to their assumption that energy performance is more valued in areas where more heating is needed. In the UK no significant effect of EPCs are found. Whereas in Austria, Belgium and Ireland the premia are significant and positive. Conversely, the results of Austria show that the effect is bigger in cities than in non-city areas, while the results of Ireland and Belgium show the opposite. This could be due to the booming housing market in Austria during that time, while the housing markets of Belgium and Ireland were stagnant an falling respectively. Another finding that relates to the diversity of the effect among regions is done by Eichholtz et al. (2010). They find a smaller premium for buildings situated in more expensive and bigger regions, which implies that as rent increases, the importance of sustainability decreases (Eichholtz et al., 2010).

Next, the importance of the observation period becomes clear when comparing Eichholtz et al.'s first and second research (2010; 2013). Their findings show how the premium relates to the economic circumstances of a country. The premium decreases from almost 17% in 2007, to 13% right after the financial crisis (Eichholtz et al., 2010; Eichholtz, Kok & Quigley, 2013). According to Eichholtz et al. (2013) energy performance becomes less important during an economic downturn because energy efficiency is seen as a luxury characteristic of a house. Similarly, when Brounen & Kok (2011) look how the premium develops over their observation period, they find that the premium drops to almost zero in the first quarter of 2009, which could be due to the economic downturn in that period.

Besides differences in time periods and regions, there is also a clear division between researches that focus on commercial buildings, especially office spaces, and residential dwellings. Naturally, there will be differences between how businesses evaluate green office spaces and how households evaluate green dwellings. An organisation will consider different aspects when choosing a building, such as the productivity of its employees and corporate reputation (Eichholtz et al., 2010; Fuerst & McAllister, 2011). The effect of energy efficiency might therefore be substantially different for residential dwellings than for office buildings. This seems to be the case when comparing the different premia found in the US between office buildings (Eichholtz et al., 2010) and residential dwellings (Kahn

& Kok, 2014). The differences between offices and dwellings is furthermore shown when looking at the premia found in the UK (Chegut, Eichholtz & Kok, 2014; Fuerst, McAllister, Nanda & Wyatt, 2015).

Furthermore, the researches in Table 2.1 differ in the certification schemes that are examined. The American Energy Star and LEED as well as the Swiss Minergie are different from the EPBD. In these programs it is only possible for buildings to be awarded with a certificate of being energy efficient. Hence, unlike the EPBD, they have no different classes in their energy-efficiency. The benchmark used by the Energy Star program *"is chosen so that the label is awarded to the top quarter of all comparable buildings"* (Eichholtz et al., 2010, p. 8). The EPBD on the other hand has a rating system ranging from label A to label G. These differences in programs probably also attribute to the wide range of premia found.

Not only do the papers differ from each other in regions, observation periods, type of buildings and certification programs, but also in more methodological aspects. Although all researches seem to use a hedonic price model for the estimation regression, which will be explained in greater detail in Chapter 3, the researchers use profoundly different approaches.

Some researchers base their results on a very small number of observations, which might make them less reliable (Bloom, Nobe & Nobe, 2011; Mudgal et al., 2013; Salvi, Horehájová & Syz, 2010). Moreover, the researchers that compare labeled buildings as the treatment-group versus non-labeled buildings as the control-group seem to have a large sample on first sight (Eichholtz et al., 2010, 2013; Fuerst & McAllister, 2011; Chegut et al., 2014; Kahn & Kok, 2014). However, in reality the data samples consist mostly of the control-group, which are non-labeled buildings. For example, the samples of Eichholtz et al. (2010) and Chegut et al. (2014), which consist of only 199 and 68 certified buildings respectively, emphasise this.

This approach of comparing labeled versus non-labeled might give rise to another issue. When requesting a label is not mandatory, it is difficult to make any assumptions on the energy performance of the control-group. Thus, the control-groups might consist of relatively energy-efficient buildings, which could influence the results. Therefore, the EPBD and other programs that have different classes of labels can be seen as more suitable in finding the effect of energy performance on housing values. Mudgal et al. (2013) do look at the effect of the EPBD, but use the EPC-rating as a continuous variable, by which they assume that the differences between the labels are constant. However, it seems more plausible that the premium is varying across the labels. This approach is not used by any of the other studies, except in one other instance to check for robustness (Hyland, Lyons & Lyons, 2013).

Looking at the findings of Bloom et al. (2011) another methodological difference appears: the study uses a linear regression function, while all other researches in the overview use a log-linear functional form. By not using a logarithm for the dependent variable, Bloom et al. (2011) assume that the effect stays the same regardless of the transaction price paid for the house. It seems more preferable to allow for a percentage effect, since the premium might change in relation to the price. Additionally, the logarithm will mitigate the outliers in the dataset.

Taking all these aspects into consideration, we are left with only a few studies that relate closely to this thesis and these will be thoroughly reviewed. Firstly, Hyland et al. (2013) find a premium of 9.3% for a label A compared to a label D. Conversely, an F or G label is worth 10,6% less. The study uses asking prices instead of final transaction prices. However, as the EPCs are especially meant to make house buyers aware of the house's energy performance and not house sellers (who already are aware) it is possible that this influences the results. Moreover, the dataset misses some essential characteristics of the dwellings. In particular, the model only controls for the size of the house by including a variable that indicates the number of rooms. However, one could argue that the number of rooms is not an adequate indicator of the size of a dwelling. Another missing control variable is the year of construction of the dwelling. It is likely that size and age of a dwelling are both related to the energy performance as well as the transaction price. Hence, the exclusion of these variables in the model could lead to substantial omitted variable biases.

The second study that is relatively comparable to this thesis is the work of Fuerst, McAllister, Nanda & Wyatt (2015), who find a significant premium in the UK between 1995-2012. As is mentioned in the paper, 44% of the sample consists of buildings that are more than 50 years old. This might be due to the fact that they use a repeated sales method, which only observes dwellings that are sold more than once in the time frame and thus exclude newly built dwellings. Since energy efficient houses are built more recently, this might explain the very low amount of observations with label A and B in their sample: 7 and 4,405 observations respectively. Furthermore, the method of matching the datasets is debatable. Initially, it is not checked whether the EPC was present at the time of sale in the matched dataset. If a house is sold before it received the EPC, but did receive the EPC within the observation period, it should not be included in the sample, because this could influence the results. Indeed, when Fuerst et al. (2015) check robustness for when the EPC is known at time of sale, their results change. The effect of labels A and B, for instance, lose their significance.

The only study assessing the effect of energy performance on the value of residential dwellings in the Netherlands has been done by Brounen & Kok (2011), which makes it the most comparable paper to this thesis. However, the observation period ranges from 2008 to 2009, which is right at the beginning of the economic downturn that followed the global financial crisis. According to Brounen & Kok this is:

> *"clearly reflected in the distribution of the transactions over the sample period: more than half of the transactions took place in the first two quarters of 2008, with transactions in the housing market virtually grinding to a halt in the third quarter of 2009"* (Brounen & Kok, 2011, p. 7).

Although the authors do not discuss the effect that this could have on their results any further, it is possible that this might therefore not be representative for an average economy. Despite this unfortunate time period, the results show significant effects: green compared to non-green dwellings sell for a premium of 3.7% and dwellings with label A compared to label D sell for a premium of 10.2%.

When looking more accurately at the methodology used in the paper, some aspects require a closer look. First, the authors use exterior maintenance as a control variable in their second regression and find that the premium decreases to 3.6%. However, the descriptive statistics reveal that they also have data on the interior maintenance. As it is likely that people who improve the energy performance of a house do this in combination with other interior renovations, it might have been more accurate to check robustness for the results by including interior maintenance instead of exterior maintenance as a control variable. Furthermore, a bigger concern throughout their research is the methodology in the control for locational effects. An often used mantra within hedonic price models is that house prices are determined solely by three factors: *'location, location, location'*. Yet, Brounen & Kok (2011) use only a few locational characteristics. The research includes province-dummies to control for province fixed effects. However, the Netherlands is divided in only 12 provinces, meaning locational variation within these provinces will remain, especially on micro-level. On neighbourhood level the study includes housing density, average monthly household income and average time on market into the model. Although these three characteristics allow for an economic interpretation of the locational effects, which might be seen as favorable, it does not control for all neighbourhood effects. Dwellings are immobile, so the location of the house is a vital determinant of housing prices and all studies previously reviewed in this paper use some method of controlling for neighbourhood fixed effects, thus the lack hereof in this study might lead to substantial biases.

Table 2.1: Overview of existing empirical researches

Authors	Year	Region	Program	Observ. Period	Type of buildings	N	Comparing	Premium
Bloom, Nobe & Nobe	2011	Colorado (US)	Energy Star	1999-2005	Residential	300	Label vs. No label	$8.66 per foot
Brounen & Kok	2011	NL	EPBD	2008-2009	Residential	31,993	Green vs. Non-green	3.7%
							Label A vs. D	10.2%
Cajias & Piazolo	2012	Germany	EPBD	2008-2010	Residential	2630	Green vs. Non-green	2.84%
Cerin, Hassel & Semenova	2014	Sweden	EPBD	2009-2010	Residential	67,559	Green vs. Non-green	6.0%
Chegut, Eichholtz & Kok	2013	UK	BREEAM	2000-2009	Office	2,103	Label vs. No label	14.7%
Eichholtz, Kok & Quigley	2013	US	LEED and Energy Star	2007 & 2009	Office	5,993	Label vs. No label	13.3%
Eichholtz, Kok & Quigley	2010	US	LEED and Energy Star	2004-2007	Office	1,813	Label vs. No label	16.8%
Fuerst & Mc Allister	2011	US	LEED and Energy Star	1999-2008	Office	6,157	Label vs. No label	30%
Fuerst, Mc Allister, Nanda, Wyatt	2015	UK	EPBD	1995-2012	Residential	333,095	Label A vs. D	5%
Hyland, Lyons & Lyons	2013	Ireland	BER	2008-2012	Residential	15,060	Label A vs. D	9.30%
Kahn & Kok	2014	California (US)	LEED, Energy Star & GreenPoint	2007-2012	Residential	4,321	Label vs. No label	9.00%
Mudgal, Lyons, Cohen, Lyons & Fedrigo-Fazio	2013	Ireland	BER	2008-2012	Residential	11,247	Continuous rating	2.80%
		Flanders, Wallonia & Brussels (BE)	CPEB	2012	Residential	18,200	Continuous rating	2.9-5.4%
		Oxford (UK)	EPBD	2012	Residential	236	Continuous rating	-
		Vienna (AT)	EPBD	2012	Residential	1,189	Continuous rating	8%
		Lille & Marseille (FR)	EPBD	2011-2012	Residential	3,178	Continuous rating	3.2-4.3%
Salvi, Horehajova & Müri	2008	Switzerland	Minergie	1998-2008	Residential	259	Label vs. No label	3.5-7%

3 Theoretical Background

Before empirically estimating the premium of a green dwelling, it is crucial to understand some of the specific characteristics of the housing market. Specifically three aspects are important for our understanding of the green versus non-green housing market. This chapter will explore these aspects in the next three sections. Firstly, the issue of imperfect information on energy performance in the housing market -and how EPCs are supposed to solve this by creating awareness- is discussed. Secondly, the mechanism that works from EPCs to housing values is explained by looking at the interaction of relative demand and relative supply of green dwellings. Lastly, the hedonic price model is introduced, which will help to establish the right empirical model for the heterogeneous housing market.

3.1 Imperfect information in the housing market

Simple economic theory predicts that the discounted energy savings that one expects from living in a green house would sum up to the premium paid in advance for the dwelling. This implies that a rational buyer is indifferent when it comes to the energy performance of a house.

However, since house buyers are not completely aware of all the characteristics of a house, there is information asymmetry. Specifically, there is a problem of imperfect information when it comes to the energy performance of a building. House buyers are often not completely informed about the energy performance of the dwelling they want to buy and thus might not be able to distinguish between a green and a non-green dwelling. This might lead to an incorrect pricing of energy performance of houses. Palmquist (2005) argues that housing values reflect the expectations of people and that these values change when people revise their expectations. However, the expectations greatly depend on the awareness of people and the information they have (Palmquist, 2005). This issue of information asymmetry can be best explained with Akerlof's (1970) 'market for lemons'.

