University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme: Industrial Economics	Spring 2020 Open
Authors: Edris Afzali and Daniel Torvaldsen Garip	
Internal supervisor: Peter Molnár External supervisor: Aslak Wigdahl Bergersen	
Title of master's thesis: "Impact of energy efficiency and temperature on house prices in Norway"	
Credits: 30	
Keywords: • House prices • Energy Performance Certificates • Energy efficiency • Climate and weather • Hedonic regression	Number of pages: 62 + supplemental material/other: 0 Stavanger, 15/6-2020 date/year

Preface

This thesis is the final work of our Master's degree in *Industrial Economics*. The thesis consists of two separate papers, where both have their own abstract, content-list, chapters and reference-list. We first present paper one: **"Impact of energy efficiency on house prices and the role of temperature"** and then paper two: **"Impact of temperature on house prices in Norway"**. The first paper has been our main focus, and is the more comprehensive work of the two papers. Writing the first paper gave us access to large and well-detailed datasets and as a result, we decided to write the second paper to utilize the data even further. While we are proud and have put a lot of effort in to both papers, paper one is more finalized compared to paper two.

Acknowledgements

We would like to direct our gratitude towards our supervisors: professor Peter Molnár at the University of Stavanger and data scientist Aslak Wigdahl Bergersen from *Alva Technologies*. Peter was always available for questions and guidance and was eager to help us with anything. Aslak boosted our thesis with valuable ideas and insights and helped us through this long process.

Additionally, we would like to thank all the people at Alva Technologies for giving us access to their well organized and highly valuable dataset. At last, we would like to acknowledge the help from *ENOVA SF* and representatives from *Multiconsult*, for valuable information and providing us with data about the Energy Performance Certificate scheme in Norway.

We have learned a lot through the writing of this thesis and with the guidance and ideas of all the collaborators, we have obtained new skills and knowledge that we will have for the rest of our lives. Thank you all for the numerous e-mails and interesting digital meetings!

Impact of energy efficiency on house prices and the role of temperature

(Paper one)

June 15, 2020

Abstract

Energy Performance Certificate, a measure of a dwelling's energy efficiency, is a legislatively approved instrument for reducing energy consumption in households. We empirically investigate the impact of energy efficiency on house prices and how the impact vary with the type of dwelling and in disparate climate zones using average temperature. To accomplish this, we utilize Norwegian microdata, through a large and detailed dataset covering 222 392 dwellings with energy ratings. We find that there are increasing price premiums with increasing energy efficiency, and the effect is more significant for detached houses compared to apartments. Furthermore, the energy efficiency of buildings is valued more for houses in cold areas.

Contents

List of Figures

List of Tables

Abbreviations

EPC	**E**nergy **P**erformance **C**ertificate
HDD	**H**eating **D**egree **D**ay

Chapter 1

Introduction

Over the past decades, human activity has caused an increase in the release of greenhouse gases (Skripnuk & Samylovskaya, 2018; Shine & Forster, 1999). One major contribution is emissions caused by energy usage in the building sector and especially within the heating of dwellings. Since 2010, building-related CO2 emissions have increased by 1% annually. Moreover, in 2015, 82% of total energy consumption in buildings was supplied by fossil fuels (Environment & Agency, 2017). In addition, 40% of total energy consumption in Europe stems from the building sector (EU, 2002). Improvements in this sector will be a major step towards achieving a reduced carbon footprint and improved environmental performance by reducing the release of greenhouse gases.

The growing concern regarding man-made climate change has resulted in increasing policy focus on reducing greenhouse gas emissions. This has rooted mitigation strategies in many markets, including real estate, through the retrofitting of buildings to achieve more energy-efficient homes. The introduction of mandatory Energy Performance Certificates (EPC) for all dwellings is one of these strategies[1], as it is intended to provide buyers with reliable information regarding the energy performance of dwellings. The desired result is increased awareness for energy performance of buildings and reducing the energy consumption, through affecting the decision-making process of home-buyers and allowing existing owners to calculate for potential savings from investing in energy efficiency improvements and contribute to the modern green energy society.

Since the launch of EPC, researchers have investigated whether the use of such certificates has had any significant impact on the real estate market (Bio Intelligence Service & IEEP, 2013). However, there is no clear consensus in the literature, and previous studies have catered to a variety of conclusions. A well-known and one of the first studies in the field of EPC and economy, Brounen & Kok (2011), empirically analyzed the residential real estate in the Netherlands. They concluded that there exists a positive

[1]The framework was introduced by the Energy Performance of Buildings Directive (directive 2002/91/EC), which is the main policy instrument to promote the energy efficiency of buildings in the European Union

price premium for dwellings with higher energy ratings such as A (10%), B (5.5%) and C (2%). Using other geographical regions, other studies confirm the positive correlation between EPC and house price, however, with varying strength (Cerin et al., 2014; Fuerst et al., 2015, 2016; Bisello et al., 2020; de Ayala et al., 2016). Similar pattern was found even for rental prices in Ireland (Hyland et al., 2013). Furthermore, Fleckinger et al. (2019) theoretically presented the short term effect of the EPC as being a tool to reduce energy use and increase willingness to invest in energy efficiency.

Although there is empirical and theoretical evidence indicating a correlation between EPCs and house prices, a survey-based study in the Netherlands found that only 10% take EPC classification into account when buying a dwelling (Murphy, 2014). Likewise, Lainé (2014) and Amecke (2012) drew similar survey-based conclusions for the UK and Germany, respectively. This might indicate that the reported correlations are driven by uninvestigated confounders, or that the phrasing of the questioning hindered to capture the true effect of EPCs on the decision-making process.

In the context of Norway, findings from the Swedish real estate market is particularly interesting for our study, mainly because it is a country comparable in size, infrastructure, and geographical location. Wahlström (2016) found no price premium related to energy consumption in dwellings, following the same path, Hårsman et al. (2016) found no price premium related to the EPCs. In contradiction, Cerin et al. (2014) investigated the impact of energy ratings in the Swedish housing market, and identified a positive price premium for higher ratings. Additionally, Wilhelmsson (2019) found evidence of EPCs being appreciated to a greater extent in the northern and colder parts of Sweden.

The intersection of energy efficiency and house prices in disparate climate zones is rarely discussed in the existing literature. In fact, we find few studies focusing on this topic; Dell'Anna et al. (2019) found evidence of EPCs being appreciated differently for Barcelona and Turin, and the differences in climate. In California, Kahn & Kok (2014) concluded that houses labeled "green" have higher price premiums in warmer areas with a greater energy need for cooling.

To the authors' best of knowledge, there has to date only been two studies conducted with Norway as a geographical area of interest, but limited to the capital, Oslo. Utilizing transaction data of dwellings, Olaussen et al. (2017) were able to perform a quasi-natural experiment by applying a hedonic price model in the periods pre- and post-implementation of the EPCs in Norway. The results revealed price premiums being present even before the implementation of the EPC, and thereby argue that the price premium found are not associated directly with the implementation of EPCs. Building on the previous study, they further analyzed the results while considering energy price, however, their initial conclusion remains (Olaussen et al., 2019).

Another limitation of Olaussen et al. (2017, 2019), is that the limited geographical location did not allow for meteorological considerations. Like Sweden, the Norwegian climate is relatively cold, and the energy efficiency of buildings is likely of greater importance than in countries with a warmer climate. Climatic conditions also vary within Norway; not only is northern Norway much colder than the southern Norway, but areas further from the ocean have a significantly different climate compared to the coastal areas. Therefore, Norway is a perfect country to study with regard to this topic.

To overcome the limitation of previous studies, we accessed open weather data from The Norwegian Meteorological Institute[2] and established corporate collaborations (*Alva Technologies* and *ENOVA SF*) to compose a complete dataset of all sales in Norway over the last ten years with energy labeling and historical meteorological information. Energy consumption is highly correlated with temperature regulation of buildings, and it is therefore of interest how house prices vary with climate and energy efficiency. Our large dataset allows us to investigate the impact of climate and energy efficiency separately for each building type. Which ultimately can be used to evaluate the current use of EPC as a measure for energy efficiency and to cater to the solution of reducing the energy usage in the building sector.

We find that in Norway, houses with higher energy efficiency are sold at higher prices. Our estimates show a clear increasing price premium as we move from the lowest rating (G) to the highest (A). This relationship is stronger in colder regions where a higher degree of heating is needed throughout the year. Moreover, our results indicate that energy efficiency does not matter equally for all building types. We find that the price of detached dwellings is significantly more dependent on energy efficiency, compared to apartments.

Part one of the thesis proceeds by first providing an explanation of the EPC scheme in Norway in Section 2. Section 3 presents an exploratory analysis of the data. Section 4 presents the methods used and discusses the results. Section 5 concludes.

[2] https://frost.met.no

Chapter 2

Background

In Norway, the EPCs were first introduced in 2010 and implemented as mandatory for all new buildings, private dwellings for sale, and all commercial buildings exceeding 1000 square meters. The Norwegian EPC is composed of two independent components:

- **The energy rating**, which ranges between the characters A and G in a letter-scale. The scale is defined such that character A indicates the highest energy efficiency, while G indicates worst. The rating is purely based on estimated delivered energy per square meter.
- **The heating rating**, which is given in the form of a color from a scale (shown in Figure 3.3) ranging from red to green, and ranks the building according to the heating system installed, i.e., the percentage of heating covered by electricity, oil, gas, etc. The green color indicates the best heating rating (1), while red indicates the worst (5) (ENOVA SF, 2019).

Moreover, there are two main types of certificates in Norway. The type of certificate issued is determined by the level of dwelling-details attached in the application.

- **Simple:** A certificate only containing the basic characteristics of a dwelling, manually inserted by the applicant. The rest of the dwelling-attributes are automated standardized values for the dwelling-type.
- **Advanced:** Certificates containing a more extensive level of details about a dwelling, ranging from the shape of each floor and walls to roof structure with a higher degree of technical information.

When applying for a certificate, the applicants can choose to do it themselves or, alternatively, have a qualified energy advisor who meets the competency requirements to aid in the process. A qualified expert is an individual who holds some form of relevant education, usually a university degree, and is approved by an accreditation agency

(Arcipowska et al., 2015). A certificate issued by an expert often contains additional information retrieved from an advanced external computational tool and the EPC will include precise computational data used to categorize the dwelling.

The vast majority of certificates issued the last decade is of the type *simple*, typically carried out through self-assessment by home-owners. Norway is among the very few countries that allow a self-assessment option. However, this option remains only for owners of existing buildings.

Originally when the EPC was introduced, the energy rating scale was designed with other energy requirements, compared to recent time. The latest change was in June 2015 where the energy requirements were tightened for obtaining a high rating(A/B), and buildings constructed according to the minimum criteria listed in *Regulations on technical requirements for building works TEK10*, published by *Direktoratet for byggkvalitet*, was set as the limit for obtaining a rating C. However, the follow-up regulations implemented in the new building regulation *TEK17* introduced in 2017 with even tighter energy requirements, are yet to be implemented in the energy rating scale.

According to ENOVA SF's official statistics[1], a total of approximately one million certificates have been issued in Norway, indicating EPC coverage of approximately 38% of all dwellings[2]. Furthermore, there is evidence indicating a lack of attention towards the EPC among home-buyers and real-estate agents in Norway (ENOVA SF, 2019).

Potential explanations to the lack of attention towards EPC are 1) home-owners may have not felt the need of improving energy features because of limited knowledge about EPC, 2) the current building legislation has made it hard for older buildings to achieve a good grade without major renovations, and 3) it is also believed that the current certificate composed of two different ratings is too complex (ENOVA SF, 2019).

Recognizing these flaws in the current system in Norway, it is evident the system is yet to reach its full potential. As a consequence, several instances led by *ENOVA SF*, has called for a change in the energy labeling scheme (*Energimerkeordningen*) and is currently working on a more accurate and simplified energy labeling scheme. Expected modifications include a merged total rating instead of two different ratings, enhanced design, and a more sophisticated calculation formula with incorporating the effect of electric power (ENOVA SF, 2019).

Not only is a renewal of the EPC scheme in Norway in development, but recently the *Climate bonds standard and certification scheme*[3] (Technical Expert Group, 2019a) was introduced. Briefly explained, this a scheme for further increasing the reduction of

[1]As of today, the state-owned enterprise *Enova SF*, is responsible for issuance and maintenance of the EPC program in Norway.

[2]There are approximately 2.6 million dwellings in Norway. Assuming a 1:1 relationship between EPC and dwelling, correspond to a coverage of 38%. However, as a dwelling may hold several certificates, the actual coverage may be significantly lower.

[3]`https://www.climatebonds.net/certification`

greenhouse gases in the building sector, while utilizing the EPC scheme to issue "green" bonds and loans for energy-efficient buildings. In the real-estate market, the scheme has introduced "green" loans with a lower mortgage rate when buying dwellings with energy ratings A and B (Technical Expert Group, 2019b), or if investing in energy improvements in new or existing dwellings.