Theoretically, when house buyers would not have any information on the energy performance of a house and thus cannot distinguish between a green dwelling and a non-green dwelling, they will pay a lower amount for a green dwelling than what it is worth. This means there is no incentive for house owners to actually invest in making a house greener. This issue is also discussed by Jaffe & Stavins (1994), who argue that due to imperfect information on future energy savings, energy-saving measures are not sufficiently taken by the market.

Since the use of (non-green) energy contributes to global warming, the social costs exceed the private costs. Due to the lack of incentive to create more green dwellings, this gap between the private optimal and the social optimal amount is difficult to reduce. Theoretically this could be solved by introducing a tax on energy which would then create a socially optimal price for energy consumption (Pigou, 1920). In practice, even though energy consumption is taxed in the Netherlands, we do not see the desired result.

If all house buyers would have been perfectly informed about the energy performance in the housing market, it would enhance the premium paid for green dwellings. The introduction of EPCs are supposed to help with this by creating awareness. According to the European Commission:

> *"the basic idea behind certificates is to create information that such actors can use to make more informed decisions and integrate energy efficiency into their decision-making process. ... It can be hypothesised that the improvement of the energy performance of a building should then also lead to higher transaction prices and rents on the market" (Mudgal et al., 2013, p. 18).*

However, the European Commission (2013) concludes from their research that the energy performance of dwellings is not yet fully capitalized in the countries that were examined, even though EPCs were present. They argue that this is due to what they call the *'information gap'*. Specifically the correct valuation of EPCs depends on four factors (Mudgal et al., 2013):

- the presence of the label;

- people's understanding of the label;
- their concern about energy savings; and
- their trust in the label and related policies.

The European Commission therefore argues that the valuation of EPCs highly depends on its implementation. They recommend campaigns and programmes to make the EPC scheme more easy and practical for people (Mudgal et al., 2013). The better the EPCs are implemented, the better people will be informed and the smaller this information gap will be. It can therefore be expected that the new policy from 2015 onwards that made the labels more accessible will have increased the premium paid for green dwellings.

In short, the premium people are willing to pay for green dwellings has shown to be dependent on people's awareness and expectations of the energy performance in the housing market. The amount of information in the market influences the discounted value of the future energy savings. The less information people have, the smaller the premium, and vice versa. Existing literature for example argues that if people anticipate a future increase in energy prices due to energy taxes, this should increase the premium (Eichholtz et al., 2010). As the EPCs try to increase people's awareness on energy performance, one would expect an increase of the premium.

3.2 Demand and supply effects

A typical characteristic for the total housing market is the inelastic supply in the short-term (Dijk, Groot & Möhlmann, 2016). The stock of buildings in the market is relatively fixed and cannot easily respond to changes in demand. Building new houses takes time and therefore the supply will adjust very slowly. Figure 3.1 illustrates this with a vertical supply curve on the short-term. Since demand responds negatively to prices, the demand curve is downward sloping. Because of this inelastic supply, it is the demand side that determines the price of buildings on the short-term (Palmquist, 2005, p. 3). In Figure

3.1 it can be seen that because of the inelastic supply in the short-term, the price P^* is determined by the demand for houses. The quantity of houses Q^* is fixed.

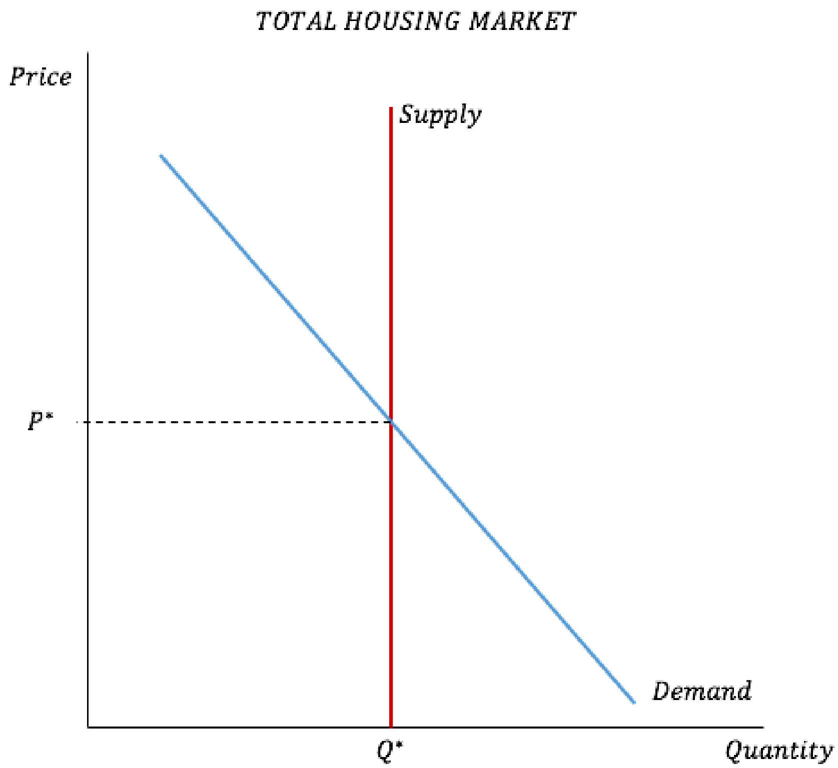


Figure 3.1: Short-term supply and demand diagram for the housing market in general

3.2.1 The green housing market

Now, turning to the green housing market, the next two diagrams illustrate the interaction with demand and supply for the green relative to the non-green housing market. The vertical axis therefore illustrates the relative price of a green dwelling compared to a non-green dwelling. Similarly, the horizontal axis illustrates the relative quantity of green dwellings compared to non-green dwellings in the housing market. Hence, the values on the axes are ratios.

The relative supply of green buildings is not as inelastic as the total supply on the housing market, as it is easier to make an existing house greener then to build a new house. Therefore, the supply curve will not be exactly vertical on the short-term. Instead it is upward sloping, because supply increases with the price. Similar to the total housing market the supply curve will flatten further when the time period increases.

Assuming rationality

Let us first assume that people are completely rational and make their decisions regarding green dwellings only based on their economic motivation, which is the same for everyone. The price people are willing-to-pay for a green dwelling, will exactly reflect the price of a non-green dwelling plus the discounted value of the expected future energy savings. A person will therefore be indifferent between a more expensive green dwelling with lower future energy costs or a less expensive non-green dwelling with higher future energy costs. This assumption implies perfectly elastic demand fixed at a certain price, which is illustrated in Figure 3.2 with a horizontal demand curve.

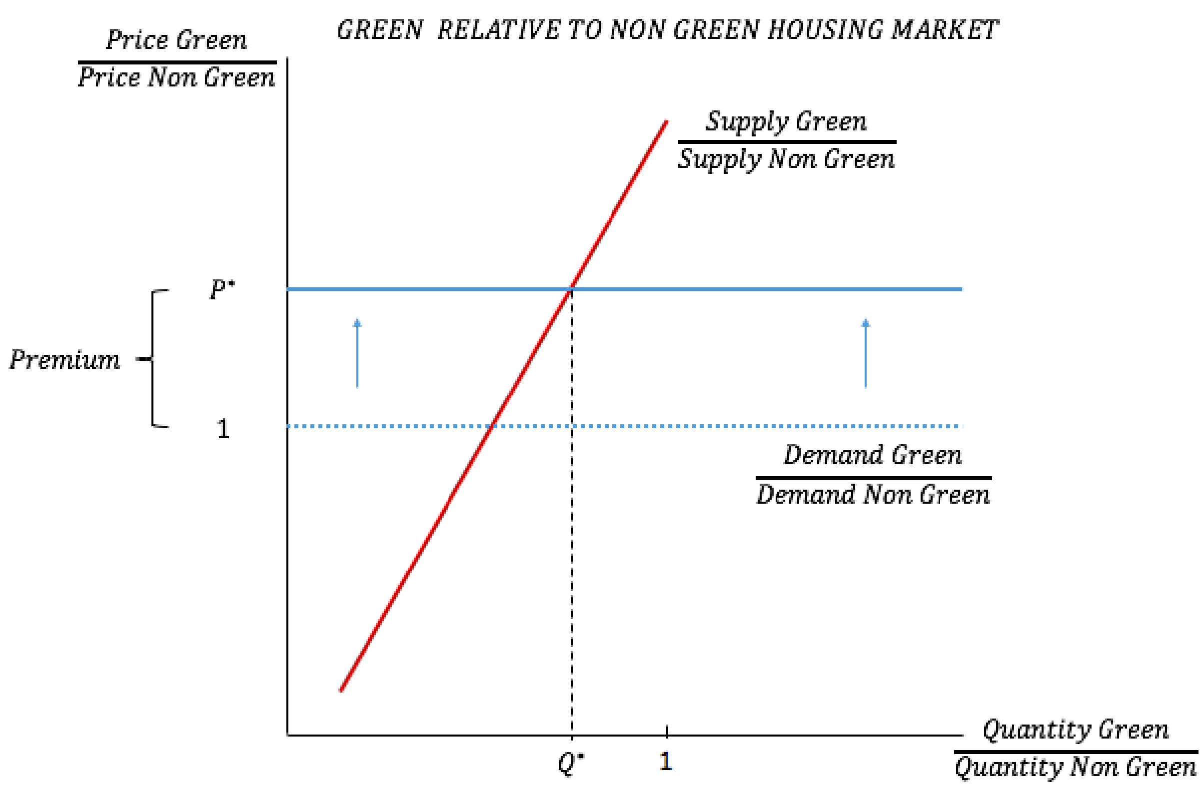


Figure 3.2: Relative supply and demand diagram of the green compared to the non-green housing market when people are completely rational

Theoretically, if there would be no information at all on the energy performance in the housing market, the price for green dwellings would be exactly the same as the price for non-green dwellings, which means the relative price would equal 1. In reality however, house buyers are always able to observe some of the energy-saving characteristics of a dwelling and will thus be able to make a rough estimation of the future energy savings. The dotted line in Figure 3.2 that equals 1 can thus be seen as the lower bound of the relative price.

The objective of the EPCs is to solve the information asymmetry regarding energy performance in the housing market. Due to the EPCs, people can more accurately estimate the expected future energy savings. Hence, as is illustrated in Figure 3.2 relative demand for green dwellings shifts upwards and the relative equilibrium price P^* increases. The premium paid for green dwellings will then be $P^* - 1$. The magnitude of the premium is thus depending on how effective the EPCs are at solving the information problem. The upper bound of this premium will be defined by a relative price, where there is complete information in the market. This implies that depending on the degree of information asymmetry in the housing market the price P^* will vary between the lower bound and the upper bound.

When the relative price increases due to the introduction of EPCs, relative supply will respond, by increasing the relative quantity of green dwellings Q^*. Hence, the quantity Q^* is completely determined by the supply of green dwellings. As long as the market has less green dwellings than non-green dwellings, Q^* will be below 1. When the supply curve flattens, changes in price will correspond with larger changes in quantity. Furthermore, when external factors would shift the relative supply curve of green dwellings compared to non-green dwellings, this would only impact the equilibrium quantity and not the premium.

To summarize the above, the policy of introducing energy labels induces a price premium and stimulates people to invest in energy-saving measures in the existing built environment, which is exactly the objective of the policy. The percentage of the premium depends on the implementation of the EPCs and how effectively the EPCs solve the issue of imperfect information. The more information on energy performance that people have in the housing market, the higher this premium will be.

Assuming personal preferences

The perfectly elastic demand, which assumes all people to be the same and rational, can be seen as too stringent. It might be more accurate to assume that people have personal preferences when choosing between green or non-green dwellings, such that some people prefer to pay more for a green dwelling than other people, even if they have equal information. Figure 3.3 demonstrates what happens to the premium when demand is not completely elastic.

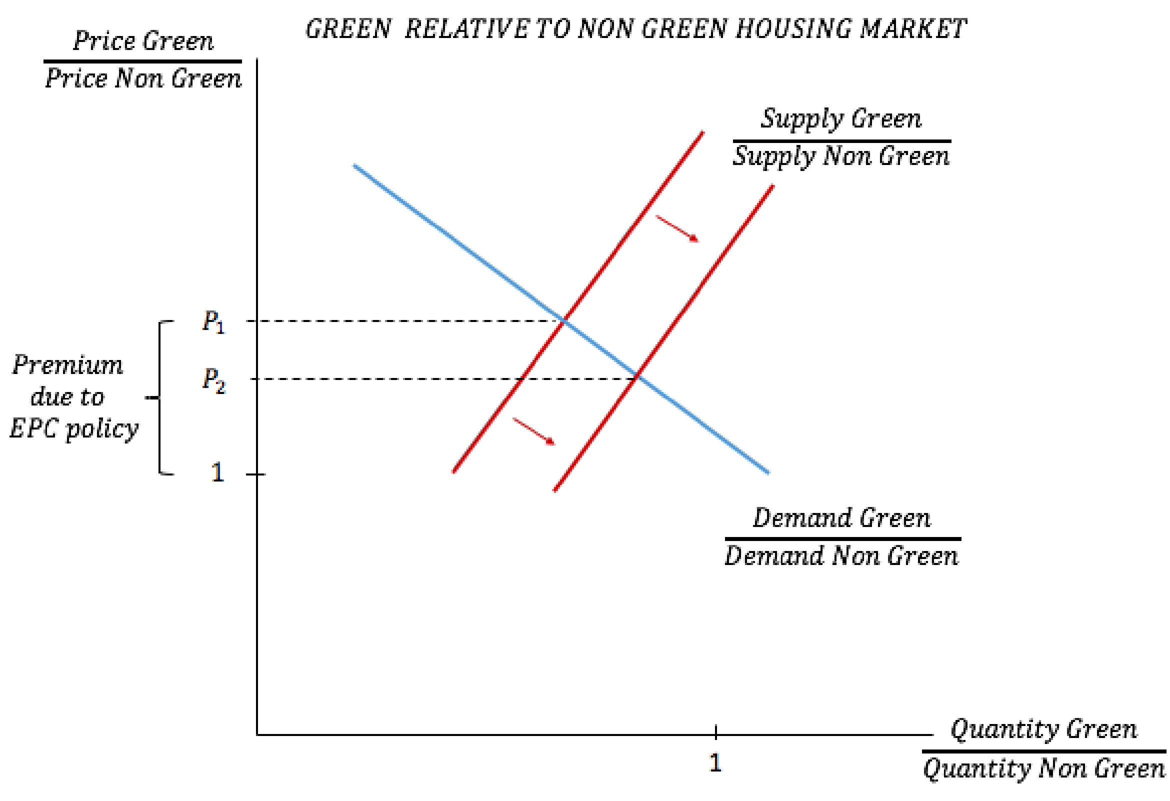


Figure 3.3: Demand for green buildings increases relative to non-green buildings due to policy introducing EPCs

The relative demand curve is downward sloping, as relative demand for green dwellings increases when the relative price decreases. When relative supply is fixed during the observation period, we can simply state that the premium reflects exactly the shift in demand due to better informed people. In other words - similarly to when we assume rationality- the premium completely reflects the effect of the EPBD policy. However, what happens when the supply of green dwellings increases simultaneously due to other reasons than the change in price?

It seems likely that external factors, e.g. other policies focused on increasing the energy performance in the built environment from the supply-side of the market, have increased the supply of green buildings during the observation period. This increase in relative supply, illustrated in Figure 3.3 as a shift of the relative supply curve to the right, would drive the relative price down to P_2. The price premium will then not only reflect the effect of the EPCs on the price, but also other *'supply-effects'*. The premium that only reflects the effect of the EPCs solving the issue of imperfect information equals $P_1 - 1$. However, empirically one would find a premium of $P_2 - 1$. Therefore, assuming personal preferences makes it harder to correctly estimate the premium due to the EPCs. When the relative supply of green dwellings has increased due to external factors, one would empirically underestimate the effect of the EPCs. The effect of the implementation of EPCs on the demand for green buildings will in reality be larger.

3.3 The hedonic price model

Another characteristic of the housing market that brings us to the empirical model used in this thesis, is heterogeneity. Every house is unique in the combination of its locational and physical characteristics. We assume that houses are sold as a complete package of these characteristics and that they cannot be sold partially. The price of such a good is thus completely depending on the price of these underlying characteristics. The hedonic price model uses the interaction between demand and supply to determine the price of a good (a dwelling in this case) by the valuation of its characteristics (Malpezzi, 2003). It is a commonly used model to determine the valuation of houses and their underlying characteristics.

As mentioned by Malpezzi (2003), who reviews the *'roots of hedonic price models'*, it dates back to Lancaster (1966), who provided the first micro-economic foundations for the model (Malpezzi, 2003, p. 74). The hedonic price model used for this research is based on Rosen's (1974) analysis on the market equilibrium for heterogeneous products, which he calls a *'package'* of characteristics that cannot be untied. Rosen states that buyers and sellers maximize their utility, which is determined by the combination of the characteristics, subject to their budget constraints. The bid functions (demand side) and the offer functions (supply side) determine the implicit price for a characteristic of the product.

Following Rosen, every house consists of a distinct set of n characteristics, which we will call $Z = (z_1, z_2, ..., z_n)$. The price of a house is determined by these characteristics and their implicit prices. Therefore, $P(Z) = P(z_1, z_2, ..., z_n)$. The hedonic price model is then used to estimate the value of buildings through these characteristics (Rosen, 1974). These characteristics are often subdivided in physical characteristics of the house (X), locational characteristics (L) and the time of measurement (T) (Francke, 2014). The basic regression model is therefore of the following form:

$$f(V) = g(X, L, T) \tag{3.1}$$

The left-hand side is a functional form of the value of a house, which in this case will be measured by transaction prices. A popular functional form is the natural logarithm. It has some clear advantages, such as that it allows for non-linear marginal effects of the characteristics, it has a clear interpretation and it mitigates heteroskedasticity of the error terms (O'Sullivan & Gibb, 2008, p. 80). The right-hand side is a functional form of characteristics of the house. According to Malpezzi (2003, p. 79) the basic variables used as physical characteristics of the house are: number of rooms, floor area, type, age and the presence of certain features (parking, elevator, basement, etc.). As will be discussed in greater detail in the next chapters, the model of this research makes a distinction between the energy performance of a building and all other physical characteristics.

Chay and Greenstone (2005), make clear why this hedonic price model is the best method to estimate the valuation of energy performance in the housing market:

> *"at each point on the hedonic price schedule, the marginal price of a housing characteristic is equal to an individual consumers marginal willingness-to-pay for that characteristic and an individual suppliers marginal cost of producing it" (Chay & Greenstone, 2005).*

However, this quote also highlights the drawback of the model, namely that the valuations of the houses used in the hedonic price model are assumed to be the result of the interactions of demand as well as supply of houses (Francke, 2014). More specifically, the transaction price lies in the area where the reservation price asked by house-sellers (supply-side) and the price that house-buyers are willing to pay (demand-side) overlap (van Dijk, Geltner & Minne, 2017). Assuming that there is no bargaining power, the price will be exactly in the middle of this area. This makes it impossible to disentangle the effect on price into a clear supply effect on the one hand and a clear demand effect on the other hand, which the previous section showed to be important for concluding whether the price premium reflects the EPC policy.

The next chapter will empirically estimate this premium paid for green dwellings. In case of a significant and positive price premium, it is therefore important to keep this in mind, as it could represent a lower bound of the real premium reflecting only the EPCs.

4 Empirical Model

In this chapter the empirical regression model used for estimating the effect of energy performance on housing values, is presented and explained. The model builds upon the theoretical framework from the previous chapter. Additionally, the chapter reviews the endogeneity issues that arise when empirically estimating the premium.

4.1 The regression model

Two separate models are used to first estimate the premium and to secondly estimate the impact of the new policy on the premium. All the variables included in the models are defined and discussed.

4.1.1 Estimating the premium paid for green dwellings

The basic regression model that this research will try to estimate is the following:

$$lnP_{int} = \alpha + \beta Green_{int} + \gamma' X_{int} + \delta_n + \zeta_t + \epsilon_{int} \tag{4.1}$$

where P_{int} is the transaction price of building i in neighbourhood n in time period t. By taking the natural logarithm of the price and - where possible - other explanatory variables, the coefficients can be seen as (semi-) elasticities, making the results easier to interpret. In addition, it mitigates the outliers in the data. $Green_{int}$ is a dummy variable which equals one in case that building i in neighbourhood n in time period t is a *green* house and zero otherwise. A green house is defined as a house with label A, B or C. The coefficient β measures the effect of energy performance on the transaction price and thus estimates the premium in percentages. Alternatively, the regression uses a categorical variable[1] indicating the label to estimate the effect per label. The controls consist of X_{int},

[1] A categorical variable exists of a binary variable per category.

which is a vector of characteristics of the building, with γ' as the corresponding coefficient vector. δ_n and ζ_t are neighbourhood specific and time specific effects respectively, which are considered fixed effects in estimation. Lastly, ϵ_{int} is the error term. To avoid ruling out any correlation between the error term across the neighbourhoods, the model will cluster standard errors at the neighbourhood level.

Regarding the vector of control variables that present the characteristics of a building, there is an extensive amount of literature discussing the variables that need to be included in hedonic price models. The characteristics used in this thesis are consistent with the variables that are most commonly used in hedonic price models. The size of the dwelling, which certainly relates to the energy performance as well as the price, is included as the volume in cubic meters as well as the number of rooms. I take the logarithm to control for outliers of the volume. The number of rooms takes the form of a categorical variable[2]. This is done to allow for non linear effects, because I argue that the effect on transaction price of going from one to two rooms will be different than the effect of going from ten to eleven rooms and it is therefore necessary to distinguish these.

The construction period will be included in the model to control for the age of the building. Since certain time periods will influence the transaction price and energy performance of a dwelling, rather than a specific year, I include categorical variables indicating the construction period of the dwelling[3] instead of controlling for the specific age of the dwelling. Additionally, an extra dummy variable indicating whether the dwelling is a monument is included, as monuments have specifically high restrictions for energy-saving measures and simultaneously spur the price. Furthermore, the type of dwelling will be controlled for by adding a categorical variable varying between a detached, semi-detached, terraced, end-of-terraced dwelling or an apartment. Finally, the dataset allows to control for the level of maintenance of the dwelling, which ranges from 1 (in a very bad condition) to 9 (best

[2]The variable for number of rooms is categorized in 9 separate categories. The first six categories are complying with the number of rooms. Thereafter, there is a seventh category for buildings with 7-10 rooms, an eighth for 11-20 rooms and lastly a ninth category for buildings with more than 20 rooms.

[3]The variable for construction period has 9 different categories, which can be viewed in Table B.1 of Appendix B.

possible condition).

Although the categorical variables represent more realistic non-linear relationships, it also means the model will include a wide range of binary variables. The (neighbourhood) fixed effects will increase this amount of binary variables even more. This is important to recognize when evaluating the explanatory power of the model, as it will probably significantly increase the R-squared of the model. Essentially, hedonic price models try to determine the valuation of a house as best as they can, so it is common for such models to have high R-squareds (van Dijk & Francke, 2008; Dröes & Koster, 2016; Eichholtz et al., 2010, 2013; Fuerst et al., 2015). However, it is important to find the right balance of on the one hand adding the right controls to avoid omitted variable biases and on the other hand of not *overfitting* the regression. To better investigate this balance Appendix C will show what happens to the results when the model is gradually simplified, by excluding or changing some of the fixed effects and controls.

4.1.2 Estimating the impact of the new policy on the premium

The model is then expanded by including an interaction term to find whether the change in policy in 2015 had an effect on the premium. This second regression model has the following form:

$$lnP_{int} = \alpha + \beta_1 Green_{int} + \beta_2 Green_{int} * After2015_{int} + \gamma' X_{int} + \delta_n + \zeta_t + \epsilon_{int} \tag{4.2}$$

where $After2015_{int}$ is a dummy variable that equals one if the year of sale was 2015 onwards (after the policy change). The coefficient β_1 reports the estimated effect of a green dwelling on transaction price before 2015, while the coefficient β_2 reports the change in effect after 2015. In other words, when β_2 is significant and positive, one can argue that the change in policy has had an impact on the effectiveness of the EPCs.