Chapter 3

Data

3.1 Datasets

The data was acquired by affiliating with two independent companies, *Alva Technologies* and *Enova SF*. *Alva Technologies* provided us with two large and well-detailed datasets containing transaction history for dwellings in Norway and several characteristics of each dwelling. Both datasets contain an internal identifier with an ID linking each entry to a dwelling.

Enova SF provided us with a dataset containing information regarding energy characteristics and performance certificates for dwellings. The dataset contains all energy performance certificates which have been quality assured and verified in the land register, thereby not containing all certificates issued in total in Norway. Additionally, a single dwelling may have several certificates, and duplicates occur. Some of the entries also contain values regarding the energy consumption of the dwelling, however, only if manually inserted by users when applying for an EPC.

In addition, we utilize freely available meteorological data for Norway. We use the MET Frost API developed by the Meteorological Institute of Norway, and extract meteorological data for each specific dwelling, in particular temperature data, which was transformed into Heating Degree Days (HDD). An overview of the datasets is displayed in Table 3.1.

Table 3.1: Overview of the Datasets

Dataset	**Type**	**size(n)**	**Source**	**Time-span**
Dwelling characteristics	CSV	2 572 317	*Alva Technologies*	
Transactions history	CSV	2 792 731	*Alva Technologies*	Jan 1991-Dec 2019
EPC	CSV	420 975	*Enova SF*	Dec 2009-Dec 2019
Meteorological data	API		*MET Frost API*	

3.2 Data Pre-Processing

3.2.1 Data Merging

The first step is to merge the datasets and have a combined dataset of all the above. The purpose is to obtain a historical dataset containing all transactions with the respective dwelling-characteristics, HDD data, and the valid energy certificate issued before the dwelling was sold. The merge process is as follows: i) First we match the entries in **Dwelling characteristics data** to the corresponding transactions in **Transactions history data**, using a common identifier *address id* located in both datasets. ii) Attach the EPC-entry in **EPC dataset** with matching land register-/building number and for some dwellings the residential unit number, to the corresponding transaction. iii) Finally, adding the wanted temperature data as HDD for each location of all the dwellings.

Step ii) is a vital step, where we match the two main components in our research, transactions and EPC. Due to the reasons and limitations discussed in Section 3.3.2, data entries are matched differently. The different merging criteria are dependent on the building type and in cases where the residential number is lacking. As a result, we classify the merging in three different classification types. Each entry obtains one of the three following classifications: **A** - Dwellings with matching land register, building and residential unit number, **B** - Detached/semi-detached dwellings with matching land register and building number, and **C** - Apartments with matching land register and building number, but with missing residential number. Additional condition in class C is the size of the dwelling comparatively equal to the size-specification in the EPC. Whereby the classification-types is ranked according to the possibility for a merge-error to occur, where A has the lowest probability and C has the highest.

Table 3.2: Overview of merged dataset

Classification	Merge Condition	% of Total Data
A	LR[1]/Building/residential number	35.0%
B	LR/Building number	45.5%
C	LR/Building number and size	19.4%

3.2.2 Data Cleaning

The merged dataset consists of multiple entries of each transaction with different energy certificate values. Thereby, the next step includes connecting the correct EPC and the removal of duplicate and extraneous transaction-entries. Which in our case is performed by first exclude all transactions completed before the first EPC was issued. Furthermore,

[1]Land Register

considering that one dwelling may have several certificates issued, we compute the difference between the issuance dates of the EPC and the official sold date, and assign the newest certificate at the time of the sale as shown in equation 3.1.

$$min((soldDate - issuedDate) >= 0) \tag{3.1}$$

After assigning correct EPC, we omit all transactions prior to 2010, i.e., transactions pre-dating the EPC. Transactions between this period, but which do not possess a certificate issued before the official sold date, are also excluded. Table 3.4 lists the attributes we extract from the merged dataset and further analyze in Section 3.3. Initially, the primary room size was to be used as a measure for a dwelling's size. However, as the percentage missing data for this variable is fairly high, consequently, we decide to use the heated usable area instead, and hereby refer to it as size[2].

Table 3.4: List of attributes selected.

Data	**Type**	**Description**
Official price	Numerical	The official sold price
Size	Numerical	The dwellings size in square meters
Number of rooms	Numerical	The dwelling's total number of rooms
Number of bathrooms	Numerical	The dwelling's total number of bathrooms
Number of wc	Numerical	The dwelling's total number of toilets
Age	Numerical	Age of the dwelling at sales date
Build year	Numerical	Building construction year
HDD	Numerical	Mean yearly HDD-value for the past decade
Postal location	Categorical	
Energy rating	Categorical	Range from A-G
Heating rating	Categorical	Range from 1-5
EPC type	Categorical	Four types: simple, advanced, XML and schema
Building unit type	Categorical	Detached, semi-detached, apartments
Building material	Categorical	Wood, brick, steel, concrete
Heating source(s)	Categorical	Type of heating source installed in the dwelling
Elevator	Categorical	Whether or not the building has an elevator

Furthermore, as shown in Table 3.5, there is a significant amount of missing data in our dataset. To cope with the missing values, we use the *imputation* method presented in Allison (2001), by imputing values for data points with missing values for number of rooms, bathrooms and WC, with the mean values of all remaining dwellings. To minimize the resulting bias, we group dwellings by size using intervals of $50m^2$ starting from $0 - 50m^2$. Further examining Table 3.5 and the correlation between the variables, substantiates the imputation and grouping method, as the correlations between the variables are relatively high.

[2]Comparing the data for entries with both variables present (Primary room/heated usable are), results in an absolute mean size difference of **10.04 square meter** and a mean relative absolute error(mRAE) of **13.6%**.

Table 3.5: Correlation and missing values of selected variables

	# Rooms	# Bathrooms	# WC	Size	% missing values
# Rooms	1.000	0.568	0.657	0.818	26.98%
# Bathrooms	0.568	1.000	0.750	0.534	35.53%
# WC	0.657	0.750	1.000	0.651	35.65%
Size	0.818	0.534	0.651	1.000	0.00%

As a final step, we clean the dataset further by excluding data points containing extreme numerical values. The steps are described below.

1. Remove all buildings that are either commercial and hence, out of range for our research, or seasonal vacation houses, as these do not represent an actual dwelling with stable annual energy consumption. This corresponds to 6.0% of the total data.

2. Remove all transactions with missing energy and heating rating. This corresponds to $< 0.1\%$ of the total data.

3. Remove all transactions outside the size range $10m^2$ - $600m^2$, which corresponds to $< 0.1\%$. These are, as observed, either dwellings with faulty inputs or abnormal building types.

4. Finally, we group dwellings by size intervals of $50m^2$ once more, and remove highly suspicious transactions within each group with the number of rooms exceeding $mean(group_i) + 3 * SD(group_i)$. This results in an additional $< 0.4\%$ trim.

3.3 Exploratory Data Analysis

This section presents and visualizes the data after performing the cleaning process. Moreover, the results of the exploratory data analysis serve as a basis for the design of models in Section 4. We focus on showcasing the important aspects of our data with regards to the research topic.

Correlated Variables

First of all, deciding which dwelling-characteristics to include in the analysis of house prices need to be carefully considered, especially when dealing with similar variables which might be highly correlated.

The first step of selecting variables was done in the cleaning process, and the next step is to examine correlation and dependencies among the numerical variables. The variables *Number of WC* and *Number of bathrooms* were excluded due to a high dependency with

 # Rooms. The correlation matrices for all remaining numerical variables are shown in Tables 3.6 (Pearson's correlation) and 3.7 (Spearman's correlation).

Table 3.6: Pearsons Correlation of numerical variables

	# Rooms	Size	Energy Rating	Heating Rating	Age	HDD
# Rooms	1	0.781	-0.020	-0.024	-0.044	0.116
Size	0.781	1	0.053	-0.040	0.011	0.133
Energy Rating	-0.020	0.053	1	0.107	0.704	0.027
Heating Rating	-0.024	-0.040	0.107	1	0.072	-0.011
Age	-0.044	0.011	0.704	0.072	1	-0.040
HDD	0.116	0.133	0.027	-0.011	-0.040	1

Table 3.7: Spearman Correlation of numerical variables

	# Rooms	Size	Energy Rating	Heating Rating	Age	HDD
# Rooms	1	0.851	-0.013	-0.169	0.002	0.112
Size	0.851	1	0.033	-0.196	0.050	0.137
Energy Rating	-0.013	0.033	1	0.042	0.853	0.022
Heating Rating	-0.169	-0.196	0.042	1	0.014	-0.056
Age	0.002	0.050	0.853	0.014	1	-0.012
HDD	0.112	0.137	0.022	-0.056	-0.012	1

Unsurprisingly, the correlations matrices show a positive correlation between *Age/Energy Rating* and *Size/# Rooms*. Having variables with high correlation imply that there is a high dependency between the variables, and the value of one can be used to predict the other. Such cases are referred to as multicollinearity (Alin, 2010), and can affect the price analysis. A concern following the selection of variables is *Omitted Variable Bias*, which is stated as the challenge of excluding and deciding which variables to include to design the best-fitted model. Excluding may cause interference and result in other variables capturing the effect of excluded variables. Such exclusion may cause bias, but it is inevitably (Clarke, 2005).

Building Type

In our processed dataset, we have 222 392 dwellings and a total of 283 173 transactions. Thus, a portion of the transactions are resales of dwellings. This portion accounts for 60 781 transactions for 51 740 dwellings, sold two or more times. In the interest of capturing the effect of energy efficiency through the EPCs and price gap in different building types, we divide the dwellings into six separate building types: *apartment (small complex), apartment (large complex), semi-detached, townhouses, Detached* and *Detached w/apartment*. Apartments are separated by the size of the building they are located in. Large complex specifies apartment-buildings with four or more floors, while small complex defines buildings with three or fewer floors. Detached houses are likewise divided into two types, depending on if the dwelling has a separate apartment within the building structure. The reasoning behind the separation of some of the building types is to capture gaps in price and/or differences in heating/energy ratings. In the way that the energy need of an apartment may differ regarding the size and form of the apartment complex (Choi et al., 2012; Danielski, 2012) and houses with a separate apartment often have a higher price, due to the possibility of rental income. Figure 3.1 illustrates all transactions by each building type and the portion of resales in each group.

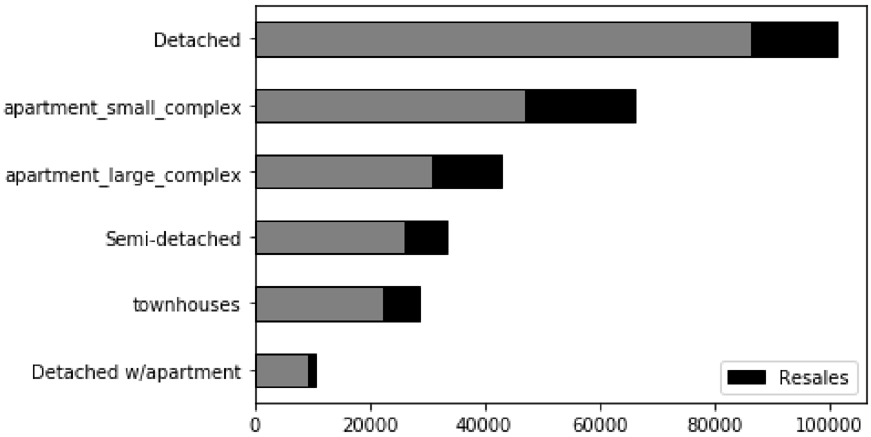


Figure 3.1: Distribution of building type

Energy/Heating Rating

Further, we present the distribution of energy and heating ratings in Figure 3.2 and 3.3. Occurrences of the highest energy ratings A/B are significantly lower compared to other ratings. A probable cause of this may be associated with the strict regulations implemented for obtaining the highest ratings. This phenomenon can also be observed in other similar studies (Brounen & Kok, 2011; Olaussen et al., 2017). The heating rating displayed in Figure 3.3 reveals that most dwellings are in category 4 and 5.

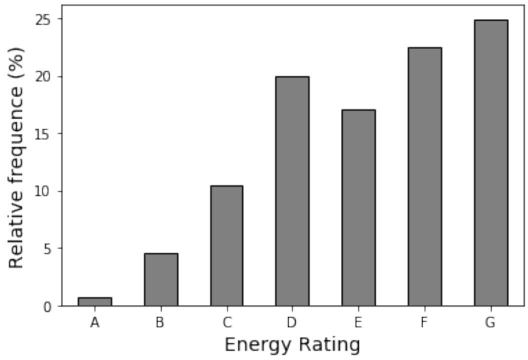


Figure 3.2: Distribution of Energy Ratings

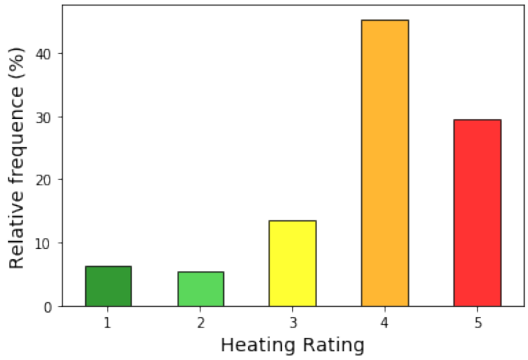


Figure 3.3: Distribution of Heating Ratings

Even though both of the ratings share similarities in the distribution, it is important to clarify that the ratings are independent, as can be observed from the close-to-zero correlation between these two variables, in Table 3.6 and 3.7. Furthermore, it is useful to examine the occurrences of the ratings combined. Table 3.8 shows the relative frequencies of all combinations of heating ratings and energy ratings in our datasets. To ensure the quality of the dataset and confirm that our dataset is a representative selection of issued EPCs in Norway, we compare the prevalence of heating/energy rating in our dataset to statistics about the official numbers produced by ENOVA as shown in Table 3.9. We observe a very similar pattern with respect to the magnitude of the frequencies.