4.2 Limitations to the model

Ideally one would like to observe the price of a house throughout a period and then go back in time to only change the energy performance to a label A, B or C and observe what would happen to the price of the house, but in reality this is not possible. OLS regression is a common used method for estimating the effect of a treatment (in this case improving the energy efficiency) and thus for finding a causal relationship. However, as is mentioned by Angrist & Pischke (2008), whenever the sample is not randomized the estimated effect can be biased and could ultimately even reverse the outcome of the study. Whether the estimated effect is unbiased depends crucially on the correlation between the error term and the independent variable, or in this case the energy performance of the dwelling (Angrist & Pischke, 2008). If the error term is correlated to the energy performance of the dwelling, it will lead to endogeneity. Depending on the correlation between the error term and the variable of interest this leads to an over- or underestimation of the regression coefficients. It is therefore important to consider the potential endogeneity errors that might bias the results.

4.2.1 Omitted variable bias

The most worrisome bias would be an overestimation of the premium, since this could lead to incorrect significance of the OLS estimate. Overestimation would exist when a variable that positively influences the value, as well as the energy performance of a house, is not controlled for. The impact of energy performance on the value of houses would then be overstated. An example of this could be when newly renovated buildings might consistently be correlated to higher energy efficiency of the building [4]. This thesis tries to control for this bias by including a maintenance variable into the model. However, it is important to keep in mind that the findings still depend on how valid this measure of maintenance is. Moreover, if the maintenance factor incorporates energy efficiency of the dwelling, the control variable could absorb some of the relevant information of a dwelling's energy

[4] It is plausible that people consistently combine the renovation measures with energy-saving measures.

performance. This could lead to an underestimation of the effect of energy performance. Therefore, in Chapter 7 extra attention is paid to this variable, to examine how it might influence the results.

On the other hand, underestimating the premium would be a less worrisome bias because it would mean the OLS estimate is a lower bound of the real effect. Any variable that would have an opposite relationship with the value of the house in comparison to the energy performance, would create an underestimation of the premium when left out of the regression. For instance, monuments might have a positive effect on the price, but a negative effect on the energy performance. We therefore also include a control variable for a houses monumental status.

By including fixed effects, one can also control for certain unobserved variables that also have an effect on the causal relationship. This research therefore controls for neighbourhood-specific variables that do not vary across years by including neighbourhood dummies. It also controls for differences in time, by including time-dummies. To address the issue of omitted variable bias more carefully, Chapter 7 will focus on several robustness checks, to show what happens to the results when certain variables are omitted from the regression.

4.2.2 Self-selection in sample prior to 2015

Since the presence of an energy label at the sale of a house was not yet mandatory before 2015, there might be self-selection into the sample. Therefore, the part of the sample between 2008 and 2015 requires some special attention. When the unobserved factors that determine whether a person requests a label or not (thus determining whether the person is part of the sample) are correlated with the unobserved factors in the regression model this could lead to biased results.

As will become clear when analyzing the data in the next chapter, the dataset is substantially bigger after 2015. When estimating the first model the issue of selection-bias cannot be eliminated. However, the part of the dataset with a potential bias will be relatively small compared to the total sample. Even though, it is therefore plausible that the impact on the results will be limited, the bias cannot be neglected and needs to be taken into consideration when interpreting the results.

As the self-selection issue is only present in the sample prior to 2015, the bias has a different impact on the results of model 4.2, where the sample before 2015 is compared with the sample after 2015.

The bias can go in two directions. If the unobserved factor determining self-selection is positively correlated with transaction prices and energy performance, or if negatively correlated with both, it could lead to overestimation of the effect of energy labels on transaction prices prior to 2015. The overestimation of the effect prior to 2015, would then imply that the change in premium due to the new policy (introduced in 2015) is underestimated. When a significant change of premium with model 4.2 is found, this will represent the lower bound of the change in premium. The question that remains in this case is whether the estimated change in premium is too conservative and if it might have been a bigger change in reality.

However, the bias could also lead to an underestimation of the effect of energy labels on transaction prices prior to 2015. This would be the case if the correlation between the unobserved factor determining self-selection and transaction prices has the opposite sign from the correlation with energy performance, e.g. if people consistently requested a label when having high energy performance but low transaction prices (or the other way around).

The presence of a selection-bias when the EPCs are not mandatory was also noticed in other empirical researches (Brounen & Kok, 2011; Hyland et al., 2013). On the one hand, Hyland et al. (2013) reason that a selection-bias could exist due to an unobserved quality factor. In this case house-owners of dwellings of higher quality have higher energy-performance and are therefore more likely to get a label. Since this quality factor would also positively influence the price of the dwelling, it would lead to overestimation. However, when checking for this bias with the Heckman method, Hyland et al. (2013) do not find any significant correlation.

On the other hand, Brounen & Kok (2010) find a negative correlation with the Heckman method, which would mean that *"the unobserved factors that make energy-labeling more likely tend to be associated with dwellings with lower transaction prices"* (Brounen & Kok, 2011, p. 11). Although this seems contrary to the reasoning of Hyland et al., Brounen & Kok do not explain this any further. Finally, both studies conclude that controlling for this bias with the Heckman method did not change any of the results substantially.

The above mentioned concerns create challenges for the empirical model and need to be kept in mind throughout this thesis. In the next chapters, I will discuss the data and results of this thesis and address the endogeneity issues mentioned by checking robustness of the results.

5 Data

5.1 Data sources

The Netherlands does not have one dataset including the EPCs as well as the housing prices and characteristics. Therefore, the data obtained in this research is coming from two different data sources.

5.1.1 Data on energy performance

For this research I received a dataset from the *Rijksdienst voor Ondernemend Nederland*[1] (RVO) with all EPCs in the Netherlands, also referred to as the final energy labels. The dataset is updated every six months. The most recent version is from January 2018 and consists of more than three million dwellings with a final energy label. It also includes the date when the label was registered, which makes it possible to check whether the EPC was present at the time of sale.

Besides the label, the RVO dataset includes for a large number of observations the Energy Index as well. As was discussed before, the index compares the energy efficiency of the building to the energy efficiency of the average building in 1990[2]. The Energy Index is based on much more detailed information and is thus a more precise proxy for the energy performance of the building. However, since the new policy made it possible to receive an energy label without having the assessor present, the information on the dwelling's energy performance is less detailed than before. Therefore, the dataset does not include Energy Indices for dwellings after 2015. The Energy Index and the energy-label nowadays comply with each other in 92% of the cases[3].

[1] A public organisation for entrepreneurs in the Netherlands.

[2] For example: a building with an index of 0.3 uses 30% of the energy used by an average building in 1990.

[3] Tweede Kamer, 2017-2018, 30 196, nr. 563.

5.1.2 Data on housing transactions

The *Nederlandse Vereniging voor Makelaars*[4] (NVM) provided the data on housing transactions. Besides characteristics on the transaction, which include the transaction price, date and time on the market, it also includes a large amount of physical characteristics of the dwelling. Characteristics that are included in the dataset are:

- measures of the size of the dwelling: in number of rooms and in square meters as well as cubic meters;
- certain facilities of the dwelling, such as an attic, parking, elevator, etc.;
- the year the dwelling was built;
- the type of dwelling (e.g. apartment, detached etc.)
- measures of how well it was maintained [5].

As the implementation of EPCs started in 2008, the observation window for this research will be from 2008 to the end of 2017. Tabel 5.1 shows the number of observations for both datasets.

Table 5.1: Number of observations

	EPC		NVM		Matched	
	N	%	*N*	%	*N*	%
Total	**3,556,972**		**1,156,896**		**334,750**	
Before 2015	1,976,544	56%	693,015	60%	69,360	21%
After 2015	1,580,428	44%	463,881	40%	265,390	79%

[4]Dutch Association of Realtors. Over 4000 realtors are member of the NVM and their combined dataset dates back to 1985 and covers almost 75% of the Dutch housing market.
[5]The measure used in this thesis ranges from 1-9 and is decided upon by the realtor.

5.2 Matching procedure

In both datasets the dwellings were identified by postcode, house number and suffix [6]. Since there is no clear notation of a suffix in the Netherlands (e.g. different cities use different notations), merging the datasets according to this variable brought some complications. Within the EPC dataset 13% of the data consisted of dwellings with more than one dwelling on the same address but with a different suffix. In only 3.4% of the cases these dwellings had different labels. This seems like a very small amount. However, if this is not randomly distributed, a bias could be created when only matching on postcode and house number. Therefore, I cleared as much of the suffices as possible in both datasets to make them as similar as possible. Afterwards both datasets were matched by postcode, house number and the *'clean'* suffix. After matching, there were more than $600,000$ observations left. However, in many of these observations the EPC was requested after the sale of the house and thus not present at the time of sale. I therefore cleared the matched dataset further by eliminating all observations that received the energy label after the sale of the house. The final sample had a total of over 300,000 observations, as can be seen from Table 5.1.

5.3 Data descriptives

According to Table 5.1 approximately 60% of the observations of both data sources are from before 2015, which almost coincides with two-thirds of the observation period (2008-2014). This is only 21% in the matched sample.

Figure 5.1 shows how the total number of transactions from the NVM datasets relative to the number of transactions with a label have developed throughout the observation period. The change of policy from 2015 onwards is clearly visible.

[6] A suffix is an addition to the house number indicating for example on which floor the dwelling is located. For example a dwelling might be located at postcode 1053 HL, house number 62, on the first floor, the suffix would then be I.

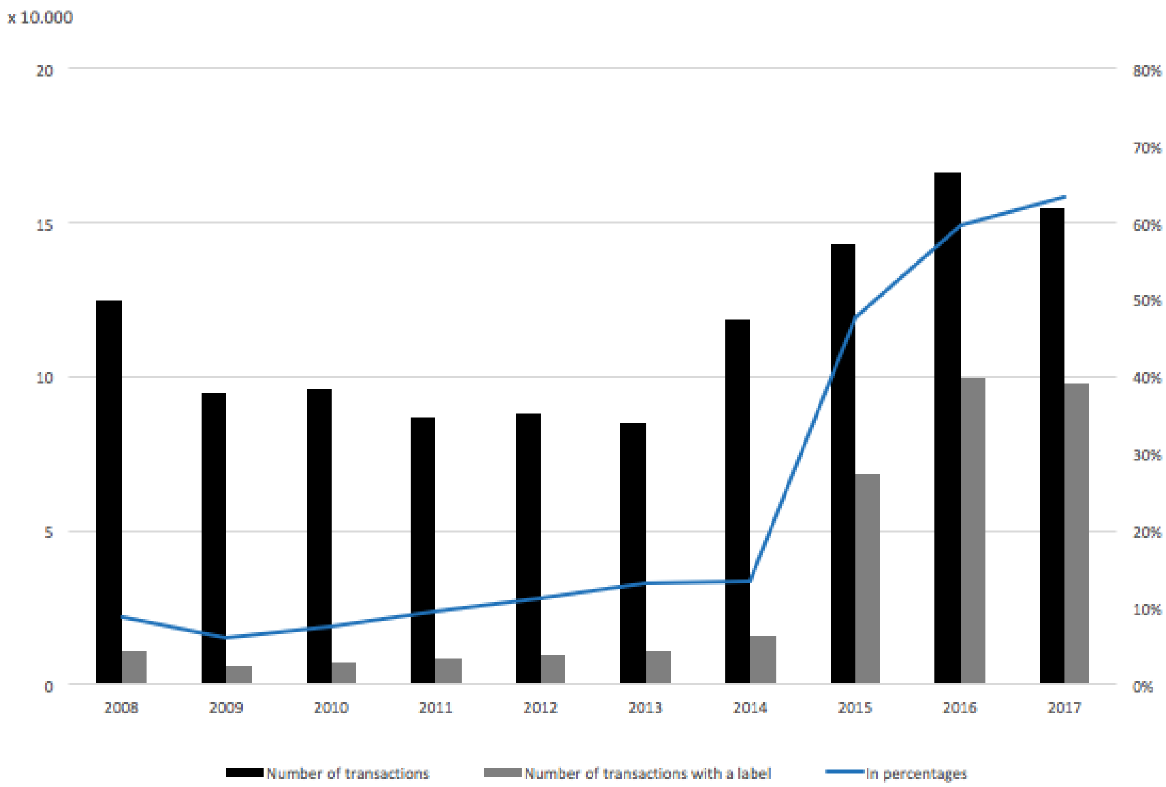


Figure 5.1: Development of transactions with labels from 2008 until 2017. The left vertical axis shows the number of observations and the right axis shows the percentage

Table 5.2 shows how the labels are distributed amongst the sample. A recently published study by the *Economisch Instituut voor de Bouw.*[7] shows the distribution of energy performance for the entire Dutch housing market (and not only for the part of the housing market that is officially labeled) (van Hoek & Koning, 2018). Comparing it to the sample of this research, the distributions are highly similar. This suggests that the total sample is representative for the entire population. Additionally, Table 5.2 compares the distribution of labels before with the distribution after the policy change. The number of dwellings with label A changed the most: these dwellings made up only 2% of the sample before 2015 and increased to 14% after the policy change. Furthermore, the distribution also shows a substantial relative decrease in labels D when comparing the distribution before and after the policy change. These changes might indicate that the EPBD has had the

[7]Economic Institute for Construction

desired effect, but it could just as well be that other external factors influenced the supply of green dwellings (e.g. supply-effects).