Table 3.8: Relative frequencies of Heating/Energy Rating observations in our dataset.

	1	2	3	4	5
A	0.2%	0.1%	0.2%	0.1%	0.1%
B	0.7%	0.6%	0.5%	1.5%	1.2%
C	1.7%	0.9%	1.2%	3.4%	3.2%
D	1.3%	1.6%	1.8%	8.3%	7.0%
E	0.4%	0.5%	3.1%	8.3%	4.8%
F	0.9%	1.0%	3.2%	11.1%	6.4%
G	1.1%	1.0%	3.4%	12.5%	6.9%

Table 3.9: Relative frequencies of Heating/Energy Rating observations in the official ENOVA dataset[3]

	1	2	3	4	5
A	0.5%	0.4%	0.7%	0.3%	0.1%
B	1.4%	1.4%	0.9%	2.1%	1.5%
C	2.3%	1.7%	1.1%	2.9%	3.1%
D	1.6%	1.7%	1.5%	5.9%	7.1%
E	0.6%	0.8%	2.4%	6.1%	6.0%
F	1.1%	1.1%	2.6%	9.0%	7.2%
G	1.3%	1.0%	2.7%	12.3%	7.8%

In addition, the distribution of energy ratings/heating rating amongst the various building types is shown in Table 3.10 and 3.11. For the heating rating, we observe that 85% of the EPC with the highest heating rating (1) are held by apartments. Similar for the highest energy rating (A), 79% are amongst apartments. Examining the tables

[3]Source: https://www.energimerking.no/no/energimerking-bygg/energimerkestatistikk/ (24.04.2020)

gives us an indication of differences in the rating distribution between the building types, which is an interesting aspect to analyze further in Section 4.

Table 3.10: Relative frequencies of heating rating by each building type

	1	2	3	4	5
Apartment (large complex)	3.5%	2.9%	0.3%	1.3%	7.2%
Apartment (small complex)	1.9%	1.5%	0.6%	7.4%	12.1%
Detached	0.5%	0.5%	9.4%	21.4%	4.0%
Detached w/apartment	< 0.1%	0.1%	0.9%	2.2%	0.6%
Semi-detached	0.1%	0.1%	1.5%	7.3%	2.7%
Townhouses	0.3%	0.3%	0.9%	5.8%	2.8%

Table 3.11: Relative frequencies of energy rating by each building type

	A	B	C	D	E	F	G
Apartment (large complex)	0.4%	0.9%	2.1%	3.7%	2.1%	3.1%	3.2%
Apartment (small complex)	0.2%	1.1%	2.4%	5.2%	3.1%	5.1%	6.7%
Detached	< 0.1%	0.9%	2.5%	5.3%	7.5%	8.6%	11.1%
Detached w/apartment	< 0.1%	0.2%	0.4%	0.9%	0.9%	0.8%	0.5%
Semi-detached	< 0.1%	0.8%	1.7%	2.8%	1.6%	2.4%	2.5%
Townhouses	0.1%	0.7%	1.3%	2.3%	2.0%	2.8%	1.1%

Geographical location

Geographical location is a major influence on the price of a dwelling, and some parts of Norway are more desired by buyers. While focusing on EPC in the different climatic zones, it is important for our dataset to include differences in the location of dwellings.

Our dataset contains transactions for all 356 municipalities in Norway. Some regions are more populated and naturally, there is a substantial difference in the total number of transactions in each municipality. Such variations are also present in our dataset. Figure 3.4 shows how the dwellings in our dataset are distributed in Norway. The plot displays a good coverage of geographical differences with observations in each municipality. The variations in the dataset are more precisely shown in the heatmap of counties in Figure 3.5. To investigate the effect of EPC, climate, and geographical location, it is imperative that we accurately access these variations in the model.

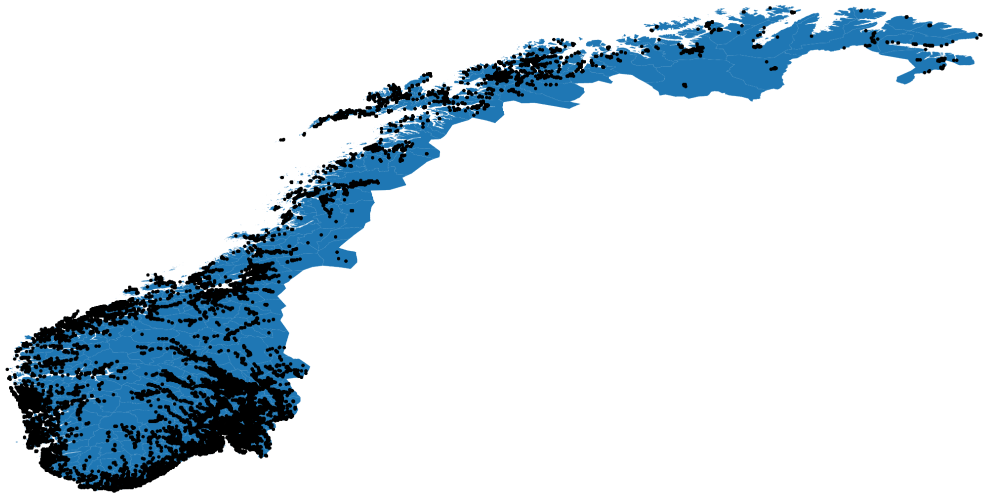


Figure 3.4: Geographical overview of all the transactions

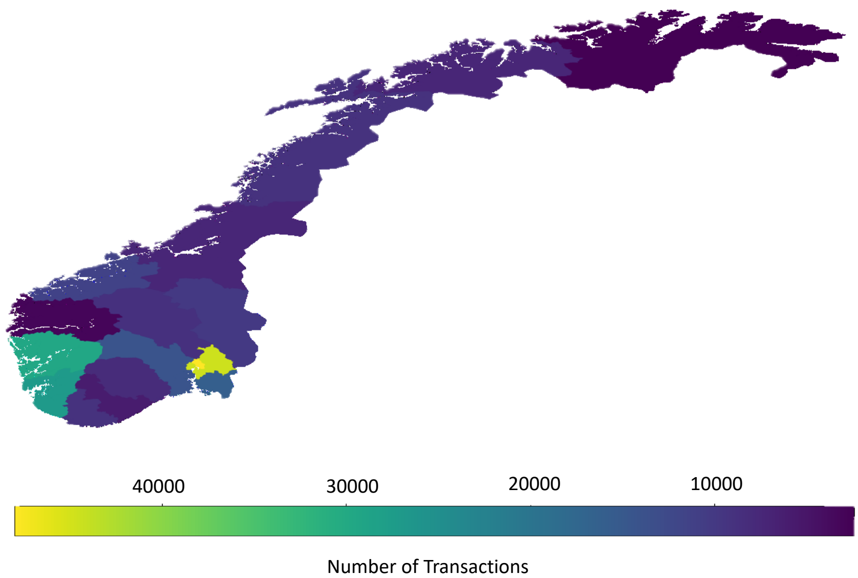


Figure 3.5: County heatmap of transactions

3.3.1 Temperature Data

When evaluating the effect of temperature on electricity demand (Do et al., 2016), electricity price (Do et al., 2019) or energy consumption in buildings (Quayle & Diaz, 1980; Eto, 1988), the concept of Heating Degree Days (HDD) is commonly used. In our study, we investigate the impact of energy efficiency on house prices, through the use of EPC. Since the impact of energy efficiency likely depends on how much heating is required in a given geographical area, it is highly useful to utilize HDD as a feature in our analysis.

Meteorological data is freely accessible in Norway, and we use the Met Frost API to collect the yearly sum of HDD with a temperature threshold of 17°C. The Met Frost API is built upon data from about 1300 weather sensor-systems in and around the Norwegian mainland. We are interested in the period 2010-2019, and retrieve the annual HDD value for each year, while omitting sensors with missing values. Then, using the geographical distance between each dwelling and the location of the sensors, we attach the closest sensor to each dwelling and the corresponding value. We choose to use the average of the annual values over the last decade, for the sake of collecting as much data as possible. The result of the process is HDD values from 387 different sensor systems. We further assign the mean yearly HDD using the data gathered as a feature for each dwelling. The descriptive statistics of the value are shown in Table 3.12. To ensure the validity of spatial geographical differences, we plot the distribution of the sensors in figure 3.6.

Table 3.12: Descriptive statistics of HDD value

Min value	Max value	Mean	St. Dev
2920	7504	3794	557

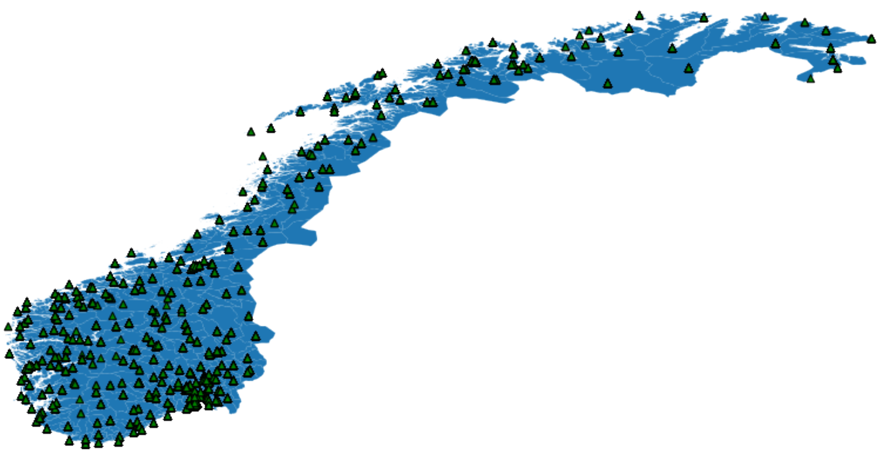


Figure 3.6: Distribution of sensors containing HDD data

3.3.2 Data Limitations and Simplifications

Using a unique identifier for each dwelling throughout each dataset was crucial in order to match each transaction with the exact dwelling, and EPC. The common identifier between dwellings in Norway is the land register number. As a consequence of land register numbers being modified and outdated, in particular with the municipality reform in 2020[4], we observe occurrences of mismatching identifiers in the EPC-dataset. To mitigate the problem, we add the building number as a merge-condition to ensure merging-correctness.

On the other hand, each property can have several different residential units, whereby the land register number stays the same. Such cases include apartment complex and other divided properties, however, each residential unit has its own unique residential number together with the other identifiers. For this reason, we merge apartments, and similar building types using the land register, building and residential number, but quickly realizing that approximately 50% of the apartment residents in the EPC- dataset is lacking this number.

Having consulted with experts at ENOVA and data scientists, we decide to account for the missing residential numbers, by merging based on the land register and building number and attach the same EPC for similar apartments within the same building with the approximately same size. This assumption is further enhanced by the EPC scheme and the written regulations for the EPC.

[4] //www.regjeringen.no/no/tema/kommuner-og-regioner/kommunereform/kommunereform/id2548377/

Chapter 4

Methodology and Results

4.1 Hedonic Pricing Method

The hedonic pricing approach is a commonly used method for the valuation of heterogeneous goods, particularly in the residential real estate market (Balk et al., 2013). Introduced by Rosen (1974), he argues that products are composites of various attributes or characteristics. The hedonic method recognizes that a product's value can be partitioned into the individual value of each attribute, and measures the marginal contribution of each part to the overall value. The main objective of this research is to determine the impact of energy efficiency on house prices, by estimating the contribution of the EPC characteristics, namely the energy rating and heating rating. Hence, we add these characteristics as explanatory variables in our model. Considering the informative dataset available, another interesting aspect to study in this context, is if the impact of energy efficiency varies with respect to building type. We therefore fit a regression model including all building types to begin with, and thereafter look more closely into each building type separately to investigate if the impact of energy efficiency remains similar across the types of buildings.

To effectively detect and interpret the impact of different attributes, we proceed by estimating several regressions. We start off with a base model including eight explanatory variables which are common for all dwellings, and add more variables in each succeeding step. Moreover, as our study includes dwellings from all over the country of Norway, we take into account regional differences by using a fixed effect model with the municipality code being the fixed parameter. The hedonic model is written as the following equation:

$$\ln(P_{ijt}) = \alpha + \sum_{k=1}^{K} \beta_k X_{ik} + \sum_{l=1}^{L} \lambda_l d_{tl} + \sum_{n=1}^{N} \delta_n m_{jn} + \epsilon_{ijt} \tag{4.1}$$

Where the dependent variable P denotes the transaction price per square meter of dwelling i in cluster j, and we use the logarithm as it allows us to easily interpret the model. X represents a set of K explanatory variables for dwelling i, and β the corresponding regression coefficients to be estimated. The additive term ($\sum_{n=1}^{N} \delta_n m_{jn}$) in eq. 4.1 accounts for the variations of housing prices in each municipality. To ensure sufficient data in each group, we merge municipalities with less than 20 transactions with the closest municipality, resulting in 295 regions (N) from the original 356. In eq. 4.1, m_i represents a set of N regions (dummies), where the corresponding code for building i is set to the binary value 1, and all the others to 0. The vector δ holds the estimated fixed-effect coefficients relative to the intercept, α. Note that this constant is derived based on all the selected baselines for the categorical variables.

Since our dataset contains transactions made in the time span from 2010 to 2019, we account for quarterly fluctuations in the house prices by the term $\sum_{l=1}^{L} \lambda_l d_{tl}$, where d_t is the set of quarterly dummies and holds a value of 1 only for the sales in period t. The estimated fixed-effect coefficients for each quarter, relative to α, is held in the vector λ.

Similar to Olaussen et al. (2017), rather than using the direct age of a dwelling, we specify the inverse of age in our model, mainly because we believe that a dwelling's age matter more at sale for newer buildings, and is less important as age increases. For the remaining categorical variables *building type* and *material*, we transform the values by utilizing dummy variables. In the first step of the regression with our base model, we define the baselines as *detached* and *wood* for the variables building type and material, respectively. The last variable *Elevator* indicates whether the building has an elevator (1) or not (0).

Additionally, our dataset can be seen as clusters of samples, with each cluster representing a municipality. To achieve more precise estimates, we want the clustered standard errors in order to capture variations within each municipality. Abadie et al. (2017) imply that in the case of clustering in sampling, cluster adjustment is important.

In the next step of the regression, we add energy performance variables and define the baseline to be the lowest rating (5) for the heating rating. In the case of the energy rating, the baseline is set to the second-lowest rating (F). The reason behind this is the same as Olaussen et al. (2017) discussed, that all dwellings that fail to insert the correct data or neglect the energy certification process, will automatically obtain the lowest rating (G).

The first two steps in our modelling follow several of the same principles and theories as other studies (Brounen & Kok, 2011), but adjusted and tweaked to fit the Norwegian real estate market and EPC scheme. In addition, we fit the models for the whole country rather than segments. Moreover in the last regression step, we utilize the temperature data and study if energy efficiency has a stronger impact on house prices in colder areas.

4.2 Base Model

The results for the base model are presented in Table 4.1. This model includes the traditional explanatory variables used in the greater portion of related literature, control variables for sales year, and municipality fixed effects to control for regional variations. The adjusted R squared is 0.771 for this model. We also calculate the mean absolute percentage error (MAPE)[1] for this model, resulting in approximately 23.1%, which is decent. All variables appear to be significant at the 1% level, with the exception of dummies *Brick, Steel, Semi-detached* and *Townhouses* for building material and type. The coefficient of *1/age* variable shows that newer dwellings have a higher price than older buildings. Dwellings with higher number of rooms have higher prices.

The time coefficients plot in Figure 4.1 visualizes the trend in the Norwegian real estate market, which has been increasing annually over the past decade. Moreover, the results indicate a small price discount for building materials *Concrete* and *Steel* compared to the baseline material, *Wood*. Apartments have a price premium of 8% compared to the baseline type *Detached*. This is not surprising, as we are modelling price per square meter, which is usually higher for smaller dwellings. Correspondingly the size coefficient is negative for the same reason. A detached house which includes a separate apartment results in approximately 7.3% price premium (per square meter) compared to one without.

[1]The MAPE is based on the difference between observed price per square meter and predicted price per square meter.

Table 4.1: Base model including the control variables, where the baselines are defined as follows: Wood (material) and detached (building type), and the age is transformed to 1/age.

	Dependent variable: log(price per square meter)
1/age	0.425***
	(0.033)
Rooms	0.015***
	(0.002)
Size	−0.003***
	(0.0001)
Elevator	0.046***
	(0.010)
Material Brick	0.024*
	(0.012)
Material Concrete	−0.059***
	(0.014)
Material Steel	−0.007
	(0.048)
Material Unknown	−0.139***
	(0.019)
Apartment (large complex)	0.080***
	(0.018)
Apartment (small complex)	0.052***
	(0.012)
Detached w/apartment	0.073***
	(0.010)
Semi-detached	−0.004
	(0.011)
Townhouse	−0.020*
	(0.011)
Region fixed effect	*Yes*
Quarterly fixed effect	*Yes*
Observations	283,170
R^2	0.772
Adjusted R^2	0.771
Note:	*p<0.1; **p<0.05; ***p<0.01

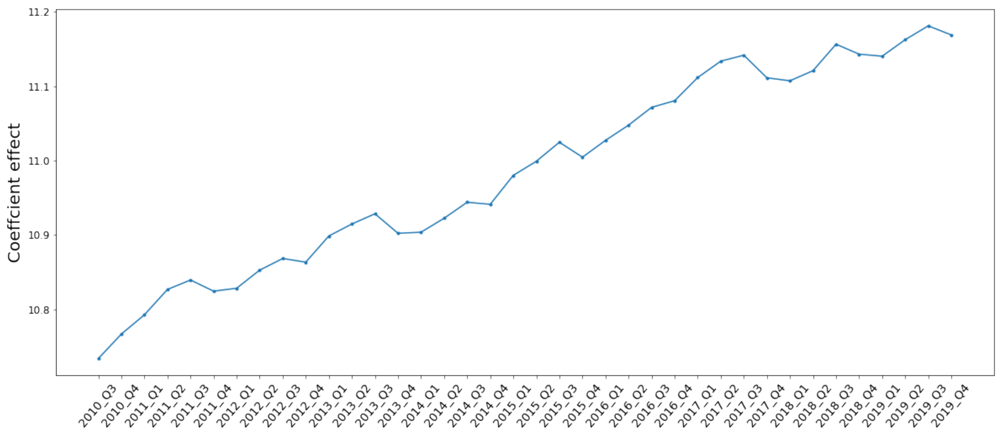


Figure 4.1: Plot of quarterly fixed effect coefficients for the base model

4.3 Models with Energy and Heating Ratings

In the next step, we run the same base model with all control variables, but add the energy- and heating rating variables in two separate models. We estimate the regression for each building type separately. The reason is that impact of explanatory variables on price per square meter can differ across building types. Moreover, we have a very large dataset, and we are therefore able to estimate models separately for each building type with sufficient precision. The results are shown in Table 4.2. In this table, we display three models for each building type: (1) the base model + energy rating, (2) base model + heating rating and (3) base model + energy + heating rating.

First of all, looking at model (1) for each building type, the results show that energy ratings are significant for almost all building types, except for energy rating G, which is significant only for detached dwellings.

In model (2) with heating rating alone, we observe that the heating rating overall is less significant and the magnitude and the sign of the coefficients is less intuitive and gives us an indicator that the heating rating may not be a decisive factor of the price of a dwelling, or capturing some other effect.

In the last model (3), when adding both ratings together, we can see the effect of the energy ratings mainly stays the same, while the heating rating becomes less significant for some types(*detached*). The effect of the heating rating is also reduced, further substantiating the results in model (2).

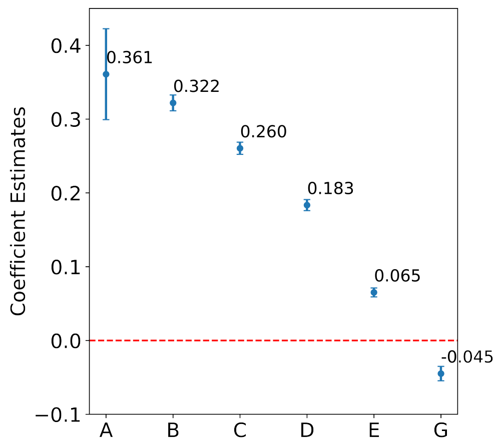
More interestingly, comparing the magnitude of energy rating coefficients amongst the building types, we observe a monotonic relationship between ratings A-G for detached dwellings. As the rating increase/decreases with respect to the baseline category F, the price of the dwelling follows. Similarly, the same relationship is present in semi-detached and detached with apartment but with reduced coefficients. However the results are not as clear in the case of apartments and townhouses, as the coefficients indicate little to none difference in impacts of ratings E or G. Furthermore, rating B has a higher effect compared to A (except for the apartment in small complex). Even though the energy rating seem to have an impact on the price in apartments and townhouses, the differences in the ratings seem to be less important, and the relationship amongst the energy ratings seem unclear. This indicates that the energy efficiency of a dwelling is appreciated differently, with regards to the type of dwelling.

In Figure 4.2, we include plots of coefficients and standard errors of the energy ratings from model (1) for each building type. The plots clearly demonstrate the differences between the building types and the price premium associated with the energy efficiency through the ratings. The standard errors indicate the estimates' accuracy and how values may overlap, and are always highest for the A rating, as there are fewer observations with this rating.

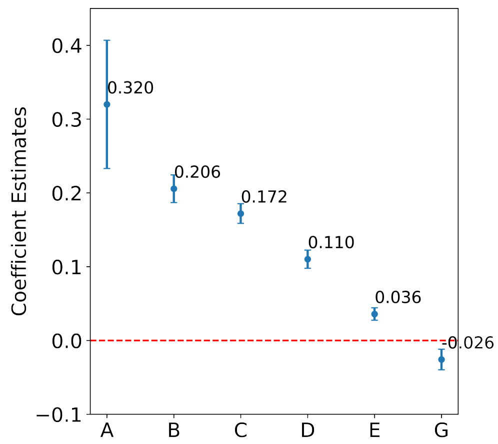
Table 4.2: Results of the base model(with the control variables) and energy/heating rating for each building type, with three models for each building type: (1) the base model + energy rating, (2) base model + heating rating and (3) base model + energy + heating rating. The baseline for each rating is F (energy) and 5 (heating).