Table 5.2: The distribution of labels

Energy-label	2008 - 2017		Before 2015		After 2015	
	N	%	*N*	%	*N*	%
A	38,914	12%	1,316	2%	37,598	14%
B	49,173	15%	6,865	10%	42,308	16%
C	98,447	29%	19,561	28%	78,886	30%
D	61,299	18%	18,343	26%	42,956	16%
E	41,369	12%	12,273	18%	29,096	11%
F	26,977	8%	7,271	10%	19,706	7%
G	18,571	6%	3,731	5%	14,840	6%
Total	**334,750**	**100%**	**69,360**	**100%**	**265,390**	**100%**

The boxplot in Figure 5.2 gives an impression of how the EPCs relate to the transaction prices of dwellings. There is no clear relationship between energy performance and transaction prices visible yet.

The complete summary statistics can be found in Table 5.3. The average transaction price is highest at label A, decreases until D/E and then slowly increases again. The opposite is true for Time on Market (TOM), where labels C, D and E have the shortest average average TOM. Regarding size, the volume variable and the number of rooms variable do not seem to follow an identical trend, although they do have the lowest means among dwellings with labels D and E. As was already expected in the previous chapter, the best maintained houses are label A with an average maintenance grade of 7.5 (out of 9). The grade then consistently decreases with the labels until an average of 6.2 for dwellings with labels G. Unsurprisingly, most monuments have labels F and G. Since the new policy makes exceptions for monuments, there are only a few monuments in the sample. Furthermore, most dwellings in the sample are apartments or mid-terraced dwellings. Lastly, more than 70 percent of the A -label dwellings are built after 2000. This result is expected, as the EPBD enforces a strategy to make all new buildings *zero-energy* by 2020 (European-Commission, 2018b).

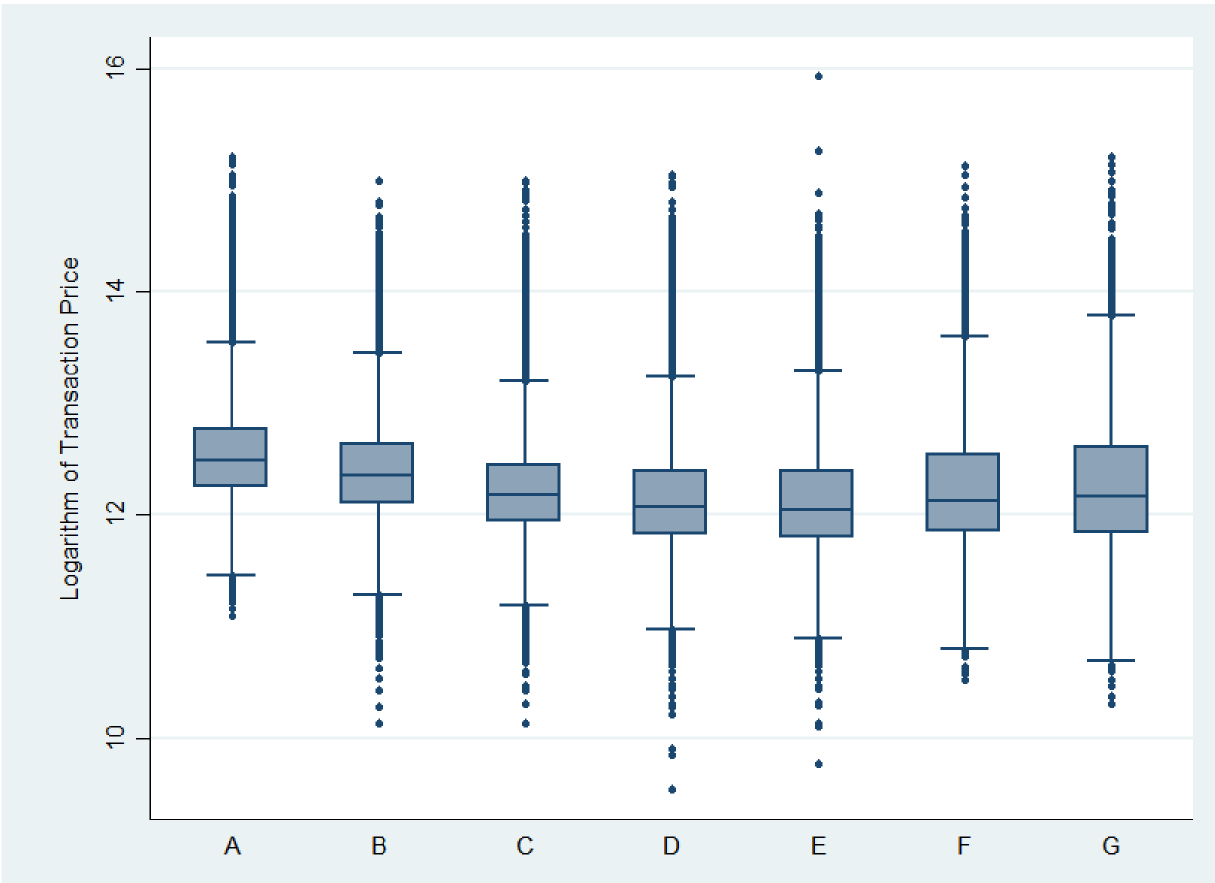


Figure 5.2: Boxplot of the logarithm of transaction price over the energy-labels

Table 5.3: Summary statistics

	Total sample		A		B		C		D		E		F		G	
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
Transaction Price	237,568.70	148,356.20	300,761.20	163,517.60	261,405.30	134,044.00	221,376.90	117,001.60	212,642.70	137,558.30	212,803.20	150,820.50	242,863.40	180,674.40	257,625.70	211,889.70
Time on Market	185.49	324.50	193.33	351.47	199.70	351.11	174.85	306.74	174.95	304.93	178.94	311.11	195.52	325.83	222.77	366.16
Size in cubic meters	363.32	311.31	431.47	190.28	390.09	658.54	353.60	182.35	333.41	178.84	321.26	171.51	358.37	216.18	400.85	295.09
Maintenance(1-9)	6.94	1.01	7.50	0.88	7.17	0.81	6.96	0.85	6.81	0.98	6.76	1.02	6.64	1.14	6.24	1.49
Number of rooms	4.45	1.47	4.61	1.45	4.47	1.43	4.48	1.37	4.31	1.45	4.25	1.47	4.50	1.60	4.67	1.81
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Monument	1,603	0.48	106	0.27	143	0.29	177	0.18	326	0.53	276	0.67	286	1.06	289	1.56
Type of Dwelling																
Apartment	103,973	31.06	10,743	27.61	15,101	30.71	26,470	26.89	22698	37.03	16,319	39.45	8,043	29.81	4,599	24.76
Mid-terraced	107,183	32.02	13,344	34.29	14,182	28.84	37,247	37.83	18810	30.69	13,028	31.49	7,525	27.89	3,047	16.41
Corner	41,660	12.45	3,879	9.97	4,852	9.87	13,844	14.06	8132	13.27	4,865	11.76	3,418	12.67	2,670	14.38
Semi-detached	45,956	13.73	5,588	14.36	8,984	18.27	13,469	13.68	6152	10.04	4,281	10.35	3,828	14.19	3,654	19.68
Detached	35,978	10.75	5,360	13.77	6,054	12.31	7,417	7.53	5507	8.98	2,876	6.95	4,163	15.43	4,601	24.78
Construction Period																
unknown	21	0.01	1	0.00	5	0.01	5	0.01	4	0.01	1	0.00	1	0.00	4	0.02
1500-1905	11,901	3.56	376	0.97	679	1.38	1,879	1.91	2579	4.21	1,932	4.67	2,163	8.02	2,293	12.35
1906-1930	31,162	9.31	523	1.34	1,130	2.30	3,703	3.76	7138	11.64	5,824	14.08	6,745	25.00	6,099	32.84
1931-1944	19,412	5.80	99	0.25	315	0.64	1,660	1.69	4285	6.99	4,351	10.52	4,693	17.40	4,009	21.59
1945-1959	29,229	8.73	123	0.32	745	1.52	4,608	4.68	7646	12.47	8,445	20.41	4,630	17.16	3,032	16.33
1960-1970	52,819	15.78	349	0.90	2,110	4.29	12,864	13.07	15918	25.97	13,082	31.62	5,984	22.18	2,512	13.53
1971-1980	57,948	17.31	617	1.59	4,483	9.12	26,864	27.29	15893	25.93	6,973	16.86	2,599	9.63	519	2.79
1981-1990	52,378	15.65	858	2.20	8,374	17.03	35,461	36.02	6897	11.25	649	1.57	91	0.34	48	0.26
1991-2000	43,931	13.12	8,484	21.80	23,974	48.75	10,620	10.79	742	1.21	59	0.14	27	0.10	25	0.13
¿ 2001	35,949	10.74	27,484	70.63	7,358	14.96	783	0.80	197	0.32	53	0.13	44	0.16	30	0.16

6 Results

6.1 The premium paid for green dwellings

The results from Model 4.1 are summarized[1] in Table 6.1 and show a strong relationship between the energy performance and the housing value. Several forms of the variable of interest are used and all of the outcomes are highly significant. Column 1 of Table 6.1 shows that a green dwelling has a significant positive effect on the transaction price, which indicates that people value the energy performance of a house. In particular a dwelling with label A, B or C transacts for a 3.1% higher price on average compared to dwellings with labels D, E, F or G, *ceteris paribus*. This result is comparable to the premia of other studies comparing green versus non-green, that were mentioned in Table 2.1 in Chapter 2 (Brounen & Kok, 2011; Cajias & Piazolo, 2012).

As is documented in Column 2, similar results are found when looking at the relationship separately per label and comparing them to (*'neutral'*) label D. All results are highly significant and as expected: labels E, F and G have an increasing negative effect (compared to label D) on the transaction price respectively. Conversely, labels A, B and C have a positive effect on the transaction price. The estimated effect of label A is smaller than label B, which might seem surprising. However, as their confidence intervals overlap the difference between the two labels is not significant. Hence, compared to dwellings with labels D, dwellings with labels A and B transact for a premium of around 3% and dwellings with labels C transact for a premium of 2% (*ceteris paribus*), whereas dwellings with labels E, F and G are sold for an average discount of 0.9%, 2% and 5.6% respectively (*ceteris paribus*). This is lower than what other studies found. Specifically, when comparing to Brounen & Kok (2011) the difference is substantial. However, as becomes clear from Table C.1 in Appendix C, leaving out some of the fixed effects crucially affects the magnitude of

[1]The complete results can be found in Table B.1 of Appendix B.

Table 6.1: 2008-2017 OLS regression results for model 1 - Energy performance proxy's on transaction prices (Dependent variable: logarithm of transaction price)

	(1)		(2)		(3)	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Green	0.031***	(0.002)				
(Label A, B or C)						
Energy Label[A]						
A			0.025***	(0.004)		
B			0.033***	(0.003)		
C			0.020***	(0.002)		
E			-0.009***	(0.002)		
F			-0.020***	(0.002)		
G			-0.056***	(0.003)		
Energy Index					-0.023***	(0.003)
Monument	0.076***	(0.012)	0.077***	(0.012)	0.059***	(0.020)
Size of dwelling	0.589***	(0.012)	0.588***	(0.012)	0.577***	(0.019)
(log of m3)						
Type of dwelling[B]						
Apartment	-0.121***	(0.005)	-0.121***	(0.005)	-0.117***	(0.006)
Corner	0.051***	(0.001)	0.053***	(0.001)	0.054***	(0.002)
Semi-detached	0.180***	(0.003)	0.182***	(0.003)	0.196***	(0.005)
Detached	0.402***	(0.005)	0.406***	(0.005)	0.454***	(0.009)
Extra controls for:						
Construction period	YES		YES		YES	
Maintenance	YES		YES		YES	
Number of rooms	YES		YES		YES	
Month FE	YES		YES		YES	
Neighbourhood FE	YES		YES		YES	
Constant	9.403***	(0.143)	9.438***	(0.143)	9.382***	(0.107)
N	334,750		334,750		123,524	
R-squared	0.853		0.853		0.857	
Adj. R-squared	0.851		0.852		0.853	

Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable. Due to the high number of binary variables, not all coefficients are included. *, **, ***: significant at 10%, 5% and 1% respectively.
[A] Reference group is label D
[B] Reference group is a mid-terraced dwelling

the premium. As Brounen & Kok (2011) do not control for neighbourhood fixed effects, this could explain the difference between the premia.