	Dependent variable: log(price per square metre)																	
	Detached			Detached w/ap.			Semi-Detached			Ap Large			Ap Small			Townhouse		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
1/age	0.156^{***}	0.586^{***}	0.142^{***}	0.204^{***}	0.458^{***}	0.195^{***}	0.193^{***}	0.338^{***}	0.188^{***}	0.050	0.188^{***}	0.017	0.305^{***}	0.358^{***}	0.307^{***}	0.227^{***}	0.339^{***}	0.235^{***}
	(0.017)	(0.026)	(0.017)	(0.036)	(0.040)	(0.037)	(0.018)	(0.018)	(0.018)	(0.043)	(0.055)	(0.042)	(0.023)	(0.028)	(0.027)	(0.044)	(0.045)	(0.045)
Rooms	0.010^{***}	0.018^{***}	0.010^{***}	0.009^{***}	0.015^{***}	0.008^{***}	0.018^{***}	0.023^{***}	0.019^{***}	0.008	0.011	0.007	0.015^{***}	0.016^{***}	0.016^{***}	0.016^{***}	0.022^{***}	0.017^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.006)	(0.008)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Size	-0.003^{***}	-0.003^{***}	-0.003^{***}	-0.002^{***}	-0.002^{***}	-0.002^{***}	-0.003^{***}	-0.004^{***}	-0.003^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Elevator	0.060^{*}	0.086^{**}	0.060^{*}	0.075	0.077	0.083	0.139^{***}	0.164^{***}	0.138^{***}	0.006	0.040^{***}	0.008	0.072^{***}	0.081^{***}	0.074^{***}	0.036	0.035	0.033
	(0.035)	(0.036)	(0.035)	(0.053)	(0.058)	(0.053)	(0.050)	(0.055)	(0.050)	(0.009)	(0.010)	(0.008)	(0.014)	(0.016)	(0.016)	(0.052)	(0.052)	(0.050)
Material Brick	0.008	-0.012	0.007	-0.028	-0.033^{*}	-0.028	0.038^{***}	0.022	0.038^{***}	0.082^{***}	0.059^{**}	0.084^{***}	0.051^{***}	0.048^{***}	0.052^{***}	0.095^{***}	0.080^{***}	0.095^{***}
	(0.010)	(0.011)	(0.010)	(0.019)	(0.019)	(0.019)	(0.014)	(0.015)	(0.014)	(0.029)	(0.030)	(0.030)	(0.008)	(0.011)	(0.010)	(0.023)	(0.023)	(0.023)
Material Concrete	-0.005	-0.006	-0.006	0.016	0.030	0.014	-0.005	-0.005	-0.006	-0.028	-0.029	-0.023	-0.056^{***}	-0.060^{***}	-0.053^{***}	-0.003	-0.012	-0.003
	(0.010)	(0.011)	(0.010)	(0.018)	(0.019)	(0.018)	(0.011)	(0.011)	(0.011)	(0.027)	(0.025)	(0.026)	(0.018)	(0.015)	(0.014)	(0.013)	(0.013)	(0.013)
Material Steel	-0.076	-0.070	-0.081	0.112^{***}	0.104	0.103^{**}	0.041	0.058	0.041	-0.020	0.022	-0.014	-0.016	-0.015	-0.018	0.001	0.008	-0.006
	(0.100)	(0.098)	(0.100)	(0.040)	(0.066)	(0.042)	(0.119)	(0.131)	(0.123)	(0.040)	(0.041)	(0.040)	(0.120)	(0.120)	(0.120)	(0.043)	(0.047)	(0.047)
Material Unknown	-0.053^{**}	-0.094^{***}	-0.057^{**}	-0.033	-0.041	-0.031	-0.068	-0.096^{*}	-0.068	0.020	0.005	0.029	-0.083^{**}	-0.070^{**}	-0.069^{**}	-0.132^{**}	-0.168^{**}	-0.125^{*}
	(0.024)	(0.026)	(0.023)	(0.086)	(0.103)	(0.088)	(0.060)	(0.051)	(0.061)	(0.029)	(0.029)	(0.028)	(0.036)	(0.031)	(0.030)	(0.065)	(0.070)	(0.064)
A	0.361^{***}		0.364^{***}	0.320^{***}		0.308^{***}	0.192^{***}		0.191^{***}	0.197^{***}		0.210^{***}	0.097		0.085	0.035		0.040
	(0.062)		(0.058)	(0.087)		(0.089)	(0.054)		(0.054)	(0.067)		(0.062)	(0.060)		(0.062)	(0.118)		(0.114)
B	0.322^{***}		0.312^{***}	0.206^{***}		0.193^{***}	0.143^{***}		0.140^{***}	0.235^{***}		0.226^{***}	0.071^{***}		0.070^{***}	0.093^{***}		0.092^{***}
	(0.011)		(0.010)	(0.019)		(0.019)	(0.014)		(0.013)	(0.043)		(0.038)	(0.014)		(0.014)	(0.024)		(0.023)
C	0.260^{***}		0.258^{***}	0.172^{***}		0.170^{***}	0.134^{***}		0.131^{***}	0.251^{***}		0.240^{***}	0.079^{***}		0.079^{***}	0.118^{***}		0.126^{***}
	(0.008)		(0.008)	(0.013)		(0.014)	(0.011)		(0.011)	(0.023)		(0.019)	(0.010)		(0.010)	(0.015)		(0.013)
D	0.183^{***}		0.182^{***}	0.110^{***}		0.109^{***}	0.107^{***}		0.104^{***}	0.179^{***}		0.179^{***}	0.039^{***}		0.040^{***}	0.093^{***}		0.096^{***}
	(0.007)		(0.007)	(0.012)		(0.012)	(0.009)		(0.009)	(0.014)		(0.010)	(0.011)		(0.011)	(0.011)		(0.011)
E	0.065^{***}		0.067^{***}	0.036^{***}		0.038^{***}	0.030^{***}		0.030^{***}	0.052^{***}		0.062^{***}	0.015^{**}		0.015^{*}	0.013		0.015
	(0.006)		(0.006)	(0.008)		(0.008)	(0.007)		(0.007)	(0.017)		(0.020)	(0.007)		(0.008)	(0.010)		(0.009)
G	-0.045^{***}		-0.047^{***}	-0.026^{*}		-0.029^{**}	0.008		0.006	0.052^{***}		0.053^{***}	0.040^{***}		0.039^{***}	0.016		0.013
	(0.010)		(0.009)	(0.014)		(0.014)	(0.011)		(0.011)	(0.012)		(0.009)	(0.012)		(0.012)	(0.012)		(0.012)
Heating Rating 1		0.022	-0.008		0.047	0.033		0.003	-0.007		0.104^{***}	0.079^{***}		0.015	0.012		-0.048	-0.069
		(0.081)	(0.075)		(0.034)	(0.033)		(0.023)	(0.022)		(0.022)	(0.015)		(0.039)	(0.038)		(0.057)	(0.053)
Heating Rating 2		0.103^{***}	0.019		0.104^{***}	0.039^{*}		0.051^{*}	0.030		0.059^{***}	0.045^{***}		-0.056	-0.057		-0.024	-0.048
		(0.036)	(0.034)		(0.023)	(0.021)		(0.028)	(0.028)		(0.018)	(0.014)		(0.057)	(0.057)		(0.061)	(0.060)
Heating Rating 3		-0.048^{***}	-0.058^{***}		-0.036^{***}	-0.044^{***}		-0.039^{***}	-0.031^{***}		-0.021	-0.022		0.013	0.009		-0.038^{**}	-0.035^{**}
		(0.011)	(0.010)		(0.011)	(0.011)		(0.010)	(0.010)		(0.054)	(0.041)		(0.021)	(0.018)		(0.015)	(0.015)
Heating Rating 4		-0.037^{***}	-0.034^{***}		-0.027^{***}	-0.018^{*}		-0.023^{***}	-0.013^{**}		0.057^{***}	0.075^{***}		0.003	0.003		-0.011	-0.004
		(0.009)	(0.009)		(0.010)	(0.010)		(0.006)	(0.006)		(0.018)	(0.024)		(0.005)	(0.005)		(0.007)	(0.007)
Quarterly fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Y*	*Yes*
Region fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	101,282	101,282	101,282	10,546	10,546	10,546	33,320	33,320	33,320	42,976	42,976	42,976	66,263	66,263	66,263	28,783	28,783	28,783
R^2	0.662	0.636	0.663	0.695	0.682	0.697	0.727	0.719	0.728	0.572	0.554	0.577	0.663	0.662	0.664	0.674	0.665	0.675
Adjusted R^2	0.661	0.635	0.662	0.687	0.673	0.688	0.725	0.717	0.725	0.570	0.552	0.576	0.661	0.660	0.662	0.671	0.662	0.672

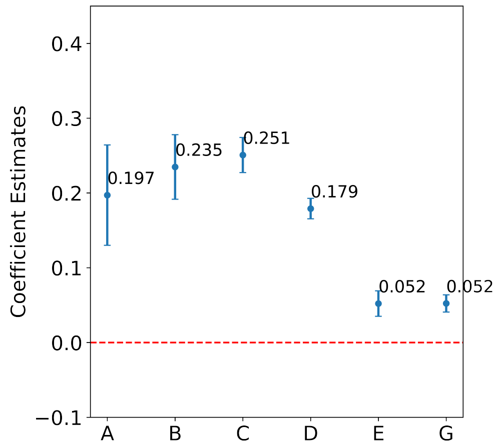
Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$



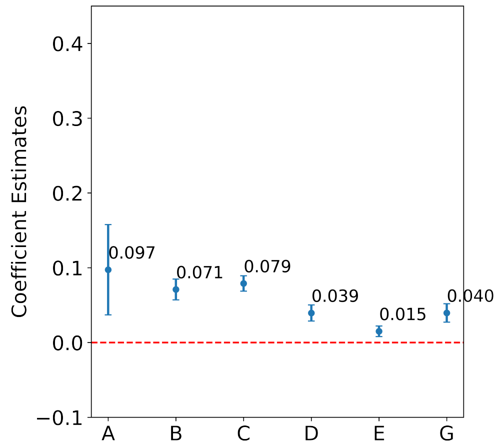
(a) Detached



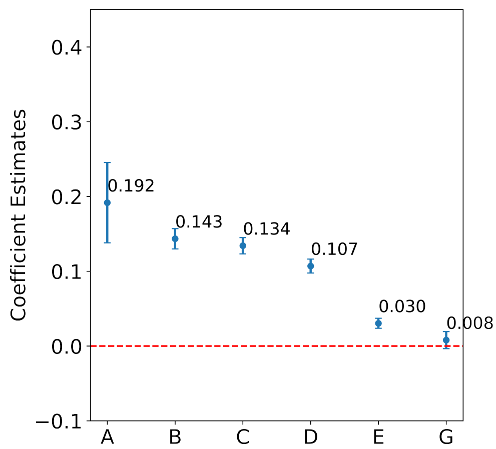
(b) Detached w/apartment



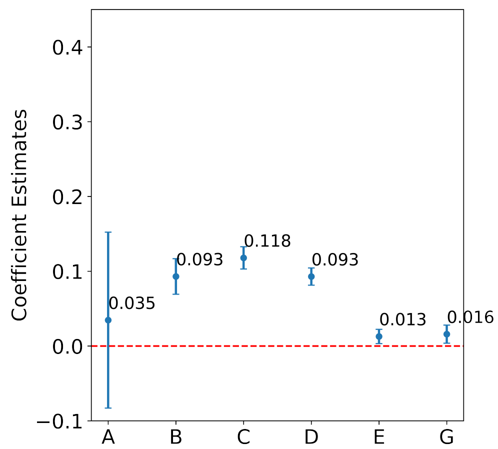
(c) Apartments large complex



(d) Apartments small complex



(e) Semi-detached



(f) Townhouse

Figure 4.2: Energy rating coefficient estimates with std. error for each building type in the base model with energy ratings. Rating F is chosen as the energy rating baseline.

4.4 Model with HDD

Having investigated the effect of attributes on house prices, we further utilize transformed temperature data as HDD-values. We introduce the HDD variable to our base model, and the results are presented in Table 4.3. As mentioned in Section 3.3.1, the HDD variable ranges from 2920 to 7504. We standardize this variable, resulting in a mean of 0 with a standard deviation of 1. This has two practical advantages; estimated regression coefficients will be more similar in magnitude to estimated coefficients of other variables, and interpretation of this coefficient will be more intuitive. The standardization is performed following the formula:

$$z = \frac{(X - \mu)}{\sigma} \tag{4.2}$$

Where z is the standardized value for the HDD value X, μ is the mean of all HDD values present in our dataset and σ is the standard deviation. The results remain almost identical to those presented previously, which means that the model is able to capture the HDD variations by a significance level of 5%. Furthermore, the coefficient for *HDD* appears with the expected sign and indicates a moderate price discount (of approximately 6% as the HDD increases by one standard deviation) for dwellings located in colder parts of Norway.

Table 4.3: Base model with the addition of HDD-values

	Dependent variable: log(price per square metre)
1/age	0.423***
	(0.034)
Rooms	0.015***
	(0.002)
Size	−0.003***
	(0.0001)
Elevator	0.047***
	(0.010)
Material Brick	0.016
	(0.014)
Material Concrete	−0.057***
	(0.013)
Material Steel	−0.010
	(0.047)
Material Unknown	−0.139***
	(0.018)
Apartment (large complex)	0.072***
	(0.017)
Apartment (small complex)	0.052***
	(0.012)
Detached w/apartment	0.072***
	(0.011)
Semi-detached	−0.003
	(0.012)
Townhouse	−0.018
	(0.011)
HDD	−0.062**
	(0.026)
Region fixed effects	*Yes*
Quarterly fixed effects	*Yes*
Observations	283,170
R^2	0.773
Adjusted R^2	0.773
Note:	*p<0.1; **p<0.05; ***p<0.01

4.5 Interaction between Energy Rating and HDD

To investigate if energy efficiency appreciation varies in different climate zones, we define the interaction of energy rating and the HDD value as a variable. This is based on the intuition that energy rating is a measure of energy efficiency, and the HDD value is a measure of the amount of heating needed. The model is explained by Equation 4.1 with an additional set of explanatory variables

$$\sum_{l=1}^{L} \zeta_l R_{il} Z_i \tag{4.3}$$

where R is the set of l energy rating dummies and Z is the standardized HDD value for dwelling i. ζ represents the corresponding coefficients to be estimated for each combination. Thus, the total impact of energy efficiency on house prices as designed in these models is not only the estimated coefficient for energy rating but also the contribution (ζ) of energy efficiency with respect to climatic conditions.

The results are presented in Table 4.4. The same control variables as before are included, and no significant deviations from previous models can be observed. All coefficients for HDD appear with the expected signs, with variations in magnitude and significance level across the different building types. The models for *All* and *Detached* indicate strong significance levels for all interaction variables except *A*HDD*, which may be due to data limitations for this category. The coefficients show a small increasing price premium for energy ratings B (and A for detached)-E compared to F, and a small discount for G in both *All* and *Detached* models. For other detached types and townhouses, the effect seem to be moderate, however, with lower significance levels for some energy ratings.

To put the coefficients into perspective, consider model (2) in Table 4.4. For instance, the price premium associated with an energy rating of B is estimated to 21% compared to the baseline, F. However, to get the total impact, we must consider the coefficient for *B*HDD* as well, resulting in a price premium of 21% + (5.7%*HDD). Hence, the actual impact varies significantly in different climate zones.