Column 3 shows the effect of the Energy Index, which has a slightly different interpretation due to its index form: the lower the index, the higher the energy efficiency of the dwelling. The coefficient in Column 3 is significant and negative, which is consistent with the results of Columns 1 and 2. The coefficient denotes a 2.3% increase of the transaction price if the Energy Index decreases by 1%.

Next, when assessing the control variables none of the coefficients are surprising. Since the coefficients are consistent among the columns, I only discuss the coefficients from Column 1. All of the included control variables in Table 6.1 are highly significant, which means they all have an effect on house values. As expected, monuments have a positive effect on the price and sell for a premium of almost 8%. The complete results, which can be found in table B.1 in Appendix B, show that the construction period of the dwelling is only significant for dwellings built between 1931 and 1970, when compared to dwellings built after 2000. Moreover, the maintenance grade has a strongly significant and positive effect on the price when compared to the lowest possible grade of 1. The only exception is a maintenance grade of 2 which is not significantly different. Regarding the size of the dwelling, we see that the coefficient of the logarithm of the dwelling's volume in cubic meters is positive and highly significant, whereas the number of rooms has a more inconsistent effect on the transaction price.

A very large part of the variation in the model is explained, namely around 85%. Indeed, this result is expected, since the objective of a hedonic price model is to explain the utmost valuation of a house through its characteristics. Additionally, Appendix C shows how the R-squared changes when the model is altered.

6.2 The impact of the new policy on the premium

Before turning to the results from Model 4.2, the regression from the first model was run separately for the period before and the period after the policy change. The results, which

Table 6.2: 2008-2017 OLS regression results for model 2 - Energy performance on transaction prices before and after policy change (Dependent variable: logarithm of transaction price)

	Model 2	
	Coefficient	Std. Error
Green (Label A, B or C)	0.012***	(0.004)
Green * After2015	0.025***	(0.004)
Monument	0.077***	(0.012)
Size of dwelling (log of m^3)	0.589***	(0.012)
Type of dwelling[A]		
Apartment	-0.121***	(0.005)
Corner	0.052***	(0.002)
Semi-detached	0.180***	(0.003)
Detached	0.403***	(0.005)
Extra controls for:		
Construction period	YES	
Maintenance	YES	
Number of rooms	YES	
Month FE	YES	
Neighbourhood FE	YES	
Constant	9.403***	(0.143)
N	334,750	
R-squared	0.853	
Adj. R-squared	0.851	

Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable. Due to the high number of binary variables, not all coefficients are included. *, **, ***: significant at 10%, 5% and 1% respectively.
[A]Reference group is a mid-terraced dwelling

can be found in table D.1 in Appendix D, give an indication on the effect of the policy change and allows to compare the premium found before and after 2015. The estimated effects are all positive and significant and are larger after 2015. The second model, defined in Chapter 4, is needed to verify whether these coefficients are significantly different from each other. The results can be found in Table 6.2.

The coefficient on the variable $Green$ can be interpreted as the effect of a green dwelling on the transaction price before the policy change. It is again highly significant and positive, but smaller. More precisely, before the policy change a green dwelling increased the price with an average of 1.2%, *ceteris paribus*. This indicates that the EPCs already solved some of the information asymmetry in the market and thus a premium was present. The coefficient on the interaction variable $Green * After2015$ is also significant and positive. According to these OLS estimates, the premium paid for a green dwelling significantly increased with another 2.5 percentage points on average after the change in policy, *ceteris paribus*. This adds up to approximately 3.7% premium for the period after 2015, which is fairly consistent with the premium found in Table D.1 when looking at the 95% confidence intervals.

Moreover, when examining the control variables little has changed compared to the basic regression model. The significance of all the coefficients stayed exactly the same. Any of the changes in the magnitude of the effects are negligible. The same is true for the R-squared of the model.

The results imply that the effect became significantly stronger after the policy-change. This increase of the premium shows how the effectiveness of the EPCs depend on their implementation. The change in policy substantially increased people's awareness and hence reduced the issue of imperfect information on energy-performance in the Dutch housing market. Naturally, it is not possible to conclude from the increase of the premium that the information problem is entirely solved. However, taking the possible future energy savings into consideration, allows us to give a rough indication. A recently published study showed that the savings on energy costs for green dwellings vary between 1,100 and 2,100 euros per year when compared to a label D (van Hoek & Koning, 2018). Discounting this to a present value[2], would imply a premium paid for green dwellings of at least 26,000. However, the premium found by this research of around 3%, indicates that a green dwelling transacts for a premium of approximately 7,000 euros compared to a non-green dwelling. Clearly, the energy-savings are not entirely capitalized into the premium that people are willing-to-pay. Although this is a very rough estimation, it does give an indication on whether the premium is enough. This means that the imperfect information problem does not yet seem to be entirely solved by the new policy.

[2]Assuming a lifespan of 25 years for the energy-saving investments and a conservative interest rate of 6%.

7 Robustness Checks

To check the robustness of the results that were found in Chapter 6, several alterations to the model and its variables are made. Overall, the results change only slightly, suggesting that the results are robust. The only exception to this is the inclusion of the maintenance factor of the dwelling. This chapter briefly discusses each robustness check.

Using the Energy Index instead of the EPCs (Column 3 of Table 6.1) has already shown that the results are robust for using a different proxy for the energy performance. Next, it could be argued that houses with lower transaction prices take longer to sell, due to a lower demand. Therefore, I use the time on the market as an alternative dependent variable (instead of the transaction price). As expected, the results in Table E.1 (Appendix E) show a negative and significant effect. A green dwelling sells on average 15 days faster than a non-green dwelling, *ceteris paribus*. This complies with the results found earlier.

In order to asses the control variables, several changes were made to the model and the subsequent effect on the results was observed. By plotting some of the variables on the transaction price the outliers of the data become visible (see Appendix F). Regarding the transaction price, I exclude observations that transacted for more than 5 million euros. Subsequently, dwellings with more than 30 rooms or a size of more than 5000 cubic meters are excluded. Next, I exclude the dwellings that were sold more than three times within the observation period. As can be seen from Figures F.1-F.4 in Appendix F, the fitted values do not seem to change significantly. In fact, reproducing the regression model without these outliers changed the results only slightly, implying that the results from the previous chapter were not driven by outliers.

Figure F.5 in Appendix F examines the relationship between the number of dwellings within the same 4-digit postcode area. In cases where there are fewer than 50 dwellings within a 4-digit postcode area, I switch to a 3-digit postcode control. This reduced the amount of locational dummies from 3745 to 2891. The results were robust for this alteration and did not change at all. Alternatively, I only include fixed effects on 3-digits postcode level, which reduced the number of locational fixed effects to 793. As can be seen in Table C.1 in Appendix C, this decreased the results slightly. Using only 500 municipal dummies instead, decreased the results further. However, the 95% confidence interval shows that there is no significant difference between the estimated effects.

Before discussing the maintenance factor, three other alterations of control variables are made to the model. Firstly, as it could be argued that house prices are affected more by the surface of the dwelling in square meters, whereas the energy-performance is more affected by the volume of the dwelling in cubic meters, the regressions are run on both. This changes none of the estimated coefficients. Secondly, instead of separating the number of rooms in categories, I used a discrete variable. Again, the effect incurs only a negligible change, while the significance stays on the 1% level. Lastly, the monthly fixed effects were replaced by yearly fixed effects (see Table C.1 in Appendix C). The results are again robust.

7.0.1 The maintenance factor

As was already mentioned a few times throughout this thesis, the quality of the dwelling has an influence on the value as well as the energy performance of houses. As expected, the results of this research are very sensitive to the way the maintenance variable is used in the model. Depending on the functional form of the variable, the model gives different results. Column 1 of table G.1 in Appendix G demonstrates that the coefficient for *Green* changes modestly when the maintenance factor is used as a discrete variable (instead of a categorical variable), which only allows for linear effects of the maintenance. The impact of a green dwelling on the transaction price increases from 3.1% to 3.2%, *ceteris paribus*. However, the 95% confidence intervals of both estimated effects overlap substantially. Column 2 reports the premium when exterior maintenance of the dwelling is used, instead of interior maintenance. This seems to substantially increase the effect to 3.5%. Again, the confidence intervals overlap. This exterior maintenance variable was used by Brounen & Kok (2011), which could explain the higher coefficients in their findings. Lastly, in Column 3 the maintenance factor is excluded from the model. Not controlling for maintenance drives up the premium to 4.7%, implying that maintenance and energy performance of dwellings are related. If the maintenance variable does not include any energy performance measures, this would indicate that a definite omitted variable bias is present when not controlling for maintenance. However, if the maintenance grade is also based on any factors that simultaneously determine the energy efficiency of the house[1], then including the variable captures some important information of the *Green* variable. In that case the premium from Chapter 6 can be seen as the lower bound of the real premium that people are willing to pay for a green dwelling.

[1] This could for example be the case if (some) realtors increase the maintenance grade when the house has double-glass windows.

8 Conclusions

This research has examined the premium that is paid for green dwellings. It looked at what the premium is and subsequently at how this premium changed due to a new policy that was initiated in 2015.

This thesis therefore provides an insight into how people value energy performance in the housing market. EPCs are supposed to inform house buyers on the energy performance of dwellings, which would lead to a price premium paid for green dwellings, due to the lower future energy costs. However, it greatly depends on how well these EPCs solve the issue of imperfect information. According to the European Commission the effectiveness of the EPCs depends on the implementation by the member states. Hence, investigating the premium for the Netherlands provides an useful insight into whether these EPCs are effective and whether the change in policy, which made the EPCs mandatory from 2015 onwards, has increased their effectiveness.

The results of this paper are affirmative. The premium that house buyers are willing to pay for a green dwelling is estimated at an average of 3.1%, *ceteris paribus*. When looking at how this has changed due to the policy, there is clearly a significant increase. Before the policy change the premium was only 1.2%, whereas after the policy change this increased significantly with another 2.5 percentage points. Therefore, I conclude that energy performance is valued in the Dutch housing market and that the implementation of EPCs have been of help with creating a more efficient price for energy efficient houses.

These findings are of specific interest for policy makers. Since this research showed that the way the EPCs are implemented by the government influences their effectiveness, it could be argued that it is the role of the government to correctly solve the information asymmetry and create an efficient premium in the housing market. Moreover, comparing the premium with the average energy savings of green dwellings has shown that these savings are not yet fully capitalized. It is therefore necessary for policy-makers to decide on what the desired premium is and how this premium is to be achieved. It might then be needed to increase the stringency of the EPC policy. More specific data on individual energy consumption could help to correctly determine what the efficient premium is. As this could advise policy-makers on whether the current policies are enough or whether stricter policies are needed, further research should address this.

Additionally, it would be economically interesting for further research to examine what the underlying neighbourhood fixed effects of this research are. However, it would be necessary to obtain a very extensive dataset with neighbourhood specific characteristics on 4-digits post code level. Running the regression model separately across different neighbourhoods would give an indication on why and when people value energy performance. This could be of much interest for policy-makers. Moreover, this would make it possible to control for locational effects that vary over time. As in this model, the time fixed effects vary only over time, but not across neighbourhoods and the neighbourhood fixed effects vary only across neighbourhoods, but not over time.

Finally, it seems likely that other external reasons or policies influencing the supply-side of the housing market have simultaneously increased the relative supply of green dwellings during the observation period. As this thesis uses a hedonic price model, which uses the interaction between supply and demand to determine the valuation of dwellings, the supply and demand effects cannot be separated. Therefore, it is possible that the premium does not only reflect the effect of the EPCs, but should thus be seen as a lower bound. It would be interesting for further research to isolate the demand effects from these supply effects.