Furthermore, the results show no strong evidence of the interaction variables having any significant effect on house prices for apartments in large and small complexes. Some mixed and counter-intuitive coefficients appear for the *A*HDD* and *G*HDD*, with possible explanations being 1) as mentioned earlier, the uncertainty associated with the G category, and 2) the wide-spread, but potentially false assumption that dwellings in larger complexes tend to keep heat more effectively compared to detached types (Wright, 2008). Another explanation could be that occupants in larger complexes are given a fixed price energy deal, and compared to other building types, apartments in large complex often have a central district heating system as heating source (ENOVA SF, 2017).

The overall result is that energy efficiency is more appreciated in colder environments, however, only for building types with few units. On the other hand, as the building structure takes a larger form, the energy efficiency seem to matter less regardless of temperature.

Table 4.4: Models with respect to energy rating, climate and prices. These are estimated for each building type to investigate the impact of energy efficiency more closely. Base variables (age, size etc.), quarterly and region fixed effects are included. The baseline for energy rating is set to F.

	Dependent variable: log(price per square metre)													
	All		Detached		Detached w/ap.		Semi-Detached		ApLarge		ApSmall		Townhouse	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
A	0.194^{***}	0.192^{***}	0.361^{***}	0.350^{***}	0.320^{***}	0.318^{***}	0.192^{***}	0.195^{***}	0.197^{***}	0.198^{***}	0.097	0.100	0.035	0.027
	(0.038)	(0.038)	(0.062)	(0.051)	(0.087)	(0.085)	(0.054)	(0.053)	(0.067)	(0.060)	(0.060)	(0.061)	(0.118)	(0.133)
B	0.210^{***}	0.210^{***}	0.322^{***}	0.326^{***}	0.206^{***}	0.207^{***}	0.143^{***}	0.144^{***}	0.235^{***}	0.224^{***}	0.071^{***}	0.068^{***}	0.093^{***}	0.100^{***}
	(0.009)	(0.009)	(0.011)	(0.011)	(0.019)	(0.020)	(0.014)	(0.015)	(0.043)	(0.047)	(0.014)	(0.014)	(0.024)	(0.022)
C	0.199^{***}	0.199^{***}	0.260^{***}	0.264^{***}	0.172^{***}	0.174^{***}	0.134^{***}	0.135^{***}	0.251^{***}	0.237^{***}	0.079^{***}	0.076^{***}	0.118^{***}	0.121^{***}
	(0.008)	(0.007)	(0.008)	(0.008)	(0.013)	(0.014)	(0.011)	(0.011)	(0.023)	(0.030)	(0.010)	(0.011)	(0.015)	(0.014)
D	0.143^{***}	0.141^{***}	0.183^{***}	0.184^{***}	0.110^{***}	0.109^{***}	0.107^{***}	0.107^{***}	0.179^{***}	0.165^{***}	0.039^{***}	0.038^{***}	0.093^{***}	0.092^{***}
	(0.007)	(0.008)	(0.007)	(0.007)	(0.012)	(0.013)	(0.009)	(0.009)	(0.014)	(0.018)	(0.011)	(0.011)	(0.011)	(0.011)
E	0.054^{***}	0.051^{***}	0.065^{***}	0.065^{***}	0.036^{***}	0.034^{***}	0.030^{***}	0.030^{***}	0.052^{***}	0.047^{***}	0.015^{**}	0.015^{**}	0.013	0.013
	(0.006)	(0.005)	(0.006)	(0.006)	(0.008)	(0.009)	(0.007)	(0.006)	(0.017)	(0.014)	(0.007)	(0.007)	(0.010)	(0.009)
G	−0.005	−0.007	-0.045^{***}	-0.044^{***}	-0.026^{*}	-0.027^{**}	0.008	0.007	0.052^{***}	0.049^{***}	0.040^{***}	0.032^{**}	0.016	0.010
	(0.014)	(0.011)	(0.010)	(0.009)	(0.014)	(0.013)	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.015)	(0.012)	(0.010)
HDD		-0.066^{**}		-0.042^{*}		−0.063		−0.051		-0.083^{**}		-0.094^{***}		-0.086^{**}
		(0.027)		(0.024)		(0.039)		(0.032)		(0.039)		(0.035)		(0.036)
A*HDD		0.010		0.109^{*}		0.007		0.012		−0.017		−0.059		0.081
		(0.027)		(0.057)		(0.073)		(0.036)		(0.032)		(0.045)		(0.088)
B*HDD		0.057^{***}		0.054^{***}		0.048^{**}		0.042^{***}		0.051		0.019		0.049^{**}
		(0.013)		(0.012)		(0.020)		(0.013)		(0.042)		(0.016)		(0.022)
C*HDD		0.041^{***}		0.037^{***}		0.030^{***}		0.022^{**}		0.043		0.021		0.039^{***}
		(0.009)		(0.009)		(0.011)		(0.009)		(0.032)		(0.014)		(0.010)
D*HDD		0.030^{**}		0.021^{***}		0.019^{*}		0.013^{*}		0.043		0.009		0.011
		(0.012)		(0.006)		(0.011)		(0.008)		(0.041)		(0.017)		(0.009)
E*HDD		0.018^{***}		0.013^{***}		0.016^{**}		0.015^{**}		0.015		0.013		0.019^{***}
		(0.004)		(0.003)		(0.008)		(0.007)		(0.009)		(0.010)		(0.007)
G*HDD		-0.024^{***}		-0.028^{***}		−0.011		−0.002		0.022		−0.001		0.007
		(0.006)		(0.005)		(0.013)		(0.008)		(0.019)		(0.006)		(0.013)
Quarterly fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Base variables	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	283,170	283,170	101,282	101,282	10,546	10,546	33,320	33,320	42,976	42,976	66,263	66,263	28,783	28,783
R^2	0.782	0.784	0.662	0.664	0.695	0.697	0.727	0.729	0.572	0.580	0.663	0.669	0.674	0.679
Adjusted R^2	0.781	0.784	0.661	0.663	0.687	0.688	0.725	0.726	0.570	0.578	0.661	0.668	0.671	0.676

Note: *p<0.1; **p<0.05; ***p<0.01

Chapter 5

Conclusion

Energy Performance Certificate (EPC) is a measure of a dwelling's energy efficiency, introduced as a tool to reduce household energy consumption. Using an empirical approach with a hedonic pricing method, we find evidence of a positive relationship between price per square meter and energy efficiency ratings from the EPC. Our estimates indicate that compared to dwellings rated F, those in band A-C are sold at a premium of approximately 20% and those in band D and E at a premium of 14% and 5%, respectively. Additionally, dwellings rated G are sold at a slight discount of 0.5% compared to F.

The energy rating estimates are further strengthened with the addition of interaction between HDD and energy rating, indicating that energy efficiency play an even more important role in colder regions of Norway. Furthermore, we find that the impact of energy efficiency on dwelling prices vary in magnitude across the different building types. We find the most significant difference to be between detached dwellings and apartments in the upper bands of energy rating. For instance, a detached dwelling with rating B is sold at a price premium of 25% compared to one with rating E, while this price premium is only 18% for apartments in large complexes.

Moreover, we find mixed results and less intuitive associations between the heating rating and price. This may have important implications for policymakers. As explained in Section 2, one drawback of the current system is thought to be the complexity with the composition of two separate ratings: energy rating and heating rating. We find no evidence of heating rating as an important determinant in house prices. In addition, the EPC scheme as of today is slightly biased towards newer dwellings[1], as indicated by the high correlation between the age of the dwelling and the energy ratings, thereby supporting the idea of abandoning heating rating, and renewal of the EPC framework in Norway.

[1]The increased energy requirements in building regulations (TEK10 and TEK17), has made it even harder for older buildings (a major part of the current building stock, dwellings built before 1981 (39 years) account for approximately 38% of the current building stock source: https://www.ssb.no/statbank/table/06266/ (08.06.2020)

The existing literature is inconclusive regarding the impact of energy performance certificates. Although EPC has been active in most countries for about ten years, the implementation varies across countries (de Ayala et al., 2016; Bio Intelligence Service & IEEP, 2013), and the data quality and quantity also varies considerably throughout the literature (Pasichnyi et al., 2019). Thus, one explanation for the discrepancies in the literature is the varying data quality. The data used in this study are comprehensive with respect to both quantity and quality, and by utilizing the data we are able to capture the price premiums of energy efficiency for all building types across the country. Besides, compared to some of the existing literature (Brounen & Kok, 2011; Bio Intelligence Service & IEEP, 2013), our dataset have an advantage of decent distribution among all the energy ratings, especially the upper-scale ratings A/B, which allow us to estimate models with sufficient data in each category.

In accordance with our results, Olaussen et al. find that houses with better energy certification are sold for higher prices. Olaussen et al. argue that the price premiums related to the EPC ratings are a proxy for general building characteristics, which may be a valid explanation. However, our results show that the effect of energy efficiency ratings are stronger in colder areas. This is a very strong reason to believe that the energy rating serve as a measure for a dwelling's energy efficiency, even though they might also be capturing the effect of some other building characteristics.

Inspecting the current state of the EPC in Norway, we notice that the relative percentages of good ratings (A, B and C) issued since 2010 has slightly increased annually, ranging from 10.5% of ABC ratings in 2010 to 31.2% in 2020[2]. Which corresponds to the findings of Franke & Nadler (2019), that the importance of EPC has increased with time. However, it is challenging to distinguish whether the increase in good ratings in Norway is a result of the stricter requirements in the building regulations (TEK10 and TEK17)[3], or if home buyers invest in better energy efficiency, nevertheless, it demonstrates how the building stock is slowly moving towards better energy efficiency.

[2]Number as of 10.06.2020, source: the official ENOVA SF statistics: `https://www.energimerking.no/no/energimerking-bygg/energimerkestatistikk/`

[3]The current baseline for obtaining the energy rating C is the requirements in TEK10, and if a building is built by TEK17, it is more likely to get a higher grade (B), assuming a high efficient heating system is installed.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017, November). *When Should You Adjust Standard Errors for Clustering?* (NBER Working Papers No. 24003). National Bureau of Economic Research, Inc.

Alin, A. (2010). Multicollinearity. *WIREs Computational Statistics*, *2*(3), 370-374.

Allison, P. D. (2001). *Missing data.* Thousand Oaks, CA: Sage Publications.

Amecke, H. (2012, 07). The impact of energy performance certificates: A survey of german home owners. *Energy Policy*, *46*.

Arcipowska, A., Mariottini, F., Anagnostopoulos, F., & Kunkel, S. (2015). *Energy performance certificates across the eu.*

Balk, B., Diewert, W. E., Fenwick, D., Prud'homme, M., & de Haan, J. (2013). *Handbook on residential property prices indices (rppis) (english).* Washington DC: Eurostat methodologies and working papers.

Bio Intelligence Service, R. L., & IEEP. (2013). Energy performance certificates in buildings and their impact on transaction prices and rents in selected eu countries, final report prepared for european commission. *DG Energy*.

Bisello, A., Antoniucci, V., & Marella, G. (2020). Measuring the price premium of energy efficiency: A two-step analysis in the italian housing market. *Energy and Buildings*, *208*, 109670.

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.

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

Choi, I. Y., Cho, S. H., & Kim, J. T. (2012). Energy consumption characteristics of high-rise apartment buildings according to building shape and mixed-use development. *Energy and Buildings*, *46*, 123 - 131. (Sustainable and healthy buildings)

Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science*, *22*(4), 341-352.

Danielski, I. (2012). Large variations in specific final energy use in swedish apartment buildings: Causes and solutions. *Energy and Buildings*, *49*, 276 - 285.

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

Dell'Anna, F., Bravi, M., Duarte, C., Bottero, M., & Chen, A. (2019, 10). Epc green premium in two different european climate zones: A comparative study between barcelona and turin. *Sustainability*, *11*, 5605.

Do, L. P. C., Lin, K.-H., & Molnár, P. (2016). Electricity consumption modelling: A case of germany. *Economic Modelling*, *55*, 92–101.

Do, L. P. C., Lyócsa, Š., & Molnár, P. (2019). Impact of wind and solar production on electricity prices: Quantile regression approach. *Journal of the Operational Research Society*, *70*(10), 1752–1768.

ENOVA SF. (2017). *Enovas byggstatistikk 2017* (Tech. Rep.). Trondheim: ENOVA SF.

ENOVA SF. (2019). *Forprosjekt ny energimerkeordning - hovedrapport* (Tech. Rep.). Trondheim: ENOVA SF.

Environment, U., & Agency, I. E. (2017). Towards a zero-emission, efficient, and resilient buildings and construction sector. global status report 2017.

Eto, J. H. (1988). On using degree-days to account for the effects of weather on annual energy use in office buildings. *Energy and Buildings*, *12*(2), 113 - 127.

EU. (2002). Directive 2002/91/ec of the european parliament and of the council of 16 december 2002 on the energy performance of buildings. *Official Journal of the European Communities*.

Fleckinger, P., Glachant, M., & Kamga], P.-H. T. (2019). Energy performance certificates and investments in building energy efficiency: A theoretical analysis. *Energy Economics*, *84*, 104604. (Eighth Atlantic Workshop on Energy and Environmental Economics)

Franke, M., & Nadler, C. (2019). Energy efficiency in the german residential housing market: Its influence on tenants and owners. *Energy Policy*, *128*, 879 - 890.

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

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

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

Hårsman, B., Daghbashyan, Z., & Chaudhary, P. (2016). On the quality and impact of residential energy performance certificates. *Energy and Buildings*, *133*, 711 - 723.

Kahn, M. E., & Kok, N. (2014). The capitalization of green labels in the california housing market. *Regional Science and Urban Economics*, *47*, 25 - 34. (SI: Tribute to John Quigley)

Lainé, L. (2014). *Room for improvement: the impact of epcs on consumer decision-making, uk.*

Murphy, L. (2014, 04). The influence of the energy performance certificate: The dutch case. *Energy Policy*, *67*, 664-672.

Olaussen, J. O., Oust, A., & Solstad, J. T. (2017). Energy performance certificates – informing the informed or the indifferent? *Energy Policy*, *111*, 246 - 254.

Olaussen, J. O., Oust, A., Solstad, J. T., & Kristiansen, L. (2019). Energy performance certificates—the role of the energy price. *Energies*, *12*(18).

Pasichnyi, O., Wallin, J., Levihn, F., Shahrokni, H., & Kordas, O. (2019). Energy performance certificates — new opportunities for data-enabled urban energy policy instruments? *Energy Policy*, *127*, 486 - 499.

Quayle, R. G., & Diaz, H. F. (1980). Heating degree day data applied to residential heating energy consumption. *Journal of Applied Meteorology*, *19*(3), 241-246.

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

Shine, K. P., & Forster, P. M. F. (1999). The effect of human activity on radiative forcing of climate change: a review of recent developments. *Global and Planetary Change*, *20*(4), 205 - 225.

Skripnuk, D. F., & Samylovskaya, E. A. (2018, aug). Human activity and the global temperature of the planet. *IOP Conference Series: Earth and Environmental Science*, *180*, 012021.

Technical Expert Group. (2019a). *Proposal for an eu green bond standard.*

Technical Expert Group. (2019b). *Taxonomy technical report.*

Wahlström, M. H. (2016). Doing good but not that well? a dilemma for energy conserving homeowners. *Energy Economics*, *60*, 197 - 205.

Wilhelmsson, M. (2019). Energy performance certificates and its capitalization in housing values in sweden. *Sustainability*, *11*(21).

Wright, A. (2008, 12). What is the relationship between built form and energy use in dwellings? energy policy. *Energy Policy*, *36*, 4544-4547. doi: 10.1016/j.enpol.2008.09.014

Impact of temperature on house prices in norway

(Paper two)

June 15, 2020

Abstract

Meteorological conditions are used as a determining factor in relocation decisions, and the cost of nice weather is often referred to as an amenity. We utilize temperature data from Norway and apply the common hedonic pricing method to determine the impact of different temperature transformations on house prices. When utilizing the temperature as a climate variable, we find that the different temperature transformations are almost equally good, with HDD being the slightly better transformation. The results suggest that overall in Norway, temperature is an amenity, and higher winter temperatures are slightly more appreciated compared to summer. Our dataset consists of well-detailed house transactions and weather data with good spatial distribution, and this research can serve as a base for future research.

Contents

List of Tables

Chapter 1

Introduction

House prices vary naturally with time, and can be described with macro- and micro economical factors. However, there are also environmental and geographical variations which are important to take into account.

There is a global trend of migration towards locations with mild winters (Rappaport, 2007), warm summers and places with less rain and wind. People seem to use, consciously or unconsciously, meteorological conditions as a determinant for choosing where to live, in combination with other socioeconomic conditions. Such migration can be found enclosed within a country, or across borders. The cost of nice weather is often referred to as price premiums related to different amenities, and existing literature uses various methods to find the impact of weather on house prices. Rosen (1974) and Roback (1982) were the first to perform hedonic regression to analyse how different geographical locations are linked to differences in house prices and wages.

Studies investigating the impact of meteorological conditions on house prices are present, but the number of geographical locations studied are limited. Due to recent increase in global temperature (Giannakopoulos et al., 2005) and climate change, meteorological factors are of increasing importance (Costa & Kahn, 2003)[1]. Butsic et al. (2011) found that a reduction in snowfall in ski resorts due to global warming lead to a reduction in house prices. Studies from Italy, Germany, and Britain have shown that there is a price premium related with better climate, e.g., milder winters and summers (Rehdanz, 2006; Maddison & Bigano, 2003; Rehdanz & Maddison, 2009). Similar studies conducted in the US found that people preferred to live in places with warm winters, cooler summers and less rainfall (Rappaport, 2007; Englin, 1996). Cheshire & Magrini (2006) investigated population growth and weather, and were able to show that weather mattered solely within a country and not across country-borders in relocation decisions.

Literature shows that climate conditions are highly correlated to the decision-making regarding where to live. We want to examine the impact of different weather variables

[1]Showed that the importance of weather on house prices has increased from 1970 to 1999.

on house prices, using Norway as the geographical location. Such cases have rarely been discussed with Norway as geographical location, and we are only able to find one Master's thesis studying something similar; Jacobsen & Røisehagen (2016) mainly investigated the impact of the weather on the bidding day on house prices in one Norwegian city, Asker. Norway is known to be a country with significant variations in regional climate. Not only is northern Norway much colder than the southern Norway, but areas further from the ocean have a significantly different climate compared to the coastal areas. Norway also has the second longest coastline in the world[2], and a lot of high mountains, making it a suitable geographical region for the purpose of this research.

We use a large dataset containing transactions of dwellings in Norway, which contains data for dwellings in all cities and further attach averaged weather variables for each dwelling. Which allow us to create a hedonic model to analyse the impact of weather variables on the price of a dwelling.

Our main findings share some similarities with other studies. Utilizing three different measures of historical temperature data, we find that all three transformations are almost equally good, with HDD being the slightly better transformation. We also find that overall in Norway, temperature is an amenity, where higher winter temperatures are slightly more appreciated compared to summer.

Part two of the thesis proceeds by Section 2 explaining the data utilized, Section 3 describes the method and the results, while Section 4 includes concluding remarks and outline future work.

[2]Measured in a straight line, the coastline is 2500 km long. However, including all the fjords and islands, the total length is approximately 100 915 km.
https://www.regjeringen.no/en/topics/climate-and-environment/biodiversity/innsiktsartikler-naturmangfold/hav-og-kyst/id2076396/

Chapter 2

Data

The data used in this research was acquired by utilizing the Met Frost API, developed by the Meteorological Institute of Norway. We use this API to fetch meteorological values and variables to use in the analyses of their impact on house prices. To obtain datasets about transactions and dwellings in Norway, we partnered up with *Alva Technologies*, who gave us access to two large and well-detailed datasets about transactions and dwelling characteristics in the time period 1991-2019. A short summary of the two datasets is shown in Table 2.1.

Table 2.1: Overview of the Datasets

Dataset	**Type**	**size(n)**	**Source**	**Time-span**
Dwelling characteristics	CSV	2 572 317	*Alva Technologies*	
Transactions history	CSV	2 792 731	*Alva Technologies*	Jan 1991-Dec 2019

To use the data in the analysis and modelling, we first merge the dwelling and transaction datasets, resulting in a total of approximately 2 793 731 transactions with the respective dwelling characteristics. We apply a simple cleaning process to exclude data containing ambiguous or missing values for critical dwelling characteristics, such as size, construction year, and number of rooms, resulting in a modified dataset of 1 610 956 transactions between 1991-2019. In Table 2.2, we list the selected attributes to form the base model with control variables for our analysis.

Table 2.2: List of dwelling attributes selected as control variables.

Data	**Type**	**Description**
Official price	Numerical	The official sold price
Size	Numerical	The dwellings size in square meters
Age	Numerical	Age of the dwelling at sales date
Build year	Numerical	Building construction year
Building unit type	Categorical	Detached, semi-detached, apartments, serial-house
Location	Categorical	The municipality
Elevator	Categorical	Whether or not the building has an elevator

Furthermore, we fetch certain weather variables from the Met Frost API. The Met frost API is built upon data from about 1300 sensor systems located in Norway. Each sensor have different amount of information stored, and as a consequence, the variation of number of sensors may vary amongst the different variables. Next, we attach the meteorological data values using the closest sensor to each dwelling, and assign the corresponding values. The values we use, are daily averages transformed into monthly normal weather where the sensors are located. We set the time span as 1991-2019, the same as for the transactions. The selected weather variables are shown in Table 2.3, with a description and number of sensors used to obtain the data.

Table 2.3: Overview of the meteorological data we fetch from the Met Frost API, we get the values in average sum of monthly values from the period 1991-2019, except for the temperature, were we get the average of each month.

Variable	**Description**	**Sensors**	**Unit**
Precipitation	Amount of precipitation (cumulative) in a given time period.	622	mm
Precipitation Days	Number of days with precipitation (> 1.0mm) in a given time period.	619	days
Overcast Days	Number of days with overcast weather defined as days with sum of cloud cover values **above** a specific limit throughout the day.	95	days
Clear Days	Number of days with clear weather defined as days with sum of cloud cover values **under** a specific limit throughout the day.	95	days
Temperature	The mean temperature throughout a day which is an arithmetic mean of 24 hour values.	412	Celsius
Heating Degree Days	A HDD value where the HDD is defined as the number of degrees the daily average temperature is below a threshold value of 17 ºC. If the daily average temperature is equal to 17 ºC or higher, the HDD value is set to zero.	387	HDD

The weather variables are some of the most common weather variables used in literature, and we use the specific variables for the purpose of being able to compare our results to existing research. To further examine the weather variables and how the values varies, we list the descriptive statistics of each monthly normal weather between 1991-2019 with respect to the whole country in Table 2.4.

Table 2.4: Descriptive statistics from the weather data of monthly weather normals throughout Norway in the form [min value - max a value] and the mean value is displayed under the range for each month.

Variable	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**
Precipitation (mm)	[0 - 489] 92	[0 - 336] 72	[0 - 302] 66	[0 - 294] 61	[1 - 200] 67	[25 - 289] 76
Precipitation Days	[0 - 23] 11	[0 - 18] 10	[0 - 21] 10	[0 - 20] 9	[0 - 17] 10	[3 - 22] 10
Overcast Days	[7 - 21] 15	[7 - 18] 13	[8 - 21] 13	[9 - 18] 12	[2 - 20] 10	[6 - 19] 11
Clear Days	[0 - 7] 3	[0 - 6] 3	[0 - 5] 3	[0 - 6] 2	[0 - 10] 3	[0 - 5] 2
Temperature	[(-15) - 6] (-2)	[(-14) - 3.4] (-2)	[(-11) - 4] 1	[(-17) - 8] 5	[(-2) - 16] 11	[1 - 19] 14

Variable	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Precipitation (mm)	[16 - 201] 80	[42 - 464] 121	[0 - 428] 118	[14 - 505] 118	[0 - 410] 115	[0 - 587] 104
Precipitation Days	[0 - 23]	[0 - 18]	[0 - 21]	[0 - 20]	[0 - 17]	[3 - 22]
Overcast Days	[5 - 21] 11	[6 - 22] 11	[6 - 21] 12	[9 - 20] 15	[8 - 22] 16	[7 - 21] 14
Clear Days	[0 - 8] 2	[0 - 4] 2	[0 - 5] 2	[0 - 6] 2	[0 - 5] 2	[0 - 6] 3
Temperature	[5 - 23] 17	[4 - 17] 15	[1 - 15] 12	[(-4) - 11] 7	[(-8) - 9] 3	[(-12) - 7] (-1)

We hypothesize that the temperature is the most important weather determinant, and gives a clear separation between months and geography, as shown in Table 2.4. Therefore, we differentiate the temperature measurement using the data gathered and transform the temperature data into three new variables. Another explanation of doing this is to have comparable measures to similar research, where some use the January and July measures, while other use the summer/winter averages. In addition we add a third temperature transformation HDD:

- **Summer and Winter temperature:** Defined by dividing the year into two seasons, winter defined as the average temperature from October to March, while summer is defined as the average temperature from April to September (Jacobsen & Røisehagen, 2016).
- **July and January average:** Defined as the average temperature in January and July to describe the most common "winter" month and "summer" month (Maddison & Bigano, 2003).
- **Heating Degree Days (HDD):** HDD is a transformation of temperature to capture the economical effect of temperature and weather on house prices(Eto, 1988). HDD is quantitative measurement for isolating the impact of weather on energy usage in buildings, and especially correlated with the heating and electricity (Do et al., 2016) need for a house. The variable is defined as the yearly sum of HDD values.

Table 2.5 presents the correlation between the temperature variables, and as expected, there is a high correlation between most of them. In addition, the January and winter variables have the highest correlation with HDD.

Table 2.5: Correlation between the temperature variables. Summer is defined as the average temperature in the warmer months of the year, April-October, and winter as the average of the remaining six months.