References

Akerlof, G. A. (1970). The market for “lemons”: Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488–500.

Angrist, J. D. & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bloom, B., Nobe, M. & Nobe, M. (2011). Valuing green home designs: A study of energy star® homes. *Journal of Sustainable Real Estate*, *3*(1), 109–126.

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.

Cajias, M. & Piazolo, D. (2012). Green performs better: Energy efficiency and financial return on buildings. *Journal of Corporate Real Estate*, *15*(1), 53–72.

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

Chay, K. Y. & Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, *113*(2), 376–424.

Chegut, A., Eichholtz, P. & Kok, N. (2014). Supply, demand and the value of green buildings. *Urban Studies*, *51*(1), 22–43.

Dijk, M., Groot, S. & Möhlmann, J. (2016). CPB Woningmarktmodel [Achtergronddocument]. *Centraal Planbureau.*

Dröes, M. I. & Koster, H. R. (2016). Renewable energy and negative externalities: The effect of wind turbines on house prices. *Journal of Urban Economics*, *96*, 121–141.

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

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

European-Commission. (2018a). *Buildings.* Retrieved 27-02-18, from `https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings`

European-Commission. (2018b). *Nearly zero-energy buildings.* Retrieved 14-05-18, from `https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings/nearly-zero-energy-buildings`

Francke, M. K. (2014). Hedonic price models. *Course reader in the topic real estate valuation and investments. University of Amsterdam (UvA) Business School*, *94*.

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.

Hyland, M., Lyons, R. C. & Lyons, S. (2013). The value of domestic building energy efficiency evidence from Ireland. *Energy Economics*, *40*, 943–952.

Jaffe, A. B. & Stavins, R. N. (1994). The energy-efficiency gap: What does it mean? *Energy Policy*, *22*(10), 804–810.

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

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. *Housing Economics and Public Policy*, 67–89.

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. *Paris: European Commission (DG Energy).*

O'Sullivan, A. & Gibb, K. (2008). *Housing economics and public policy.* John Wiley & Sons.

Palmquist, R. B. (2005). Property value models. *Handbook of Environmental Economics*, *2*, 763–819.

Pigou, A. C. (1920). *The economics of welfare.* London, Palgrave MacMillan.

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

Salvi, M., Horehájová, A. & Syz, J. (2010). *Are households willing to pay for green buildings? An empirical study of the Swiss minergie label* (Tech. Rep.). European Real Estate Society (ERES).

van Dijk, D. W. & Francke, M. K. (2008). Internet search behavior, liquidity and prices in the housing market. *Real Estate Economics*, *46*(2), 368–403.

van Dijk, D. W., Geltner, D. & Minne, A. (2017). Revisiting supply and demand indices in real estate. *Seminar of De Nederlandsche Bank*.

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

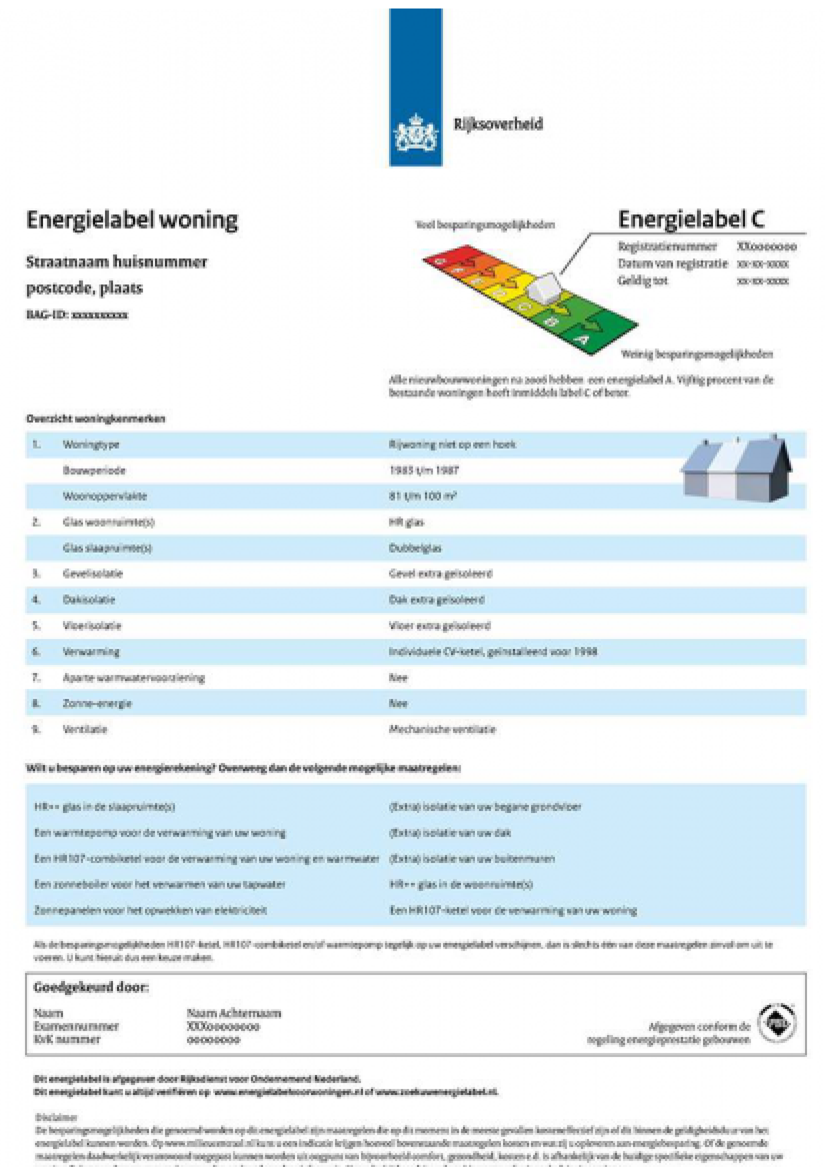
van Eck, H. (2015). *Implementation of the EPBD in the Netherlands: Status november 2015.* Rijksdienst voor Ondernemend Nederland.

van Ekerschot, F. & Heinemans, M. (2008). *Implementation of the EPBD in the Netherlands: Status and planning in june 2008.* Rijksdienst voor Ondernemend Nederland.

Appendices

A The Dutch Energy Label

Figure A.1: Example of a Dutch energy-label



Rijksoverheid

Energielabel woning

Straatnaam huisnummer
postcode, plaats
BAG-ID: xxxxxxxxxx

Veel besparingsmogelijkheden

Energielabel C

Registratienummer	XXoooooooo
Datum van registratie	xx-xx-xxxx
Geldig tot	xx-xx-xxxx

Weinig besparingsmogelijkheden

Alle nieuwbouwwoningen na 2006 hebben een energielabel A. Vijftig procent van de bestaande woningen heeft inmiddels label C of beter.

Overzicht woningkenmerken

1.	Woningtype	Rijwoning niet op een hoek
	Bouwperiode	1985 t/m 1987
	Woonoppervlakte	81 t/m 100 m²
2.	Glas woonruimte(s)	HR glas
	Glas slaapruimte(s)	Dubbelglas
3.	Gevelisolatie	Gevel extra geïsoleerd
4.	Dakisolatie	Dak extra geïsoleerd
5.	Vloerisolatie	Vloer extra geïsoleerd
6.	Verwarming	Individuele CV-ketel, geïnstalleerd voor 1998
7.	Aparte warmwatervoorziening	Nee
8.	Zonne-energie	Nee
9.	Ventilatie	Mechanische ventilatie

Wilt u besparen op uw energierekening? Overweeg dan de volgende mogelijke maatregelen:

HR++ glas in de slaapruimte(s)	(Extra) isolatie van uw begane grondvloer
Een warmtepomp voor de verwarming van uw woning	(Extra) isolatie van uw dak
Een HR107-combiketel voor de verwarming van uw woning en warmwater	(Extra) isolatie van uw buitenmuren
Een zonneboiler voor het verwarmen van uw tapwater	HR++ glas in de woonruimte(s)
Zonnepanelen voor het opwekken van elektriciteit	Een HR107-ketel voor de verwarming van uw woning

Als de besparingsmogelijkheden HR107-ketel, HR107-combiketel en/of warmtepomp tegelijk op uw energielabel verschijnen, dan is slechts één van deze maatregelen zinvol om uit te voeren. U kunt hieruit dus een keuze maken.

Goedgekeurd door:

Naam	Naam Achternaam
Examennummer	XXXoooooooo
KvK nummer	oooooooo

Afgegeven conform de regeling energieprestatie gebouwen

Dit energielabel is afgegeven door Rijksdienst voor Ondernemend Nederland.
Dit energielabel kunt u altijd verifiëren op www.energielabelvoorwoningen.nl of www.zoekuwenergielabel.nl.

Disclaimer

Figure A.2: Dutch info-graphic explaining the energy-label

Figure A.3: Dutch info-graphic explaining how to get an energy-label after the policy change

B The Complete Results

Table B.1: Complete results of the basic regression including all control variables

	ln (Transaction Price)	
	Coefficient	Std. Error
Green	0.031***	(0.002)
(Label A, B or C)		
Monument	0.076***	(0.012)
Size of dwelling	0.589***	(0.012)
(log of m^3)		
Type of dwelling[A]		
Apartment	-0.121***	(0.005)
Corner	0.051***	(0.001)
Semi-detached	0.180***	(0.003)
Detached	0.402***	(0.005)
Construction period[B]		
unknown	-0.137	(0.128)
1500-1905	-0.184	(0.128)
1906-1930	-0.154	(0.128)
1931-1944	-0.242*	(0.128)
1945-1959	-0.270**	(0.128)
1960-1970	-0.237*	(0.128)
1971-1980	-0.202	(0.128)
1981-1990	-0.098	(0.128)
1991-2000	-0.040	(0.128)

(to be continued on next page)

Maintenance[C]		
2	0.016	(0.027)
3	0.077***	(0.017)
4	0.114***	(0.018)
5	0.148***	(0.017)
6	0.180***	(0.017)
7	0.266***	(0.017)
8	0.335***	(0.017)
9	0.368***	(0.017)
Number of Rooms[D]		
2	-0.054***	(0.014)
3	0.011	(0.012)
4	0.032***	(0.012)
5	0.050***	(0.012)
6	0.083***	(0.012)
6-10	0.131***	(0.012)
10-19	0.204***	(0.018)
> 19	0.073	(0.103)
Month FE	YES	
Neighbourhood FE	YES	
Constant	9.403***	(0.143)
N	334,750	
R-squared	0.853	
Adj. R-squared	0.851	

Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable.
*, **, ***: significant at 10%, 5% and 1% respectively.
[A]Reference group is a mid-terraced dwelling
[B]Reference group are dwellings built > 2000
[C]Reference group is maintenance grade 1
[D]Reference group is dwelling with 1 room

C The Fixed Effects

Table C.1: OLS regression results for variable of interest and R-squared for modifications of the model

		Green variable		
		Coefficient	Std. Error	R-squared
(1)	Basic model:			
	Green	0.123***	(0.009)	0.016
(2)	Controls:			
	Green + X	0.051***	(0.005)	0.555
(3)	Time fixed effects:			
	Green + X + year-dummies	0.041***	(0.005)	0.574
	Green + X + month-dummies	0.041***	(0.005)	0.575
(4)	Locational fixed effects:			
	Green + X + month-dummies + municipal-dummies	0.026***	(0.003)	0.788
	Green + X + month-dummies + postcode-3-dummies	0.030***	(0.002)	0.825
	Green + X + month-dummies + postcode-4-dummies	0.031***	(0.002)	0.853

Notes: This table reports the results for the variable Green and the R-squared when the model is gradually expanded. The first model regresses only the variable of interest (Green) on the dependent variable (log of price). Second model adds physical characteristics as control variables. Third model adds time fixed effects. Fourth model adds locational fixed effects. ***: significant at 1%.