	HDD	Summer	Winter	January	July
HDD	1	-0.740	-0.871	-0.787	-0.498
Summer	-0.740	1	0.382	0.283	0.937
Winter	-0.871	0.382	1	0.965	0.081
January	-0.787	0.283	0.965	1	-0.008
July	-0.498	0.937	0.081	-0.008	1

Chapter 3

Methodology and Results

The hedonic approach is a widely recognized method for examining price determinants in the real estate proportion of existing literature. The idea is to conceptually break down a property into its various attributes, and estimate the contribution of each to the property's price (Mander et al., 2007). Using this method, we are able to isolate the individual influence of the different weather conditions. The log-linear model is written:

$$\ln(P_{itm}) = \alpha + \lambda_t + \delta_m + \sum_{k=1}^{K} \beta_k X_{ik} + \epsilon_{itm} \tag{3.1}$$

Where the dependent variable P_{itm} is the price of dwelling i, located in municipality m and sold in the time period t. We apply the log-transformation in order to easily interpret the coefficients in terms of average percentages. Furthermore, X is the set of explanatory variables for dwelling i, β is the corresponding coefficients to be estimated and ϵ is the error term. The explanatory variables include the weather condition variables, as well as all dwelling characteristics listed in Table 2.2, because we see these as the most important for our study.

To control for regional variations and price development over time, we specify both quarterly time and municipality fixed effects. Accordingly, the terms λ_t and δ_m in Equation 3.1 are the coefficients for sales period t and municipality m, associated with dwelling i. Finally, standard errors are clustered at the municipality level to control for spatial autocorrelation in prices within municipalities.

While some studies like Maddison & Bigano (2003) utilize wage data to control for economic inequality, we believe the dummies for each municipality are able to capture the effect of differences in wages and job opportunities.

3.1 Results Temperature Transformations

The different temperature transformations are added to the model with all the control variables, and we design models with different combinations of the variables. Thereby, we are able to define which impact the different variables have on price, and which variable to utilize when adding the other climatic variables. The results are shown in Table 3.1.

The results show that all the different temperature transformations are significant when put alone in separate models, and that the magnitude of the coefficients varies. We can see that the HDD in model (1) has the expected sign and that a high HDD value, which implies cold temperatures and an increasing need for energy consumption, are related to a decrease in house prices. While model (2) and (3) displays that a higher temperature in both summer and winter periods are related to higher house price, where the effect is higher for the winter period. In models (5) and (6), the variables representing average temperature in January and July indicate that higher average temperature results in higher price, and the effect is nearly of the same magnitude for both variables. Moreover, in model (4) with the summer and winter temperature put together, both variables seem to be insignificant, which may be due to the high correlation between the variables shown in Table 2.5. While adding the July and January average together in model (7), both are still significant and the impact seems to be quite similar; warmer July and January temperatures indicate a higher house price.

The results show small differences between the temperature transformations, and imply that the use of each variables is almost equally good. In addition, the results show that temperature in the winter matters more, compared to the summer temperature, and is further substantiated by the HDD.

Table 3.1: Results of regression with control variables and three different temperature transformation. Age is transformed in to 1/Age, which indicates that the age matter less for really old houses. We control for different building types buy have a dummy for each type, with the type *Detached* as the baseline

	Dependent variable: log(price per square metre)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	0.181***	0.182***	0.182***	0.182***	0.184***	0.184***	0.185***
	(0.014)	(0.014)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
Rooms	0.008***	0.008***	0.008***	0.008***	0.008***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Size per $100m^2$	−0.286***	−0.287***	−0.287***	−0.287***	−0.287***	−0.287***	−0.287***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Elevator	0.079***	0.078***	0.077***	0.077***	0.077***	0.078***	0.076***
	(0.014)	(0.014)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
Apartment	0.048***	0.045***	0.046***	0.045***	0.044***	0.046***	0.044***
	(0.012)	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.014)
Semi-detatched	−0.006	−0.007	−0.006	−0.006	−0.007	−0.007	−0.007
	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.013)
Serial-house	−0.017	−0.017	−0.016	−0.016	−0.017	−0.016	−0.016
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)
HDD	−0.058***						
	(0.016)						
Average summer temp.		0.044***		0.005			
		(0.010)		(0.034)			
Average winter temp.			0.064**	0.060			
			(0.026)	(0.053)			
Average July temp.					0.048***		0.030**
					(0.015)		(0.012)
Average January temp.						0.051***	0.032***
						(0.014)	(0.011)
Quarterly fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Municipality fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1,610,956	1,610,956	1,610,956	1,610,956	1,610,956	1,610,956	1,610,956
R^2	0.788	0.789	0.790	0.790	0.790	0.790	0.790
Adjusted R^2	0.788	0.789	0.790	0.790	0.790	0.790	0.790
Note:							*p<0.1; **p<0.05; ***p<0.01

3.2 Building Type Specific Results

Considering the characteristic differences among building types, which is also reflected in the coefficients from previous model, we investigate the impact of temperature on price for each dwelling type individually. We further adapt the model presented in Equation 3.1. First, we transform the variable *age* to age group dummies, as we are not able to capture the variations with a continuous variable for some of the dwelling types. Second, we proceed with only the average temperatures for the months January and July. These two variables are, compared to the summer and winter variables, significantly less correlated as shown in Table 2.5, and the estimated coefficients from previous model are more intuitive and significant.

Regression results for each building type are presented in Table 3.2. All of the coefficients for dwelling characteristics appear with expected signs and are significant at one of the three levels, except the age group dummy 10-20 for detached houses. The coefficient for this variable shows a negligible discount compared to the base age group 0-10, indicating that the two groups have similar average effects on dwelling prices. Moreover, the discounts associated with age 40-50 are greater than for 50+ for all building types, which is not hard to believe considering that the older a dwelling is beyond this point, the higher possibility that the dwelling has undergone significant renovation, or has additional value due to age.

More interestingly, temperature is a significant price determinant for both dwelling types, however, the impact is slightly greater for apartments. The latter is counter intuitive, as houses will likely have a greater energy usage compared to apartments in a colder climate. Hence, we expected *HDD* to play a more significant role for houses, but as it seems, people use the temperature not only as a determinant for economical gain, but in a combination with the preference of living in warmer places.

Finally, although the *HDD* varies in magnitude between the building types, it is the temperature transformation which has the greater impact for both types.

Table 3.2: Results of regression with control variables and average temperatures in January and July as indicators for winter and summer.

	Dependent variable: log(price per square metre)							
		Apartments				Houses		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age (10-20)	−0.135***	−0.137***	−0.135***	−0.136***	−0.052***	−0.052***	−0.052***	−0.052***
	(0.023)	(0.021)	(0.021)	(0.021)	(0.006)	(0.006)	(0.006)	(0.006)
Age (20-30)	−0.200***	−0.199***	−0.197***	−0.196***	−0.150***	−0.150***	−0.151***	−0.150***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.009)	(0.009)	(0.009)	(0.008)
Age (30-40)	−0.244***	−0.241***	−0.243***	−0.240***	−0.204***	−0.204***	−0.205***	−0.204***
	(0.018)	(0.018)	(0.019)	(0.019)	(0.010)	(0.010)	(0.010)	(0.010)
Age (40-50)	−0.222***	−0.220***	−0.220***	−0.219***	−0.228***	−0.228***	−0.229***	−0.228***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age (50+)	−0.062***	−0.076***	−0.077***	−0.080***	−0.194***	−0.196***	−0.197***	−0.197***
	(0.022)	(0.018)	(0.020)	(0.019)	(0.031)	(0.031)	(0.030)	(0.030)
Rooms	−0.017***	−0.015***	−0.016***	−0.016***	0.010***	0.010***	0.010***	0.010***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Size per $100m^2$	−0.192***	−0.192***	−0.192***	−0.191***	−0.282***	−0.283***	−0.283***	−0.283***
	(0.033)	(0.034)	(0.032)	(0.033)	(0.016)	(0.015)	(0.016)	(0.015)
Elevator	0.081***	0.077***	0.076***	0.075***	0.200***	0.196***	0.198***	0.196***
	(0.007)	(0.009)	(0.009)	(0.009)	(0.030)	(0.028)	(0.029)	(0.029)
HDD	−0.067***				−0.051***			
	(0.018)				(0.019)			
Average January temp.		0.052***		0.021**		0.043***		0.032**
		(0.010)		(0.009)		(0.016)		(0.015)
Average July temp.			0.059***	0.045**			0.034**	0.021*
			(0.017)	(0.019)			(0.014)	(0.012)
Quarterly fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Municipality fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	810,465	810,465	810,465	810,465	800,491	800,491	800,491	800,491
R^2	0.713	0.715	0.717	0.717	0.766	0.767	0.767	0.767
Adjusted R^2	0.712	0.715	0.716	0.717	0.766	0.767	0.767	0.767

Note: *p<0.1; **p<0.05; ***p<0.01

Chapter 4

Conclusion

Even though there are multiple intricate aspects to determining a dwelling's price, there is substantial evidence of weather being a contributing factor. To determine the implicit price of climate, in our case the temperature, we utilize the hedonic price model. Our results show that in Norway, people pay a price premium for higher temperature in the summer, and even more for higher temperatures in the winter. We analyse the temperature effects even further, and investigate how the impact varies among the building types. The results show differences among building types, where apartments have the highest positive impact from high temperatures. Furthermore, we find that all three temperature transformations are good measures. However, the estimated coefficients for HDD are consistently (and slightly) stronger, meaning it allows us to capture more climatic variations compared to the other temperature measures.

Compared to other research, our findings are similar to what Cragg & Kahn (1999) found, with migration towards warm summers and winters. While others (Rappaport, 2007) found evidence of preference for mild summers, which most likely is due to in some places in the US, the summers are much warmer compared to Norway, but as the global temperature increase, there are possibilities of increased summer temperatures in Norway.

Literature using Norway as the geographical location is lacking, and the master's thesis by Jacobsen & Røisehagen (2016) mainly focused on the weather on the bidding day in one city, Asker. In comparison, we focus on the long term weather effects using weather normals, but we can still draw similar conclusion that warmer weather impacts the short term and long term decision-making.

In this study, we present several different weather variables, but decide to focus on the temperature explicitly. However, given the increased focus on climate change, additional research should include models with additional climatic variables regarding snow, as some regions in Norway has extreme winters. Besides, an important factor to bear in mind when investigating the value of an amenity, is the economic inequality among consumers. We attempt to account for wage differences by controlling for municipalities (Section 3), although there are huge variations even within municipalities, and especially in larger cities. Ideally, controlling for subdivisions of municipalities, or alternatively, zip codes, would perhaps produce more accurate estimates.

Not only can such models presented in this research show the current impact of climate on house prices, but it can be used in forecasting how the house prices will change with respect to increase in global temperature and other climatic conditions in the future. Like Costa & Kahn (2003) showed, the value of weather has increased with time. We may see further increase in the future to come, with increasing extreme weather and temperature all around the world.

References

Butsic, V., Hanak, E., & Valletta, R. G. (2011). Climate change and housing prices: Hedonic estimates for ski resorts in western north america. *Land Economics*, *87*(1), 75–91.

Cheshire, P. C., & Magrini, S. (2006). Population growth in european cities: Weather matters – but only nationally. *Regional Studies*, *40*(1), 23-37.

Costa, D. L., & Kahn, M. E. (2003, May). The rising price of nonmarket goods. *American Economic Review*, *93*(2), 227-232.

Cragg, M. I., & Kahn, M. E. (1999). Climate consumption and climate pricing from 1940 to 1990. *Regional Science and Urban Economics*, *29*(4), 519 - 539.

Do, L. P. C., Lin, K.-H., & Molnár, P. (2016). Electricity consumption modelling: A case of germany. *Economic Modelling*, *55*, 92–101.

Englin, J. (1996). Estimating the amenity value of rainfall. *Ann Reg Sci*, *30*, 273–283.

Eto, J. H. (1988). On using degree-days to account for the effects of weather on annual energy use in office buildings. *Energy and Buildings*, *12*(2), 113 - 127.

Giannakopoulos, C., Bindi, M., Moriondo, M., LeSager, P., & Tin, T. (2005, 07). *Climate change impacts in the mediterranean resulting from a 2°c global temperature rise* (Vol. 1; Tech. Rep.).

Jacobsen, S. K., & Røisehagen, S. (2016). *Boligpriser og vær* (Unpublished master's thesis). Høgskolen i Oslo og Akershus.

Maddison, D., & Bigano, A. (2003). The amenity value of the italian climate. *Journal of Environmental Economics and Management*, *45*(2), 319 - 332.

Mander, Ü., Wiggering, H., & Helming, K. (2007). *Multifunctional land use: Meeting future demands for landscape goods and services.* Springer Berlin Heidelberg.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics*, *37*(3), 375 - 398.

Rehdanz, K. (2006). Hedonic pricing of climate change impacts to households in great britain. *Climatic Change*, *74*, 413–434.

Rehdanz, K., & Maddison, D. (2009, 02). The amenity value of climate to households in germany. *Oxford Economic Papers*, *61*, 150-167.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, *90*(6), 1257–1278.

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