D The Premium Before and After the Policy Change

Table D.1: 2008-2017 OLS regression results - Comparing the effect before and after policy change (Dependent variable: logarithm of transaction price)

	Total Period		Before 2015		After 2015	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Green	0.031***	(0.002)	0.013***	(0.003)	0.042***	(0.002)
(Label A, B or C)						
Monument	0.076***	(0.012)	0.082***	(0.023)	0.076***	(0.010)
Size of dwelling	0.589***	(0.012)	0.546***	(0.025)	0.599***	(0.013)
(log of m$\hat{3}$)						
Type of dwelling[A]						
Apartment	-0.121***	(0.005)	-0.107***	(0.008)	-0.124***	(0.005)
Corner	0.051***	(0.001)	0.052***	(0.003)	0.051***	(0.002)
Semi-detached	0.180***	(0.003)	0.202***	(0.006)	0.177***	(0.003)
Detached	0.402***	(0.005)	0.470***	(0.011)	0.397***	(0.005)
Extra controls for:						
Construction period	YES		YES		YES	
Maintenance	YES		YES		YES	
Number of rooms	YES		YES		YES	
Month FE	YES		YES		YES	
Neighbourhood FE	YES		YES		YES	
Constant	9.403***	(0.143)	9.391***	(0.127)	9.252***	(0.182)
N	334,750		69,360		265,390	
R-squared	0.853		0.863		0.853	
Adj. R-squared	0.851		0.857		0.851	

Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable. Due to the high number of binary variables, not all coefficients are included. *, **, ***: significant at 10%, 5% and 1% respectively.
[A] Reference group is a mid-terraced dwelling

E *Time on Market*

Table E.1: 2008-2017 OLS regression results with time on market in days as dependent variable

	Time on market	
	Coefficient	Std. Error
Green	-14.992***	(1.622)
(Label A, B or C)		
Monument	8.003	(8.553)
Size of dwelling	60.439***	(3.255)
(log of m^3)		
Type of dwelling[A]		
Apartment	73.473***	(2.886)
Corner	0.393	(1.539)
Semi-detached	30.914***	(2.262)
Detached	157.285***	(3.550)
Extra controls for:		
Construction period	YES	
Maintenance	YES	
Number of rooms	YES	
Month FE	YES	
Neighbourhood FE	YES	
Constant	-404.145***	(34.102)
N	334,750	
R-squared	0.853	
Adj. R-squared	0.851	

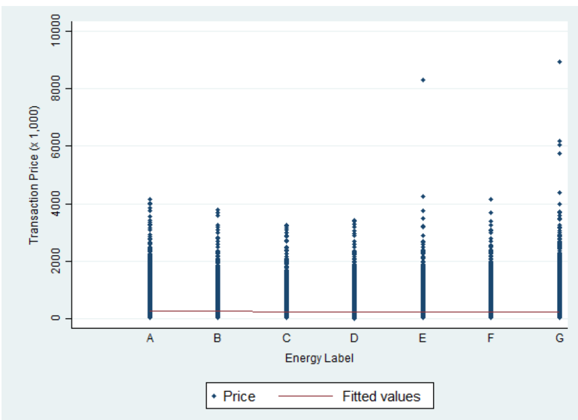
Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable. Standard errors are clustered per 4-digit post-code.
*, **, ***: significant at 10%, 5% and 1% respectively.
[A]Reference group is a mid-terraced dwelling

F Scatter Plots

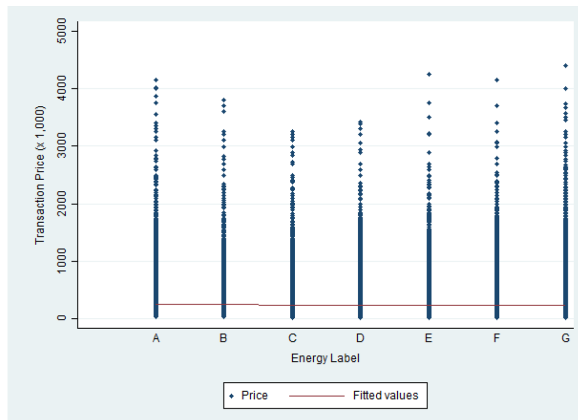
Scatter Plots: looking for outliers

Transaction Price

Figure F.1: Scatter Plot between Transaction Price and Energy Labels



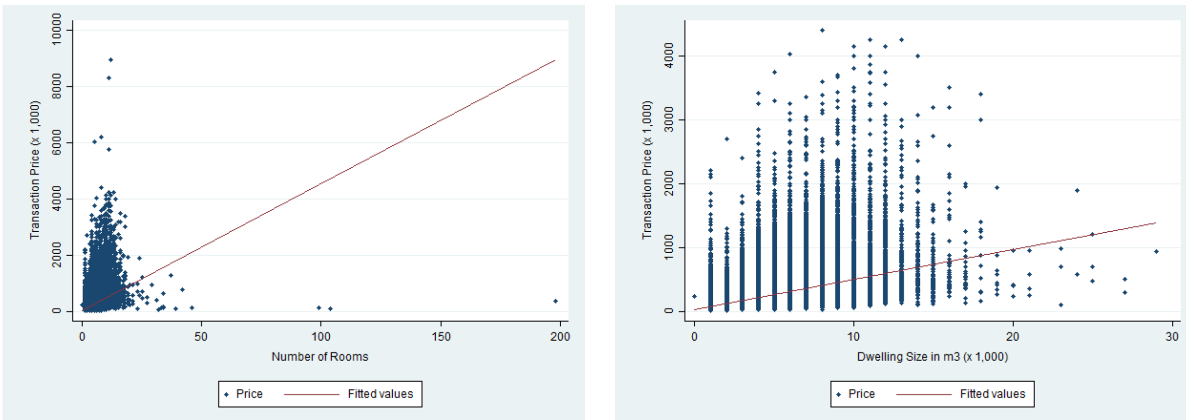
(a) Total sample



(b) Exclude outliers where the transaction price exceeds 5 mln euro's

Number of Rooms

Figure F.2: Scatter Plot between Transaction Price and Number of Rooms

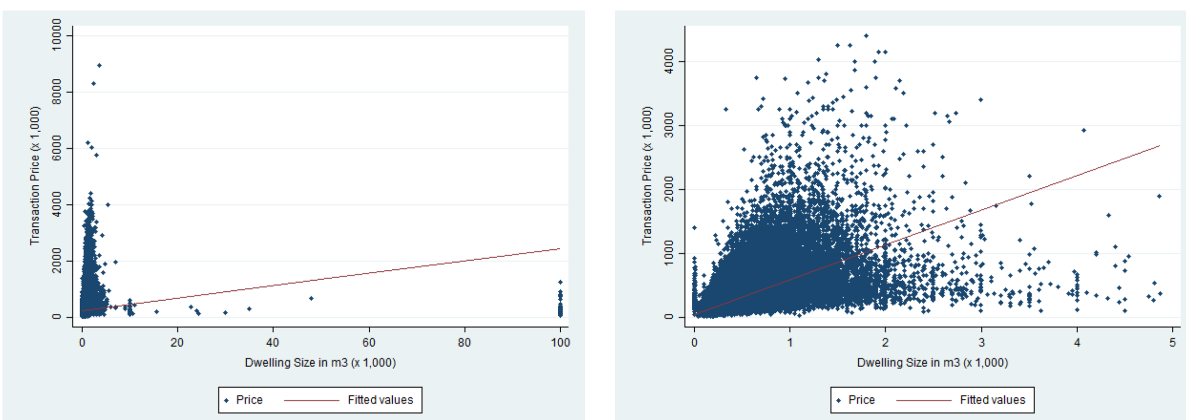


(a) Total sample

(b) Exclude outliers where transaction price exceeds 5 mln euro's and number of rooms exceed 30

Size of Dwelling

Figure F.3: Scatter Plot between Transaction Price and Size of Dwelling

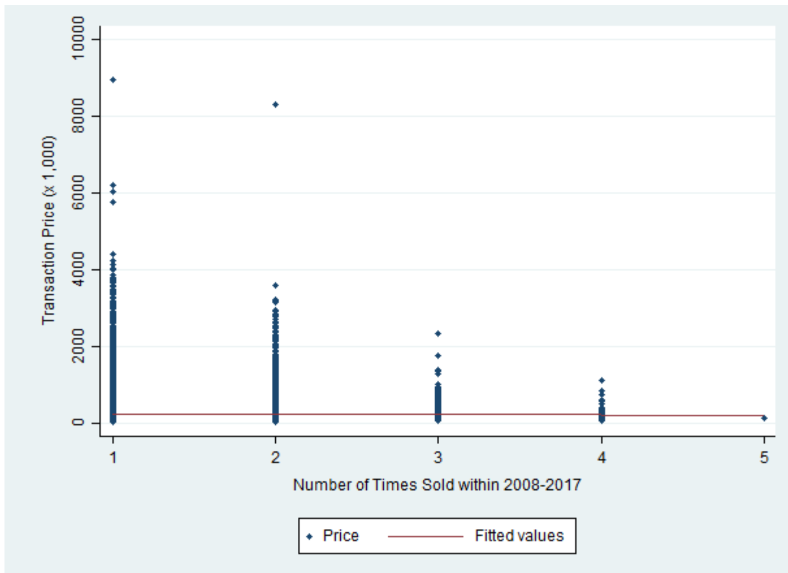


(a) Total sample

(b) Exclude outliers where transaction price exceeds 5 mln euro's and size of dwelling exceeds 5000 m^3

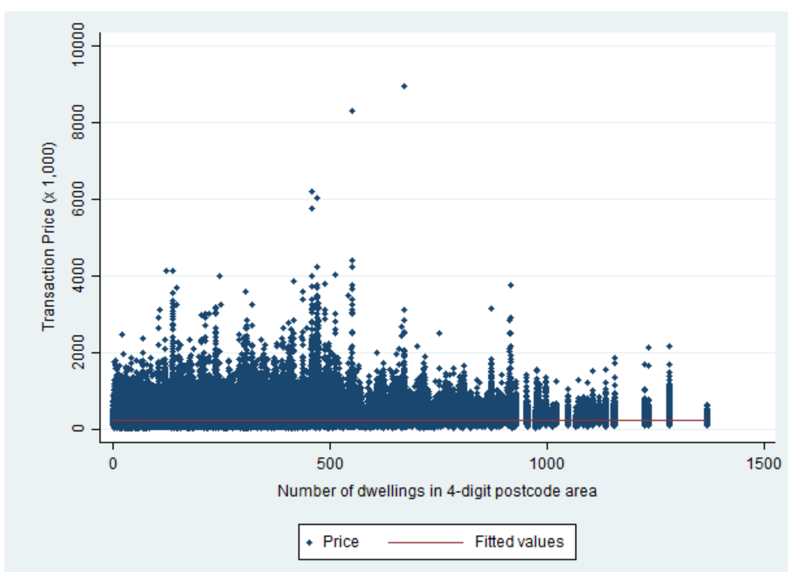
Repeated sales

Figure F.4: Scatter Plot between Transaction Price and Number of Times Dwelling was Sold



4-digits Post-code

Figure F.5: Scatter Plot between Transaction Price and number of dwellings within the same 4-digits post-code area



G The Maintenance Factor

Table G.1: Robustness check: controls for maintenance factor

	(1)		(2)		(3)	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Green (Label A, B or C)	0.032***	(0.002)	0.035***	(0.002)	0.047***	(0.002)
Maintenance[A]	0.054***	(0.001)				
Exterior Maintenance[B]						
2			0.029	(0.039)		
3			0.132***	(0.027)		
4			0.171***	(0.028)		
5			0.218***	(0.026)		
6			0.260***	(0.026)		
7			0.340***	(0.026)		
8			0.409***	(0.026)		
9			0.436***	(0.026)		
Monument	0.076***	(0.012)	0.071***	(0.012)	0.078***	(0.011)
Size of dwelling (log of m3)	0.590***	(0.012)	0.596***	(0.012)	0.611***	(0.012)
Extra controls for:						
Construction period	YES		YES		YES	
Type of dwelling	YES		YES		YES	
Number of rooms	YES		YES		YES	
Month FE	YES		YES		YES	
Neighbourhood FE	YES		YES		YES	
Constant	9.285***	(0.142)	9.230***	(0.146)	9.558***	(0.148)
N	334,750		334,750		334,750	
R-squared	0.852		0.849		0.841	
Adj. R-squared	0.851		0.847		0.840	

Notes: This table reports the regression results with the logarithm of transaction price as the dependent variable. Due to the high number of binary variables, not all coefficients are included. Standard errors are clustered per 4-digit post-code. *, **, ***: significant at 10%, 5% and 1% respectively.
[A] Discrete variable ranging from 1-9.
[B] As a categorical variable: reference is 1.